

*Benutzerhandbuch zu IBM SPSS Mo-
deler 18.2.1*

IBM

Hinweis

Vor Verwendung dieser Informationen und des darin beschriebenen Produkts sollten die Informationen unter „Bemerkungen“ auf Seite 259 gelesen werden.

Produktinformation

Diese Ausgabe bezieht sich auf Version 18, Release 2, Modifikation 0 von IBM SPSS Modeler und alle nachfolgenden Releases und Modifikationen, bis dieser Hinweis in einer Neuausgabe geändert wird.

Inhaltsverzeichnis

Kapitel 1. Informationen zu IBM SPSS

Modeler	1
IBM SPSS Modeler-Produkte	1
IBM SPSS Modeler	1
IBM SPSS Modeler Server	1
IBM SPSS Modeler Administration Console	2
IBM SPSS Modeler Batch	2
IBM SPSS Modeler Solution Publisher	2
IBM SPSS Modeler Server-Adapter für IBM SPSS Collaboration and Deployment Services	2
IBM SPSS Modeler-Editionen	2
Dokumentation	3
SPSS Modeler Professional-Dokumentation	3
SPSS Modeler Premium-Dokumentation	4
Anwendungsbeispiele	4
Ordner "Demos"	5
Lizenzüberwachung	5

Kapitel 2. Neue Funktionen in IBM SPSS

Modeler 18.2.1	7
---------------------------------	----------

Kapitel 3. Produktübersicht 9

Erste Schritte	9
Starten von IBM SPSS Modeler	9
Starten über die Befehlszeile	9
Verbinden mit IBM SPSS Modeler Server	10
Verbinden mit Analytic Server	12
Ändern des temporären Verzeichnisses	13
Starten mehrerer IBM SPSS Modeler-Sitzungen	13
Schnittstelle von IBM SPSS Modeler auf einen Blick	14
Streamerstellungsbereich von IBM SPSS Modeler	14
Knotenpalette.	15
IBM SPSS Modeler-Manager.	16
IBM SPSS Modeler-Projekte	18
IBM SPSS Modeler-Symbolleiste	18
Anpassen der Symbolleiste	20
Anpassen des IBM SPSS Modeler-Fensters	20
Ändern der Symbolgröße für einen Stream	21
Verwenden der Maus in IBM SPSS Modeler	22
Verwenden von Tastenkombinationen	22
Drucken	23
Automatisieren von IBM SPSS Modeler	24

Kapitel 4. Data-Mining 25

Data-Mining - Übersicht	25
Bewerten der Daten	26
Strategie für Data-Mining.	27
Prozessmodell "CRISP-DM"	28
Typen von Modellen	29
Data-Mining-Beispiele	35

Kapitel 5. Erstellen von Streams. 37

Streamerstellung - Übersicht.	37
Erstellen von Datenstreams	37

Arbeiten mit Knoten	38
Arbeiten mit Streams	43
Streambeschreibungen.	57
Ausführen von Streams	59
Arbeiten mit Modellen	59
Hinzufügen von Kommentaren und Anmerkungen zu Knoten und Streams	60
Speichern von Datenstreams.	66
Laden von Dateien	67
Zuordnen von Datenstreams.	68
Tipps und Tastenkombinationen	70

Kapitel 6. Arbeiten mit Daten 73

Erstellen von Diagrammen	73
Layout und Begriffe	73
Erstellen eines Diagramms aus der Galerie	74
Diagrammtypen	74
Dashboard.	97
Globale Darstellungsvorgaben	98

Kapitel 7. Arbeiten mit Ausgaben 101

Viewer	101
Ein- und Ausblenden von Ergebnissen	101
Verschieben, Löschen und Kopieren von Ausgaben.	102
Ändern der anfänglichen Ausrichtung	102
Ändern der Ausrichtung von Ausgabeelementen	102
Gliederung für Viewer	102
Hinzufügen von Elementen zum Viewer	103
Suchen und Ersetzen von Informationen im Viewer.	104
Kopieren von Ausgaben in andere Anwendungen	105
Interaktive Ausgabe	106
Ausgabe exportieren	107
HTML-Optionen	108
Webberichtsoptionen	109
Word/RTF-Optionen	109
Excel-Optionen	110
PowerPoint-Optionen.	111
PDF-Optionen	112
Textoptionen.	112
Nur Grafikoptionen	113
Grafikformatoptionen.	113
Drucken von Viewer-Dokumenten	114
Drucken von Ausgaben und Diagrammen.	115
Seitenansicht	115
Seitenattribute: Kopf- und Fußzeilen.	115
Seitenattribute: Optionen	116
Speichern der Ausgabe	116
Speichern eines Viewer-Dokuments	116
Pivot-Tabellen	118
Pivot-Tabellen	118
Bearbeiten von Pivot-Tabellen	118
Arbeiten mit Schichten	121
Ein- und Ausblenden von Elementen	121

Tabellenvorlagen	122
Tabelleneigenschaften.	123
Zelleneigenschaften	125
Fußnoten und Titelzeilen	126
Breite der Datenzellen	128
Ändern der Spaltenbreite	128
Anzeigen der ausgeblendeten Rahmen in einer Pivot-Tabelle	128
Auswählen von Zeilen, Spalten und Zellen in Pivot-Tabellen	128
Drucken von Pivot-Tabellen	129
Erstellen eines Diagramms aus einer Pivot-Tabelle	130
Aus früheren Versionen übernommene Tabellen	130
Optionen	130
Optionen	130
Allgemeine Optionen.	131
Viewer-Optionen	131
Optionen: Pivot-Tabellen	131
Ausgabeoptionen	132

Kapitel 8. Verarbeiten fehlender Werte 133

Fehlende Werte - Übersicht	133
Umgang mit fehlenden Werten	134
Behandeln von Datensätzen mit fehlenden Werten	134
Behandeln von Feldern mit fehlenden Werten	134
Behandeln von Datensätzen mit systemdefiniert fehlenden Werten	135
Imputieren oder Auffüllen fehlender Werte	137
CLEM-Funktionen für fehlende Werte	138

Kapitel 9. Erstellen von CLEM-Ausdrücken 139

Informationen zu CLEM.	139
CLEM-Beispiele	139
Werte und Datentypen	141
Ausdrücke und Bedingungen	142
Stream-, Sitzungs- und Superknotenparameter	142
Arbeiten mit Zeichenfolgen.	143
Umgang mit Leerzeichen und fehlenden Werten	144
Arbeiten mit Zahlen	144
Arbeiten mit Zeit- und Datumsangaben	144
Zusammenfassen mehrerer Felder	145
Arbeiten mit Mehrfachantwortdaten.	146
Expression Builder	147
Zugriff auf Expression Builder.	147
Erstellen von Ausdrücken	147
Auswählen von Funktionen	148
Auswählen von Feldern, Parametern und globalen Variablen	151
Anzeigen oder Auswählen von Werten	151
Überprüfen von CLEM-Ausdrücken	151
Suchen und Ersetzen	152

Kapitel 10. CLEM-Sprachreferenz . . . 155

CLEM-Referenz - Überblick	155
CLEM-Datentypen.	155
Ganze Zahlen	155
Reelle Zahlen	156

Zeichen	156
Zeichenfolgen	156
Listen	156
Felder	157
Datumsangaben	157
Zeit	158
CLEM-Operatoren.	158
Funktionsreferenz	160
Konventionen bei Funktionsbeschreibungen	161
Informationsfunktionen	162
Konvertierungsfunktionen	163
Vergleichsfunktionen	164
Logische Funktionen	166
Numerische Funktionen	167
Trigonometrische Funktionen	168
Wahrscheinlichkeitsfunktionen.	169
Räumliche Funktionen	169
Bitweise Operationen mit ganzen Zahlen	170
Zufallsfunktionen	171
Zeichenfolgelfunktionen	172
SoundEx-Funktionen	177
Datums- und Uhrzeitfunktionen	177
Sequenzfunktionen	182
Globale Funktionen	187
Funktionen zum Umgang mit Leerstellen und Nullwerten	188
Sonderfelder.	189

Kapitel 11. Verwenden von IBM SPSS Modeler mit einem Repository 191

Informationen zu IBM SPSS Collaboration and Deployment Services Repository	191
Speichern und Bereitstellen von Repository-Objekten	192
Verbinden mit dem Repository	192
Eingabe der Berechtigungsnachweise für das Repository	193
Suchen von Repository-Berechtigungsnachweisen	193
Durchsuchen des Repository-Inhalts.	193
Speichern von Objekten im Repository	193
Festlegen der Objekteigenschaften	194
Speichern von Streams	196
Temporäres Speichern von Projekten	196
Speichern von Knoten	197
Speichern von Ausgabeobjekten	197
Speichern von Modellen und Modellpaletten	197
Abrufen von Objekten aus dem Repository	198
Auswählen eines abzurufenden Objekts	199
Auswählen einer Objektversion	199
Suchen nach Objekten im Repository	199
Ändern von Repository-Objekten.	201
Erstellen, Umbenennen und Löschen von Ordnern	201
Sperren und Entsperren von Repository-Objekten	201
Löschen von Repository-Objekten	201
Verwalten der Eigenschaften von Repository-Objekten	202
Anzeigen von Ordneigenschaften	202

Anzeigen und Bearbeiten von Objekteigenschaften	203
Verwalten von Objektversionsbeschriftungen	204
Bereitstellen von Streams	204
Bereitstellungsoptionen für den Stream	205
Scoring-Verzweigung	207

Kapitel 12. Exportieren in externe Anwendungen 211

Exportieren in externe Anwendungen	211
Öffnen eines Streams in IBM SPSS Modeler Advantage	211
Importieren und Exportieren von Modellen als PMML	212
Modelltypen, die PMML unterstützen	213

Kapitel 13. Projekte und Berichte . . . 215

Einführung in Projekte	215
CRISP-DM-Ansicht	215
Klassenansicht	216
Erstellen von Projekten	216
Erstellen eines neuen Projekts	216
Hinzufügen zu einem Projekt	217
Übertragen von Projekten an IBM SPSS Collaboration and Deployment Services Repository	218
Festlegen der Projekteigenschaften	218
Erstellen von Anmerkungen für ein Projekt	218
Objekteigenschaften	219
Schließen von Projekten	220
Generieren von Berichten	220
Speichern und Exportieren von generierten Berichten	221

Kapitel 14. Anpassen von IBM SPSS Modeler 223

Anpassen von IBM SPSS Modeler-Optionen	223
Festlegen von IBM SPSS Modeler-Optionen	223
Systemoptionen	223

Festlegen von Standardverzeichnissen	224
Festlegen von Benutzeroptionen	224
Anpassen der Knotenpalette	232
Anpassen des Palettenmanagers	233
Ändern der Ansicht einer Palettenregisterkarte	235

Kapitel 15. Leistungsaspekte für Streams und Knoten 237

Knotenreihenfolge	237
Knotencaches	238
Leistung: Prozessknoten	239
Leistung: Modellierungsknoten	240
Leistung: CLEM-Ausdrücke	241

Kapitel 16. Eingabehilfen in IBM SPSS Modeler 243

Übersicht über die Eingabehilfen in IBM SPSS Modeler	243
Arten der Unterstützung von Eingabehilfen	243
Zugriffsmöglichkeiten für Sehbehinderte	243
Zugriffsmöglichkeiten für Blinde	244
Tastatureingabehilfen	245
Verwenden eines Sprachausgabeprogramms	253
Tipps zur Verwendung	254
Kompatibilität mit anderer Software	254
JAWS und Java	255
Verwenden von Diagrammen in IBM SPSS Modeler	255

Kapitel 17. Unicode-Unterstützung 257

Unicode-Unterstützung in IBM SPSS Modeler	257
---	-----

Bemerkungen 259

Marken	260
Bedingungen für Produktdokumentation	260

Index 263

Kapitel 1. Informationen zu IBM SPSS Modeler

IBM® SPSS Modeler ist ein Set von Data-Mining-Tools, mit dem Sie auf der Grundlage Ihres Fachwissens schnell und einfach Vorhersagemodelle erstellen und zur Erleichterung der Entscheidungsfindung in die Betriebsabläufe einbinden können. Das Produkt IBM SPSS Modeler, das auf der Grundlage des den Industrienormen entsprechenden Modells CRISP-DM entwickelt wurde, unterstützt den gesamten Data-Mining-Prozess, von den Daten bis hin zu besseren Geschäftsergebnissen.

IBM SPSS Modeler bietet eine Vielzahl von Modellbildungsmethoden, die aus dem maschinellen Lernen, der künstlichen Intelligenz und der Statistik stammen. Mit den in der Modellierungspalette verfügbaren Methoden können Sie aus Ihren Daten neue Informationen ableiten und Vorhersagemodelle erstellen. Jede Methode hat ihre speziellen Stärken und eignet sich besonders für bestimmte Problemtypen.

SPSS Modeler kann als Standalone-Produkt oder als Client in Verbindung mit SPSS Modeler Server erworben werden. Außerdem ist eine Reihe von Zusatzoptionen verfügbar, die in den folgenden Abschnitten kurz zusammengefasst werden. Weitere Informationen finden Sie unter <https://www.ibm.com/analytics/us/en/technology/spss/>.

IBM SPSS Modeler-Produkte

Zur IBM SPSS Modeler-Produktfamilie und der zugehörigen Software gehören folgende Elemente.

- IBM SPSS Modeler
- IBM SPSS Modeler Server
- IBM SPSS Modeler Administration Console (im Lieferumfang von IBM SPSS Deployment Manager enthalten)
- IBM SPSS Modeler Batch
- IBM SPSS Modeler Solution Publisher
- IBM SPSS Modeler Server-Adapter für IBM SPSS Collaboration and Deployment Services

IBM SPSS Modeler

SPSS Modeler ist eine funktionell in sich abgeschlossene Produktversion, die Sie auf Ihrem PC installieren und ausführen können. Sie können SPSS Modeler im lokalen Modus als Standalone-Produkt oder im verteilten Modus zusammen mit IBM SPSS Modeler Server verwenden, um bei Datasets die Leistung zu verbessern.

Mit SPSS Modeler können Sie schnell und intuitiv genaue Vorhersagemodelle erstellen, und das ohne Programmierung. Mithilfe der speziellen grafischen Schnittstelle können Sie den Data-Mining-Prozess auf einfache Weise visualisieren. Mit der Unterstützung der in das Produkt eingebetteten erweiterten Analyseprozesse können Sie zuvor verborgene Muster und Trends in Ihren Daten aufdecken. Sie können Ergebnisse modellieren und Einblick in die Faktoren gewinnen, die Einfluss auf diese Ergebnisse haben, wodurch Sie in die Lage versetzt werden, Geschäftschancen zu nutzen und Risiken zu mindern.

SPSS Modeler ist in zwei Editionen erhältlich: SPSS Modeler Professional und SPSS Modeler Premium. Weitere Informationen finden Sie in „IBM SPSS Modeler-Editionen“ auf Seite 2.

IBM SPSS Modeler Server

SPSS Modeler verwendet eine Client/Server-Architektur zur Verteilung von Anforderungen für ressourcenintensive Vorgänge an leistungsstarke Serversoftware, wodurch bei größeren Datasets eine höhere Leistung erzielt werden kann.

SPSS Modeler Server ist ein separat lizenziertes Produkt, das durchgehend im Modus für verteilte Analysen auf einem Server-Host in Verbindung mit einer oder mehreren IBM SPSS Modeler-Installationen ausgeführt wird. Auf diese Weise bietet SPSS Modeler Server eine herausragende Leistung bei großen Data-sets, da speicherintensive Vorgänge auf dem Server ausgeführt werden können, ohne Daten auf den Client-Computer herunterladen zu müssen. IBM SPSS Modeler Server bietet außerdem Unterstützung für SQL-Optimierung sowie Möglichkeiten zur Modellierung innerhalb der Datenbank, was weitere Vorteile hinsichtlich Leistung und Automatisierung mit sich bringt.

IBM SPSS Modeler Administration Console

Modeler Administration Console ist eine grafische Benutzerschnittstelle zur Verwaltung einer Vielzahl von SPSS Modeler Server-Konfigurationsoptionen, die auch mithilfe einer Optionsdatei konfiguriert werden können. Die Konsole gehört zum Lieferumfang von IBM SPSS Deployment Manager, kann zum Überwachen und Konfigurieren Ihrer SPSS Modeler Server-Installationen verwendet werden und stehen aktuellen SPSS Modeler Server-Kunden kostenlos zur Verfügung. Die Anwendung kann nur unter Windows installiert werden. Der von ihr verwaltete Server kann jedoch auf einer beliebigen unterstützten Plattform installiert sein.

IBM SPSS Modeler Batch

Das Data-Mining ist zwar in der Regel ein interaktiver Vorgang, es ist jedoch auch möglich, SPSS Modeler über eine Befehlszeile auszuführen, ohne dass die grafische Benutzerschnittstelle verwendet werden muss. Beispielsweise kann es sinnvoll sein, langwierige oder sich wiederholende Aufgaben ohne Eingreifen des Benutzers durchzuführen. SPSS Modeler Batch ist eine spezielle Version des Produkts, die die vollständigen Analysefunktionen von SPSS Modeler ohne Zugriff auf die reguläre Benutzerschnittstelle bietet. SPSS Modeler Server ist für die Verwendung von SPSS Modeler Batch erforderlich.

IBM SPSS Modeler Solution Publisher

SPSS Modeler Solution Publisher ist ein Tool, mit dem Sie eine gepackte Version eines SPSS Modeler-Streams erstellen können, der durch eine externe Runtime-Engine ausgeführt oder in eine externe Anwendung eingebettet werden kann. Auf diese Weise können Sie vollständige SPSS Modeler-Streams für die Verwendung in Umgebungen veröffentlichen und bereitstellen, in denen SPSS Modeler nicht installiert ist. SPSS Modeler Solution Publisher wird als Teil des Diensts für IBM SPSS Collaboration and Deployment Services - Scoring verteilt, für den eine separate Lizenz erforderlich ist. Mit dieser Lizenz erhalten Sie SPSS Modeler Solution Publisher Runtime, womit Sie die veröffentlichten Streams ausführen können.

Weitere Informationen zu SPSS Modeler Solution Publisher finden Sie in der Dokumentation zu IBM SPSS Collaboration and Deployment Services. Das IBM SPSS Collaboration and Deployment Services Knowledge Center enthält die Abschnitte "IBM SPSS Modeler Solution Publisher" und "IBM SPSS Analytics Toolkit".

IBM SPSS Modeler Server-Adapter für IBM SPSS Collaboration and Deployment Services

Für IBM SPSS Collaboration and Deployment Services ist eine Reihe von Adaptern verfügbar, mit denen SPSS Modeler und SPSS Modeler Server mit einem Repository von IBM SPSS Collaboration and Deployment Services interagieren können. Auf diese Weise kann ein im Repository bereitgestellter SPSS Modeler-Stream von mehreren Benutzern gemeinsam verwendet werden. Auch der Zugriff über die Thin-Client-Anwendung IBM SPSS Modeler Advantage ist möglich. Sie installieren den Adapter auf dem System, das als Host für das Repository fungiert.

IBM SPSS Modeler-Editionen

SPSS Modeler ist in den folgenden Editionen erhältlich.

SPSS Modeler Professional

SPSS Modeler Professional bietet sämtliche Tools, die Sie für die Arbeit mit den meisten Typen von strukturierten Daten benötigen, beispielsweise in CRM-Systemen erfasste Verhaltensweisen und Interaktionen, demografische Daten, Kaufverhalten und Umsatzdaten.

SPSS Modeler Premium

SPSS Modeler Premium ist ein separat lizenziertes Produkt, das SPSS Modeler Professional für die Arbeit mit spezialisierten Daten sowie für die Arbeit mit unstrukturierten Textdaten erweitert. SPSS Modeler Premium schließt IBM SPSS Modeler Text Analytics ein:

IBM SPSS Modeler Text Analytics verwendet hoch entwickelte linguistische Technologien und die Verarbeitung natürlicher Sprache (Natural Language Processing, NLP), um eine schnelle Verarbeitung einer großen Vielfalt an unstrukturierten Textdaten zu ermöglichen, um die Schlüsselkonzepte zu extrahieren und zu ordnen und um diese Konzepte in Kategorien zusammenzufassen. Extrahierte Konzepte und Kategorien können mit vorhandenen strukturierten Daten, beispielsweise demografischen Informationen, kombiniert und mithilfe der vollständigen Suite der Data-Mining-Tools von IBM SPSS Modeler auf die Modellierung angewendet werden, um bessere und fokussiertere Entscheidungen zu ermöglichen.

IBM SPSS Modeler Subscription

IBM SPSS Modeler Subscription stellt dieselbe Vorhersageanalysefunktionalität bereit wie der konventionelle IBM SPSS Modeler-Client. Mit der Subscription-Edition können Sie regelmäßig Produktaktualisierungen herunterladen.

Dokumentation

Dokumentation ist über das Hilfemenü in SPSS Modeler verfügbar. Dadurch wird das Online-Knowledge Center geöffnet, das stets außerhalb des Produkts verfügbar ist.

Die vollständige Dokumentation für die einzelnen Produkte (einschließlich Installationsanweisungen) ist über den Produktdownload in einem separaten komprimierten Ordner auch im PDF-Format verfügbar. Die neuesten PDF-Dokumente können auch über das Web unter <http://www.ibm.com/support/docview.wss?uid=ibm10874788> heruntergeladen werden.

SPSS Modeler Professional-Dokumentation

Die SPSS Modeler Professional -Dokumentationssuite (ohne Installationsanweisungen) umfasst folgende Dokumente:

- **IBM SPSS Modeler Benutzerhandbuch.** Allgemeine Einführung in die Verwendung von SPSS Modeler, in der u. a. die Erstellung von Datenstreams, der Umgang mit fehlenden Werten, die Erstellung von CLEM-Ausdrücken, das Arbeiten mit Projekten und Berichten und das Packen von Streams für die Bereitstellung von IBM SPSS Collaboration and Deployment Services oder IBM SPSS Modeler Advantage beschrieben werden.
- **IBM SPSS Modeler Quellen-, Prozess- und Ausgabeknoten.** Beschreibung aller Knoten, die zum Lesen, zum Verarbeiten und zur Ausgabe von Daten in verschiedenen Formaten verwendet werden. Im Grunde sind sie alle Knoten, mit Ausnahme der Modellierungsknoten.
- **IBM SPSS Modeler Modellierungsknoten.** Beschreibungen sämtlicher für die Erstellung von Data-Mining-Modellen verwendeter Knoten. IBM SPSS Modeler bietet eine Vielzahl von Modellbildungsmethoden, die aus dem maschinellen Lernen, der künstlichen Intelligenz und der Statistik stammen.
- **IBM SPSS Modeler Anwendungshandbuch.** Die Beispiele in diesem Handbuch bieten eine kurze, gezielte Einführung in bestimmte Modellierungsmethoden und -verfahren. Eine Online-Version dieses Handbuchs kann auch über das Hilfemenü aufgerufen werden. Weitere Informationen finden Sie im Thema „Anwendungsbeispiele“ auf Seite 4.

- **IBM SPSS Modeler Python Handbuch für Scripterstellung und Automatisierung.** Informationen zur Automatisierung des Systems über Python-Scripting, einschließlich der Eigenschaften, die zur Bearbeitung von Knoten und Streams verwendet werden können.
- **IBM SPSS Modeler Bereitstellungshandbuch.** Informationen zum Ausführen von IBM SPSS Modeler-Streams als Schritte bei der Verarbeitung von Jobs im IBM SPSS Deployment Manager.
- **IBM SPSS Modeler CLEF Entwicklerhandbuch.** CLEF bietet die Möglichkeit, Drittanbieterprogramme, wie Datenverarbeitungsroutinen oder Modellierungsalgorithmen, als Knoten in IBM SPSS Modeler zu integrieren.
- **IBM SPSS Modeler Datenbankinternes Mining.** Informationen darüber, wie Sie Ihre Datenbank dazu einsetzen, die Leistung zu verbessern, und wie Sie die Palette der Analysefunktionen über Drittanbieteralgorithmen erweitern.
- **IBM SPSS Modeler Server Verwaltungs- und Leistungshandbuch.** Informationen zur Konfiguration und Verwaltung von IBM SPSS Modeler Server.
- **IBM SPSS Deployment Manager Benutzerhandbuch.** Informationen zur Verwendung der zum Lieferumfang von Deployment Manager gehörenden Benutzerschnittstelle der Administrationskonsole zum Überwachen und Konfigurieren von IBM SPSS Modeler Server.
- **IBM SPSS Modeler CRISP-DM Handbuch.** Schritt-für-Schritt-Anleitung für das Data Mining mit SPSS Modeler unter Verwendung der CRISP-DM-Methode.
- **IBM SPSS Modeler Batch Benutzerhandbuch.** Vollständiges Handbuch für die Verwendung von IBM SPSS Modeler im Stapelmodus, einschließlich Details zur Ausführung des Stapelmodus und zu Befehlszeilenargumenten. Dieses Handbuch steht nur im PDF-Format zur Verfügung.

SPSS Modeler Premium-Dokumentation

Die SPSS Modeler Premium-Dokumentationssuite (ohne Installationsanweisungen) umfasst folgende Dokumente:

- **SPSS Modeler Text Analytics Benutzerhandbuch.** Informationen zur Verwendung von Textanalysen mit SPSS Modeler, unter Behandlung der Text Mining-Knoten, der interaktiven Workbench sowie von Vorlagen und anderen Ressourcen.

Anwendungsbeispiele

Mit den Data-Mining-Tools in SPSS Modeler kann eine große Bandbreite an geschäfts- und unternehmensbezogenen Problemen gelöst werden; die Anwendungsbeispiele dagegen bieten jeweils eine kurze, gezielte Einführung in spezielle Modellierungsmethoden und -verfahren. Die hier verwendeten Datasets sind viel kleiner als die großen Datenbestände, die von einigen Data-Mining-Experten verwaltet werden müssen, die zugrunde liegenden Konzepte und Methoden können jedoch auch auf reale Anwendungen übertragen werden.

Klicken Sie im Menü "Hilfe" in SPSS Modeler auf die Option **Anwendungsbeispiele**, um auf die Beispiele zuzugreifen.

Die Datendateien und Beispielstreams wurden im Ordner Demos, einem Unterordner des Produktinstallationsverzeichnis, installiert. Weitere Informationen finden Sie in „Ordner "Demos"“ auf Seite 5.

Beispiele für die Datenbankmodellierung. Die Beispiele finden Sie im Handbuch *IBM SPSS Modeler Datenbankinternes Mining*.

Scriptbeispiele. Die Beispiele finden Sie im *IBM SPSS Modeler Handbuch für Scripterstellung und Automatisierung*.

Ordner "Demos"

Die in den Anwendungsbeispielen verwendeten Datendateien und Beispielstreams wurden im Ordner Demos, einem Unterordner des Produktinstallationsverzeichnis (beispielsweise C:\Programme\IBM\SPSS\Modeler\Datei > Stream öffnen zugreifen.

Lizenzüberwachung

Bei der Verwendung von SPSS Modeler wird die Lizenznutzung überwacht und in regelmäßigen Intervallen protokolliert. Es werden die Lizenzmetriken *AUTHORIZED_USER* und *CONCURRENT_USER* protokolliert und der Typ der protokollierten Metrik ist von Ihrem Lizenztyp für SPSS Modeler abhängig.

Die erstellten Protokolldateien können vom Produkt IBM License Metric Tool verarbeitet werden, über das Sie Lizenznutzungsberichte generieren können.

Die Lizenzprotokolldateien werden in demselben Verzeichnis erstellt, in dem SPSS Modeler Client-Protokolldateien aufgezeichnet werden (standardmäßig %ALLUSERSPROFILE%/IBM/SPSS/Modeler/<Version>/log).

Kapitel 2. Neue Funktionen in IBM SPSS Modeler 18.2.1

In diesem Release wurden IBM SPSS Modeler die folgenden Funktionen hinzugefügt.

Nachdem Sie mit der rechten Maustaste auf einen Datenknoten geklickt und "Daten anzeigen" ausgewählt haben, werden Sie feststellen, dass der Registerkarte "Diagramme" die folgenden neuen Diagramme hinzugefügt wurden. Das Dashboard wurde ebenfalls neu gestaltet.

- Kreispackung
- Blase
- ThemeRiver
- Zeitdiagramm

Kapitel 3. Produktübersicht

Erste Schritte

Als Data-Mining-Anwendung bietet IBM SPSS Modeler eine strategische Methode zum Auffinden nützlicher Beziehungen in großen Datasets. Im Unterschied zu herkömmlicheren statistischen Methoden müssen Sie zu Beginn nicht unbedingt wissen, wonach Sie suchen. Sie können Ihre Daten durch Anpassen verschiedener Modelle und Überprüfen unterschiedlicher Beziehungen so lange untersuchen, bis Sie auf nützliche Informationen stoßen.

Starten von IBM SPSS Modeler

Klicken Sie zum Starten der Anwendung auf:

Start > [Alle] Programme > IBM SPSS Modeler <Version> > IBM SPSS Modeler <Version>

Nach einigen Sekunden wird das Hauptfenster angezeigt.

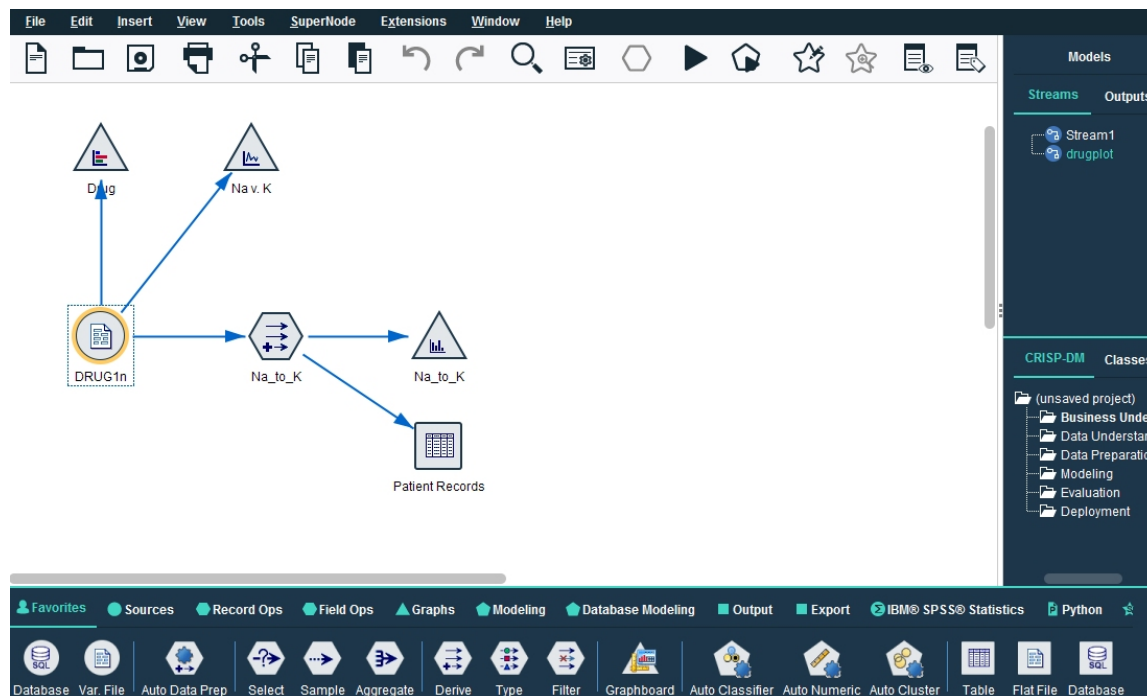


Abbildung 1. Hauptanwendungsfenster von IBM SPSS Modeler

Starten über die Befehlszeile

Sie können IBM SPSS Modeler wie folgt über die Befehlszeile Ihres Betriebssystems starten:

1. Öffnen Sie auf einem Computer, auf dem IBM SPSS Modeler installiert ist, ein DOS- oder Befehlszeilenfenster.
2. Um die IBM SPSS Modeler-Schnittstelle im interaktiven Modus zu starten, geben Sie den Befehl `modelerclient` gefolgt von den erforderlichen Argumenten ein. Beispiel:

```
modelerclient -stream report.str -execute
```

Mithilfe der verfügbaren Argumente (Flags) können Sie eine Verbindung zu einem Server herstellen, Streams laden, Scripts ausführen oder je nach Bedarf weitere Parameter angeben.

Verbinden mit IBM SPSS Modeler Server

IBM SPSS Modeler kann als eigenständige Anwendung oder als Client ausgeführt werden, der direkt mit IBM SPSS Modeler Server oder über das Plug-in Coordinator of Processes von IBM SPSS Collaboration and Deployment Services mit einer IBM SPSS Modeler Server-Instanz oder einem Server-Cluster verbunden ist. Der aktuelle Verbindungsstatus wird unten links im IBM SPSS Modeler-Fenster angezeigt.

Wenn Sie eine Verbindung zu einem Server herstellen möchten, können Sie den Namen des Servers, mit dem eine Verbindung hergestellt werden soll, manuell eingeben oder einen zuvor definierten Namen auswählen. Wenn Sie IBM SPSS Collaboration and Deployment Services verwenden, können Sie im Dialogfeld für die Anmeldung beim Server eine Liste von Servern bzw. Server-Clustern durchsuchen. Die Möglichkeit, die auf einem Netz ausgeführten Statistics-Dienste zu durchsuchen, wird über den Coordinator of Processes bereitgestellt.

So stellen Sie eine Verbindung mit einem Server her:

1. Klicken Sie im Menü "Tools" auf die Option **Anmelden beim Server**. Das Dialogfeld "Anmelden beim Server" wird geöffnet. Alternativ können Sie auf den Bereich des Verbindungsstatus im IBM SPSS Modeler-Fenster doppelklicken.
2. Legen Sie in diesem Dialogfeld die Optionen zum Verbinden mit dem lokalen Server-Computer fest oder wählen Sie eine Verbindung in der Tabelle aus.
 - Klicken Sie auf **Hinzufügen** bzw. **Bearbeiten**, um eine Verbindung hinzuzufügen bzw. zu bearbeiten. Weitere Informationen finden Sie in „Hinzufügen und Bearbeiten der IBM SPSS Modeler Server-Verbindung“ auf Seite 11.
 - Klicken Sie auf **Suche**, um auf einen Server bzw. Server-Cluster in Coordinator of Processes zuzugreifen. Weitere Informationen finden Sie in „Suchen nach Servern in IBM SPSS Collaboration and Deployment Services“ auf Seite 11.

Servertabelle. Diese Tabelle enthält die Menge der definierten Serververbindungen. In der Tabelle werden die Standardverbindung, der Servername sowie die Beschreibung und Portnummer angegeben. Sie können manuell eine neue Verbindung hinzufügen sowie eine bestehende Verbindung auswählen bzw. danach suchen. Um einen bestimmten Server als Standardverbindung einzurichten, aktivieren Sie in der Tabelle für die Verbindung das Kontrollkästchen in der Spalte "Standard".

Standarddatenpfad. Geben Sie einen Pfad an, der für Daten auf dem Server-Computer verwendet wird. Mit der Auslassungsschaltfläche (...) wechseln Sie zum gewünschten Verzeichnis.

Berechtigungsnaachweise festlegen. Lassen Sie dieses Kontrollkästchen unausgewählt, um die Funktion für **Single Sign-on** zu aktivieren. Diese versucht, Sie mithilfe Ihres lokalen Benutzernamens und Kennworts beim Server anzumelden. Falls ein Single Sign-on nicht möglich ist oder Sie das Kontrollkästchen zur Inaktivierung des Single Sign-on aktivieren (z. B. zur Anmeldung an einem Administratorkonto), wird ein weiteres Fenster angezeigt, in dem Sie aufgefordert werden, Ihre Berechtigungsnaachweise einzugeben.

Benutzer-ID. Geben Sie den Benutzernamen ein, mit dem die Anmeldung beim Server erfolgen soll.

Kennwort. Geben Sie das Kennwort ein, das zum angegebenen Benutzernamen gehört.

Domäne. Geben Sie die Domäne an, mit der die Anmeldung beim Server erfolgen soll. Ein Domänenname ist nur dann erforderlich, wenn sich der Server-Computer in einer anderen Windows-Domäne befindet als der Client-Computer.

3. Klicken Sie auf **OK**, um die Verbindung herzustellen.

So trennen Sie eine Verbindung mit einem Server:

1. Klicken Sie im Menü "Tools" auf die Option **Anmelden beim Server**. Das Dialogfeld "Anmelden beim Server" wird geöffnet. Alternativ können Sie auf den Bereich des Verbindungsstatus im IBM SPSS Modeler-Fenster doppelklicken.

2. Wählen Sie im Dialogfeld den lokalen Server aus und klicken Sie auf **OK**.

Hinzufügen und Bearbeiten der IBM SPSS Modeler Server-Verbindung

Serververbindungen können manuell im Dialogfeld "Anmelden beim Server" bearbeitet bzw. hinzugefügt werden. Durch Klicken auf "Hinzufügen" können Sie auf ein leeres Dialogfeld vom Typ "Server hinzufügen/bearbeiten" zugreifen, in dem Sie Details zur Serververbindung eingeben können. Durch Auswahl einer bestehenden Verbindung und Klicken auf "Bearbeiten" im Dialogfeld "Anmelden beim Server" wird das Dialogfeld "Server hinzufügen/bearbeiten" mit den Details für die betreffende Verbindung geöffnet, sodass Sie etwaige Änderungen vornehmen können.

Anmerkung: Sie können eine Serververbindung, die über IBM SPSS Collaboration and Deployment Services hinzugefügt wurde, nicht bearbeiten, da der Name, der Port und andere Details in IBM SPSS Collaboration and Deployment Services definiert sind. Es ist ein bewährtes Verfahren, für die Kommunikation mit IBM SPSS Collaboration and Deployment Services und dem SPSS Modeler-Client dieselben Ports zu verwenden. Diese können als `max_server_port` und `min_server_port` in der Datei `options.cfg` festgelegt werden.

So fügen Sie Serververbindungen hinzu:

1. Klicken Sie im Menü "Tools" auf die Option **Anmelden beim Server**. Das Dialogfeld "Anmelden beim Server" wird geöffnet.
2. Klicken Sie in diesem Dialogfeld auf **Hinzufügen**. Das Dialogfeld "Anmelden beim Server: Server hinzufügen/bearbeiten" wird angezeigt.
3. Geben Sie die Details für die Serververbindung ein und klicken Sie auf **OK**, um die Verbindung zu speichern und zum Dialogfeld "Anmelden beim Server" zurückzukehren.
 - **Server.** Geben Sie einen verfügbaren Server an oder wählen Sie einen aus der Liste aus. Der Server-Computer lässt sich anhand eines alphanumerischen Namens (z. B. *meinserver*) oder der dem Server-Computer zugewiesenen IP-Adresse (z. B. 202.123.456.78) identifizieren.
 - **Port.** Geben Sie die Portnummer an, die der Server überwacht. Wenn der Standardwert nicht funktioniert, fragen Sie Ihren Systemadministrator nach der richtigen Portnummer.
 - **Beschreibung.** Geben Sie eine optionale Beschreibung für diese Serververbindung ein.
 - **Verbindung verschlüsseln (mit SSL).** Legt fest, ob eine SSL-Verbindung (**Secure Sockets Layer**) verwendet werden soll. SSL ist ein weit verbreitetes Protokoll zum Schutz der über ein Netz versendeten Daten. Um diese Funktion verwenden zu können, muss SSL auf dem Server, auf dem sich IBM SPSS Modeler Server befindet, aktiviert sein. Wenden Sie sich gegebenenfalls an den lokalen Administrator, wenn Sie weitere Informationen benötigen.

So bearbeiten Sie Serververbindungen:

1. Klicken Sie im Menü "Tools" auf die Option **Anmelden beim Server**. Das Dialogfeld "Anmelden beim Server" wird geöffnet.
2. Wählen Sie in diesem Dialogfeld die zu bearbeitende Verbindung aus und klicken Sie dann auf **Bearbeiten**. Das Dialogfeld "Anmelden beim Server: Server hinzufügen/bearbeiten" wird angezeigt.
3. Ändern Sie die Details für die Serververbindung und klicken Sie auf **OK**, um die Änderungen zu speichern und zum Dialogfeld "Anmelden beim Server" zurückzukehren.

Suchen nach Servern in IBM SPSS Collaboration and Deployment Services

Anstatt eine Serververbindung manuell einzugeben, können Sie einen im Netz verfügbaren Server oder Server-Cluster über Coordinator of Processes auswählen. Diese Funktion ist in IBM SPSS Collaboration and Deployment Services verfügbar. Ein Server-Cluster ist eine Gruppe von Servern, aus denen Coordinator of Processes den Server ermittelt, der am besten für die Beantwortung einer Verarbeitungsanforderung geeignet ist.

Sie können zwar auch manuell Server im Dialogfeld "Anmelden beim Server" hinzufügen, durch die Suche nach verfügbaren Servern können Sie jedoch eine Verbindung zu Servern herstellen, ohne den richti-

gen Servernamen und die Portnummer kennen zu müssen. Diese Informationen werden automatisch bereitgestellt. Allerdings benötigen Sie auch bei dieser Variante die richtigen Anmeldeinformationen, wie Benutzername, Domäne und Kennwort.

Hinweis: Wenn Sie keinen Zugriff auf die Funktion Coordinator of Processes haben, können Sie dennoch den Namen des Servers, mit dem eine Verbindung hergestellt werden soll, manuell eingeben oder einen zuvor definierten Namen auswählen. Weitere Informationen finden Sie in „Hinzufügen und Bearbeiten der IBM SPSS Modeler Server-Verbindung“ auf Seite 11.

So suchen Sie nach Servern und Clustern:

1. Klicken Sie im Menü "Tools" auf die Option **Anmelden beim Server**. Das Dialogfeld "Anmelden beim Server" wird geöffnet.
2. Klicken Sie in diesem Dialogfeld auf **Suche**, um das Dialogfeld "Nach Servern suchen" zu öffnen. Wenn Sie versuchen, Coordinator of Processes zu durchsuchen, ohne bei IBM SPSS Collaboration and Deployment Services angemeldet zu sein, werden Sie zur Anmeldung aufgefordert.
3. Wählen Sie den Server bzw. Server-Cluster in der Liste aus.
4. Klicken Sie auf **OK**, um das Dialogfeld zu schließen und diese Verbindung der Tabelle im Dialogfeld "Anmelden beim Server" hinzuzufügen.

Verbinden mit Analytic Server

Wenn mehrere Analytic Server-Instanzen verfügbar sind, können Sie das Dialogfeld **Analytic Server-Verbindung** verwenden, um mehr als einen Server für die Verwendung in IBM SPSS Modeler zu definieren. Ihr Administrator hat möglicherweise bereits eine Analytic Server-Standardinstanz in der Datei `<Modeler-Installationspfad>/config/options.cfg` eingerichtet. Sie können jedoch auch andere verfügbare Server verwenden, nachdem Sie sie definiert haben. Wenn Sie beispielsweise die Analytic Server-Knoten "Quelle" und "Export" verwenden, sollten Sie verschiedene Analytic Server-Verbindungen in unterschiedlichen Verzweigungen eines Datenstroms verwenden, sodass jede Verzweigung bei der Ausführung eine eigene Analytic Server-Instanz verwendet und keine Daten in IBM SPSS Modeler Server extrahiert werden. Beachten Sie, dass die Daten aus den Analytic Server-Instanzen in IBM SPSS Modeler Server extrahiert werden, wenn eine Verzweigung mehr als eine Analytic Server-Verbindung enthält. Weitere Informationen einschließlich Einschränkungen finden Sie in „Analytic Server-Streameigenschaften“ auf Seite 50.

Um eine neue Analytic Server-Verbindung zu erstellen, rufen Sie **Tools > Analytic Server-Verbindungen** auf und geben Sie die erforderlichen Informationen in den folgenden Abschnitten des Dialogfelds an.

Verbindung

URL. Geben Sie die URL für Analytic Server im Format `https://Hostname:Port/Kontextstammverzeichnis` ein. Dabei ist `Hostname` die IP-Adresse oder der Hostname von Analytic Server, `Port` die zugehörige Portnummer und `Kontextstammverzeichnis` das Kontextstammverzeichnis von Analytic Server.

Nutzer. Geben Sie den Namen des Nutzers ein, dem IBM SPSS Modeler Server als Mitglied angehört. Wenden Sie sich an Ihren Administrator, wenn Sie den Nutzer nicht kennen.

Authentifizierung

Modus. Wählen Sie einen der folgenden Authentifizierungsmodi aus.

- Für **Benutzername und Kennwort** müssen Sie den Benutzernamen und das Kennwort eingeben.
- Für **Gespeicherter Berechtigungsnachweis** müssen Sie einen Berechtigungsnachweis aus IBM SPSS Collaboration and Deployment Services Repository auswählen.
- Für **Kerberos** müssen Sie den Name des Serviceprinzips und den Pfad zur Konfigurationsdatei eingeben. Wenden Sie sich an Ihren Administrator, wenn Sie diese Informationen nicht kennen.

Benutzername. Geben Sie den Analytic Server-Benutzernamen ein.

Bereiche. Wählen Sie den Bereich aus, der für die Analytic Server-Verbindung verwendet werden soll.

Kennwort. Geben Sie das Analytic Server-Kennwort ein.

Verbinden. Klicken Sie auf **Verbinden**, um die neue Verbindung zu testen.

Verbindungen

Nachdem Sie die obigen Informationen angegeben und auf **Verbinden** geklickt haben, wird die Verbindung in der Tabelle **Verbindungen** hinzugefügt. Wenn Sie eine Verbindung entfernen müssen, wählen Sie diese aus und klicken Sie auf **Entfernen**.

Wenn Ihr Administrator eine Analytic Server-Standardverbindung in der Datei `options.cfg` definiert hat, können Sie auf **Standardverbindung hinzufügen** klicken, um sie auch Ihren verfügbaren Verbindungen hinzuzufügen. Sie werden zur Angabe des Benutzernamens und Kennworts aufgefordert.

Ändern des temporären Verzeichnisses

Für einige von IBM SPSS Modeler Server durchgeführte Operationen müssen möglicherweise temporäre Dateien erstellt werden. IBM SPSS Modeler verwendet standardmäßig das temporäre Systemverzeichnis zum Erstellen von temporären Dateien. Sie können das temporäre Verzeichnis wie folgt ändern:

1. Erstellen Sie ein neues Verzeichnis namens `spss` und ein Unterverzeichnis namens `servertemp`.
2. Bearbeiten Sie die Datei `options.cfg`, die sich im Unterverzeichnis `/config` Ihres IBM SPSS Modeler-Installationsverzeichnisses befindet. Bearbeiten Sie den Parameter `temp_directory` in diesem Feld, damit er wie folgt lautet: `temp_directory, "C:/spss/servertemp"`.
3. Anschließend müssen Sie den IBM SPSS Modeler Server-Dienst neu starten. Dies ist möglich, wenn Sie auf die Registerkarte **Dienste** Ihrer Windows-Systemsteuerung klicken. Halten Sie den Dienst einfach an und starten Sie ihn erneut, um die von Ihnen durchgeführten Änderungen zu aktivieren. Durch das Neustarten des Systems wird auch der Dienst neu gestartet.

Alle temporären Dateien werden nun in dieses neue Verzeichnis geschrieben.

Anmerkung:

- Normale Schrägstriche müssen verwendet werden.
- Die Einstellung `temp_directory` gilt nicht, wenn Evaluierungsstreams über IBM SPSS Collaboration and Deployment Services-Jobs ausgeführt werden. Wenn Sie einen solchen Job ausführen, wird eine temporäre Datei erstellt. Die Datei wird standardmäßig im Installationsverzeichnis des IBM SPSS Modeler-Servers gespeichert. Sie können den Standardordner, in dem die temporären Dateien gespeichert werden, ändern, wenn Sie die Verbindung zum IBM SPSS Modeler-Server in IBM SPSS Modeler erstellen.

Starten mehrerer IBM SPSS Modeler-Sitzungen

Wenn Sie mehr als eine IBM SPSS Modeler-Sitzung gleichzeitig starten müssen, müssen Sie einige Änderungen an Ihren IBM SPSS Modeler- und Windows-Einstellungen vornehmen. Möglicherweise müssen Sie dies z. B. dann tun, wenn Sie über zwei separate Serverlizenzen verfügen und zwei Streams mit zwei verschiedenen Servern von demselben Client-Computer ausführen wollen.

So ermöglichen Sie mehrere IBM SPSS Modeler-Sitzungen:

1. Klicken Sie auf Folgendes:
Start > [Alle] Programme > IBM SPSS Modeler
2. Klicken Sie mit der rechten Maustaste auf die Verknüpfung **IBM SPSS Modeler** (die mit dem Symbol) und wählen Sie **Eigenschaften** aus.
3. Fügen Sie in dem Textfeld **Ziel** an das Ende der Zeichenfolge `-noshare` hinzu.

4. Wählen Sie in Windows Explorer Folgendes aus:
Tools > Ordneroptionen...
5. Wählen Sie auf der Registerkarte **Dateitypen** die Option **IBM SPSS Modeler-Stream** aus und klicken Sie auf **Erweitert**.
6. Wählen Sie im Dialogfeld **Dateityp bearbeiten** die Option **Mit IBM SPSS Modeler öffnen** aus und klicken Sie auf **Bearbeiten**.
7. Fügen Sie im Textfeld **Anwendung für diesen Vorgang** vor dem Argument **-stream** die Angabe **-noshare** hinzu.

Schnittstelle von IBM SPSS Modeler auf einen Blick

An jedem Punkt des Data-Mining-Prozesses können Sie über die benutzerfreundliche Benutzerschnittstelle von IBM SPSS Modeler Ihr spezielles Fachwissen einbringen. Modellierungsalgorithmen, wie Vorhersage, Klassifizierung, Segmentierung und Assoziationserkennung, gewährleisten leistungsstarke und genaue Modelle. Die Modellergebnisse können problemlos angewendet und in Datenbanken, in IBM SPSS Statistics und in einer Vielzahl anderer Anwendungen eingelesen werden.

Die Datenarbeit mit IBM SPSS Modeler besteht aus drei Schritten.

- Zunächst lesen Sie Daten in IBM SPSS Modeler ein.
- Anschließend unterziehen Sie die Daten einer Reihe von Bearbeitungen.
- Schließlich senden Sie die Daten an ein Ziel.

Diese Reihenfolge wird als **Datenstream** bezeichnet, da die Daten Datensatz für Datensatz von der Quelle durch jeden Bearbeitungsschritt zum Ziel fließen, was entweder zu einer Datenausgabe vom Typ "Modell" oder "Typ" führt.



Abbildung 2. Ein einfacher Stream

Streamerstellungsbereich von IBM SPSS Modeler

Der Streamerstellungsbereich ist der größte Bereich des IBM SPSS Modeler-Fensters. Hier erstellen und bearbeiten Sie Datenstreams.

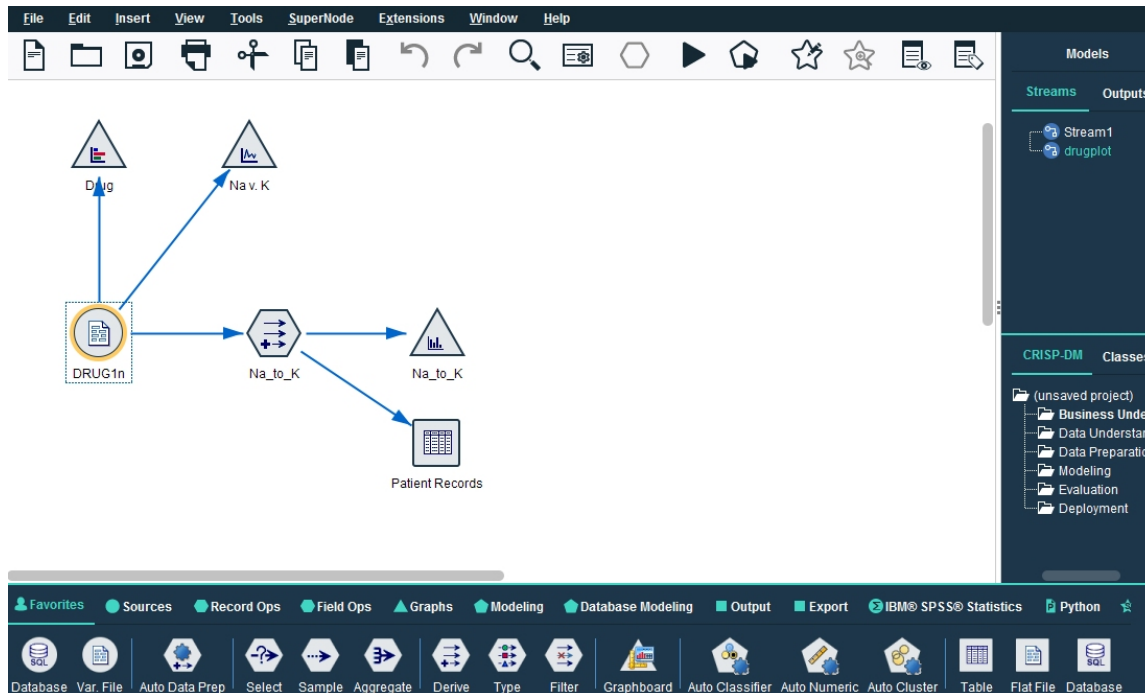


Abbildung 3. IBM SPSS Modeler-Arbeitsbereich (Standardansicht)

Die Erstellung von Streams erfolgt, indem im Haupterstellungsbereich der Benutzerschnittstelle Diagramme gezeichnet werden, in denen die für Ihren Geschäftsbetrieb relevanten Datenoperationen enthalten sind. Jede Operation wird durch ein Symbol oder einen **Knoten** dargestellt und die Knoten sind in einem **Stream** miteinander verbunden, der den Datenfluss durch jede Operation darstellt.

Sie können in IBM SPSS Modeler mit mehreren Streams gleichzeitig arbeiten, entweder in demselben Streamerstellungsbereich oder durch Öffnen eines neuen Streamerstellungsbereichs. Während einer Sitzung werden die Streams im Stream-Manager (rechts oben im IBM SPSS Modeler-Fenster) gespeichert.

Anmerkung: Wenn Sie ein MacBook verwenden, auf dem die Einstellung **Kräftiger Klick und haptisches Feedback** des integrierten Trackpads aktiviert ist, kann das Ziehen und Übergeben von Knoten aus der Knotenpalette in den Streamerstellungsbereich dazu führen, dass dem Erstellungsbereich doppelte Knoten hinzugefügt werden. Um dieses Problem zu vermeiden, wird empfohlen, die Systemvorgabe **Kräftiger Klick und haptisches Feedback** für das Trackpad zu inaktivieren.

Knotenpalette

Die meisten Daten und Modellierungstools in SPSS Modeler sind über die *Knotenpalette* (unten im Fenster) unterhalb des Streamerstellungsbereichs verfügbar.

Die Palettenregisterkarte **Datensatzoperationen** beispielsweise enthält Knoten, mit denen Sie Operationen auf die *Datensätze* anwenden können, wie beispielsweise Auswählen, Zusammenführen (Mergen) und Anhängen.

Wenn Sie Knoten zum Erstellungsbereich hinzufügen wollen, doppelklicken Sie in der Knotenpalette auf die entsprechenden Symbole oder ziehen Sie sie auf den Erstellungsbereich. Anschließend verbinden Sie sie, um einen *Stream* zu erstellen, der den Datenfluss darstellt.

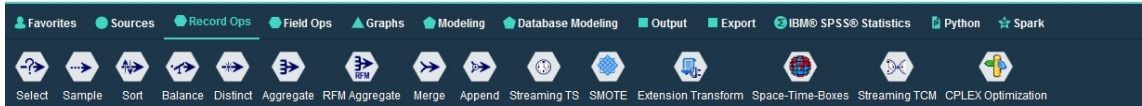


Abbildung 4. Registerkarte "Datensatzoperationen" in der Knotenpalette

Jede Palettenregisterkarte enthält eine Sammlung verwandter Knoten, die für verschiedene Phasen der Streamoperationen verwendet werden, wie:

- **Quellenknoten** lesen Daten in SPSS Modeler ein.
- **Datensatzoperationsknoten** führen Operationen an *Datensätzen* durch, wie beispielsweise Auswählen, Zusammenführen (Mergen) und Anhängen.
- **Feldoperationsknoten** führen Operationen an *Datenfeldern* durch, wie beispielsweise Filtern, Ableiten neuer Felder und Festlegen des Messniveaus für bestimmte Felder.
- **Diagrammknoten** bieten eine grafische Darstellung der Daten vor und nach der Modellierung. Diagramme umfassen Plots, Histogramme, Netzdiagrammknoten und Evaluierungsdiagramme.
- **Modellierungsknoten** verwenden die in SPSS Modeler verfügbaren Modellierungsalgorithmen, wie neuronale Netze, Entscheidungsbäume, Clusteralgorithmen und Datensequenzierung.
- **Datenbankmodellierungsknoten** verwenden die Modellierungsalgorithmen, die in Microsoft SQL Server-, IBM Db2-, Oracle- und Netezza-Datenbanken verfügbar sind.
- **Ausgabeknoten** erzeugen verschiedene Ausgabetyper für Daten, Diagramme und Modellergebnisse, die in SPSS Modeler angezeigt werden können.
- **Exportknoten** erzeugen verschiedene Ausgabetyper, die in externen Anwendungen angezeigt werden können, z. B. in IBM SPSS Data Collection oder Excel.
- **IBM SPSS Statistics-Knoten** importieren Daten aus IBM SPSS Statistics oder exportieren sie nach IBM SPSS Statistics und führen IBM SPSS Statistics-Prozeduren aus.
- **Python-Knoten** können zum Ausführen von Python-Algorithmen verwendet werden.
- **Spark-Knoten** können zum Ausführen von Spark-Algorithmen verwendet werden.

Je vertrauter Sie im Umgang mit SPSS Modeler werden, desto besser können Sie den Paletteninhalt für Ihre eigene Verwendung anpassen.

Links in der Knotenpalette können Sie die angezeigten Knoten filtern, indem Sie **Überwacht**, **Assoziation** oder **Segmentierung** auswählen.

Der Berichtsbereich, der sich unter der Knotenpalette befindet, bietet Feedback zum Fortschritt verschiedener Operationen, z. B. wann die Daten in den Datenstream eingelesen werden. Außerdem bietet der Statusbereich, der sich ebenfalls unterhalb der Knotenpalette befindet, Informationen zur aktuellen Aktivität der Anwendung sowie Anweisungen, wann ein Benutzerfeedback erforderlich ist.

Anmerkung: Wenn Sie ein MacBook verwenden, auf dem die Einstellung **Kräftiger Klick und haptisches Feedback** des integrierten Trackpads aktiviert ist, kann das Ziehen und Übergeben von Knoten aus der Knotenpalette in den Streamerstellungsbereich dazu führen, dass dem Erstellungsbereich doppelte Knoten hinzugefügt werden. Um dieses Problem zu vermeiden, wird empfohlen, die Systemvorgabe **Kräftiger Klick und haptisches Feedback** für das Trackpad zu inaktivieren.

IBM SPSS Modeler-Manager

Oben rechts im Fenster befindet sich der Managerbereich. Dieser enthält drei Registerkarten, die zum Verwalten von Streams, Ausgaben und Modellen verwendet werden.

Sie können die Registerkarte "Streams" verwenden, um die in einer Sitzung erstellten Streams zu öffnen, umzubenennen, zu speichern und zu löschen.



Abbildung 5. Registerkarte "Streams"

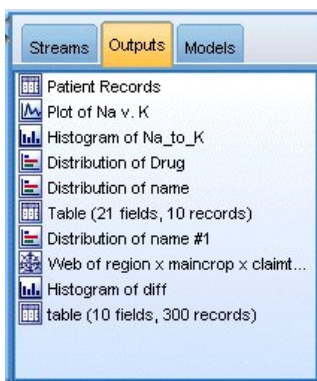


Abbildung 6. Registerkarte "Ausgaben"

Die Registerkarte "Ausgaben" enthält eine Reihe von Dateien, wie beispielsweise Diagramme und Tabellen, die von den Streamoperationen in IBM SPSS Modeler erstellt wurden. Sie können die auf dieser Registerkarte aufgeführten Tabellen, Diagramme und Berichte anzeigen, speichern, umbenennen und schließen.



Abbildung 7. Registerkarte "Modelle" mit Modellnuggets

Die Registerkarte "Modelle" ist die leistungsstärkste der Manager-Registerkarten. Diese Registerkarte enthält sämtliche **Modellnuggets**, die die in IBM SPSS Modeler erstellten Modelle enthalten, für die aktuelle Sitzung. Auf der Registerkarte "Modelle" können diese Modelle direkt durchsucht oder dem Stream im Erstellungsbereich hinzugefügt werden.

IBM SPSS Modeler-Projekte

Unten rechts im Fenster befindet sich der Projektbereich, der zum Erstellen und Verwalten von Data-Mining-**Projekten** verwendet wird (Gruppen von Dateien, die in Bezug zu einer Data-Mining-Aufgabe stehen). Es gibt zwei Methoden zur Ansicht von in IBM SPSS Modeler erstellten Projekten - die Ansicht "Klassen" und die Ansicht "CRISP-DM".

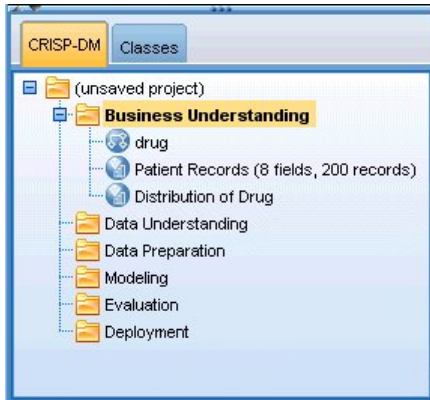


Abbildung 8. Ansicht "CRISP-DM"

Die Registerkarte "CRISP-DM" bietet ein Verfahren zur Organisation von Projekten gemäß dem Cross-Industry Standard Process for Data Mining, einer in der Branche bewährten, nicht eigentumsrechtlich geschützten Methode. Sowohl erfahrene Daten-Mining-Experten als auch Neulinge können mit dem CRISP-DM-Tool ihre Arbeit besser organisieren und an andere weitergeben.

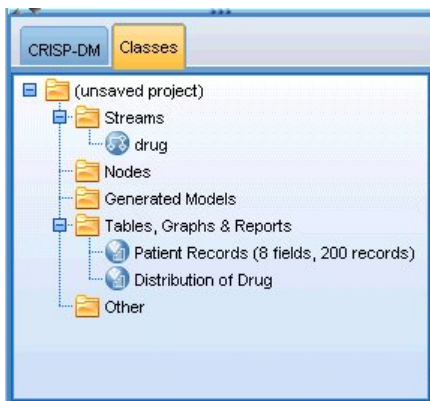


Abbildung 9. Ansicht "Klassen"

Die Registerkarte "Klassen" bietet eine Methode, mit der Sie Ihre Arbeit in IBM SPSS Modeler (nach den erstellten Objekttypen) in Kategorien organisieren können. Diese Ansicht ist hilfreich für das Inventarisieren von Daten, Streams und Modellen.

IBM SPSS Modeler-Symbolleiste

Oben im IBM SPSS Modeler-Fenster sehen Sie eine Symbolleiste, die eine Vielzahl nützlicher Funktionen bietet. Im Folgenden werden die Schaltflächen der Symbolleiste und ihre Funktionen beschrieben.



Neuen Stream erstellen



Stream öffnen



Stream speichern



Aktuellen Stream drucken



Ausschneiden & in die Zwischenablage verschieben



In Zwischenablage kopieren



Auswahl einfügen



Letzte Aktion rückgängig machen



Wiederholen



Knoten suchen



Streameigenschaften bearbeiten



Vorschau für SQL-Generierung



Aktuellen Stream ausführen



Streamauswahl ausführen



Stream anhalten (wird nur während der Streamausführung aktiv)



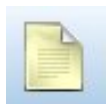
Superknoten hinzufügen



Vergrößern (nur Superknoten)



Verkleinern (nur Superknoten)



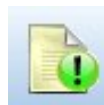
Kein Markup in Stream



Kommentar einfügen



Stream-Markup ausblenden (falls vorhanden)



Ausgeblendeten Stream-Markup einblenden



Stream in IBM SPSS Modeler Advantage öffnen

Stream-Markup umfasst Streamkommentare, Modellverknüpfungen und die Anzeige von Scoring-Verzweigungen.

Eine Beschreibung zu Modellverknüpfungen finden Sie im Handbuch *IBM SPSS Modellierungsknoten*.

Anpassen der Symbolleiste

Sie können zahlreiche Aspekte der Symbolleiste ändern, z. B.:

- ob die Symbolleiste angezeigt wird
- ob für die Symbole eine QuickInfo verfügbar ist
- ob große oder kleine Symbole angezeigt werden

So schalten Sie die Anzeige der Symbolleiste ein bzw. aus:

1. Klicken Sie im Hauptmenü auf Folgendes:

Ansicht > Symbolleiste > Anzeigen

So ändern Sie die Einstellungen für QuickInfo oder Symbolgröße:

1. Klicken Sie im Hauptmenü auf Folgendes:

Ansicht > Symbolleiste > Anpassen

Klicken Sie wie gewünscht auf **QuickInfos einblenden** oder **Große Symbole**.

Anpassen des IBM SPSS Modeler-Fensters

Mit den Trennlinien zwischen den verschiedenen Bereichen der SPSS Modeler-Benutzerschnittstelle können Sie je nach Bedarf die Größe von Tools ändern oder Tools schließen. Wenn Sie z. B. mit einem großen Stream arbeiten, können Sie die kleinen Pfeile, die sich auf den Teilern befinden, zum Schließen der Knotenpalette, des Managerbereichs und des Projektbereichs verwenden. So wird die Größe des Streamerstellungsbereichs vergrößert und bietet somit genügend Arbeitsfläche für große Streams oder mehrere Streams.

Alternativ können Sie auch im Menü "Ansicht" auf **Knotenpalette**, **Manager** oder **Projekt** klicken, um die Anzeige dieser Elemente ein- oder auszuschalten.

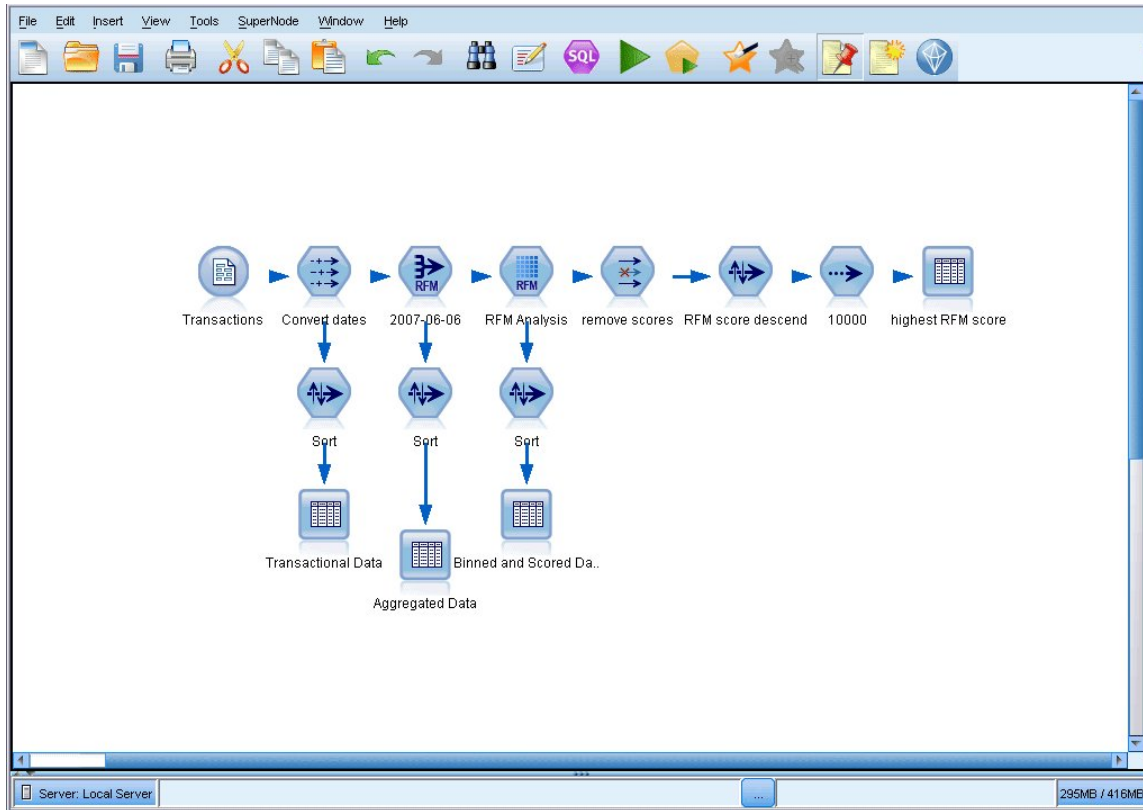


Abbildung 10. Maximierter Streamerstellungsbereich

Als Alternative zum Schließen der Knotenpalette und des Manager- und Projektbereichs können Sie den Streamerstellungsbereich als verschiebbare Seite verwenden, indem Sie die Bildlaufleisten an der Seite und unten im SPSS Modeler-Fenster vertikal und horizontal bewegen.

Sie können auch die Anzeige der Anzeigemarkup steuern, die Streamkommentare, Modellverknüpfungen und die Anzeige von Scoring-Verzweigungen umfasst. Um diese Anzeige ein- oder auszublenden, klicken Sie auf:

Ansicht > Stream-Markup

Ändern der Symbolgröße für einen Stream

Sie können die Größe der Streamsymbole auf folgende Weisen ändern:

- Über eine Einstellung für Streameigenschaften
- Über ein Popup-Menü im Stream
- Über die Tastatur

Sie können die gesamte Streamansicht auf eine Anzahl von Größenwerten zwischen 8 und 200 % der Standardsymbolgröße skalieren.

Skalierung des gesamten Streams (Methode mit Streameigenschaften)

1. Wählen Sie im Hauptmenü Folgendes aus:
Tools > Streameigenschaften > Optionen > Layout.
2. Wählen Sie im Menü "Symbolgröße" die gewünschte Größe aus.
3. Klicken Sie auf **Anwenden**, um das Ergebnis anzuzeigen.
4. Klicken Sie auf **OK**, um die Änderung zu speichern.

Skalierung des gesamten Streams (Methode über Menü)

1. Klicken Sie mit der rechten Maustaste auf den Streamhintergrund im Erstellungsbereich.
2. Wählen Sie die Option **Symbolgröße** und anschließend die gewünschte Größe aus.

Skalierung des gesamten Streams (Methode über Tastatur)

1. Drücken Sie auf der Haupttastatur die Tastenkombination "Strg+[-]", um auf die nächstkleinere Größe umzuschalten.
2. Drücken Sie auf der Haupttastatur die Tastenkombination "Strg+Umschalt+[+]", um auf die nächstgrößere Größe umzuschalten.

Diese Funktion ist besonders nützlich, um einen allgemeinen Überblick über einen komplexen Stream zu erhalten. Sie können damit auch die für das Drucken eines Streams erforderliche Anzahl an Seiten minimieren.

Verwenden der Maus in IBM SPSS Modeler

Am häufigsten wird die Maus in IBM SPSS Modeler wie folgt verwendet:

- **Einmalklicken.** Verwenden Sie entweder die rechte oder die linke Maustaste, um Optionen in den Menüs auszuwählen, Popup-Menüs zu öffnen und verschiedene andere Standardsteuerelemente und Optionen zu verwenden. Klicken Sie und halten Sie die Maustaste gedrückt, um Knoten zu verschieben und zu ziehen.
- **Doppelklicken.** Doppelklicken Sie mit der linken Maustaste, um Knoten auf dem Streamerstellungsbereich abzulegen und die bereits vorhandenen Knoten zu bearbeiten.
- **Klicken mit der mittleren Maustaste.** Klicken Sie auf die mittlere Maustaste und ziehen Sie den Cursor, um Knoten im Streamerstellungsbereich miteinander zu verbinden. Doppelklicken Sie auf die mittlere Maustaste, um die Verbindung eines Knotens zu lösen. Wenn Sie nicht über eine Maus mit drei Tasten verfügen, können Sie diese Funktion simulieren, indem Sie die Taste Alt drücken und gleichzeitig mit der Maus klicken und ziehen.

Verwenden von Tastenkombinationen

Für viele visuelle Programmieroperationen in IBM SPSS Modeler gibt es Tastenkombinationen. Sie können z. B. einen Knoten löschen, indem Sie darauf klicken und die Löschtaste auf der Tastatur drücken. Auf ähnliche Weise können Sie einen Stream speichern, indem Sie auf die Taste "S" und gleichzeitig die Steuertaste (Taste Strg) drücken. Steuerbefehle wie dieser werden durch eine Kombination aus Strg und einer anderen Taste, z. B. Strg+S, angezeigt.

Es gibt eine Reihe von Tastenkombinationen, die für Windows-Standardoperationen verwendet werden, z. B. Strg+X zum Ausschneiden. Diese Tastenkombinationen werden in IBM SPSS Modeler zusammen mit den folgenden anwendungsspezifischen Tastenkombinationen unterstützt.

Anmerkung: In manchen Fällen stehen alte, in IBM SPSS Modeler verwendete Zugriffstasten mit den Windows-Standardzugriffstasten in Konflikt. Diese alten Zugriffstasten werden durch Hinzufügen der Taste Alt unterstützt. So kann z. B. Strg+Alt+C zum Aktivieren und Inaktivieren des Cache verwendet werden.

Tabelle 1. Unterstützte Tastenkombinationen

Tastenkombination	Funktion
Strg+A	Alles markieren
Strg+X	Ausschneiden
Strg+N	Neuer Stream
Strg+O	Stream öffnen
Strg+P	Drucken

Tabelle 1. Unterstützte Tastenkombinationen (Forts.)

Tastenkombination	Funktion
Strg+C	Kopieren
Strg+V	Einfügen
Strg+Z	Rückgängig
Strg+Q	Alle Knoten unterhalb des ausgewählten Knotens auswählen
Strg+W	Auswahl aller nachgeordneten Knoten aufheben (wechselt mit Strg+Q)
Strg+E	Vom ausgewählten Knoten ausführen
Strg+S	Aktuellen Stream speichern
Alt+Pfeiltasten	Ausgewählte Knoten im Streamerstellungsbereich in die Richtung verschieben, in die die verwendete Pfeiltaste zeigt
Umschalt+F10	Popup-Menü für den ausgewählten Knoten öffnen

Tabelle 2. Unterstützte Tastenkombinationen für alte Hot Keys

Tastenkombination	Funktion
Strg+Alt+D	Knoten duplizieren
Strg+Alt+L	Knoten laden
Strg+Alt+R	Knoten umbenennen
Strg+Alt+U	Benutzereingabeknoten erstellen
Strg+Alt+C	Cache aktivieren/inaktivieren
Strg+Alt+F	Cache löschen
Strg+Alt+X	Superknoten erweitern
Strg+Alt+Z	Vergrößern/Verkleinern
Entf	Knoten oder Verbindung löschen

Drucken

Folgende Objekte können in IBM SPSS Modeler gedruckt werden:

- Streamdiagramme
- Grafiken
- Tabellen
- Berichte (über den Berichtsknoten und die Projektberichte)
- Scripts (über die Dialogfelder "Streameigenschaften", "Standalone-Script" oder "Superknotenscript")
- Modelle (Modellbrowser, Registerkarten des Dialogfelds mit aktuellem Fokus, Baumansichten)
- Anmerkungen (unter Verwendung der Registerkarte "Anmerkungen" für die Ausgabe)

So drucken Sie ein Objekt:

- Um ohne Vorschau zu drucken, klicken Sie auf die Schaltfläche "Drucken" in der Symbolleiste.
- Um die Seite vor dem Drucken einzurichten, wählen Sie im Menü "Datei" die Option **Seite einrichten** aus.
- Um vor dem Drucken eine Vorschau anzuzeigen, wählen Sie im Menü "Datei" die Option **Druckvorschau** aus.
- Um das Standarddialogfeld zum Drucken mit Optionen zur Auswahl von Druckern anzuzeigen und die Optionen für das Aussehen festzulegen, wählen Sie im Menü "Datei" die Option **Drucken** aus.

Automatisieren von IBM SPSS Modeler

Da es sich beim erweiterten Data-Mining um einen komplexen und manchmal langwierigen Vorgang handeln kann, bietet IBM SPSS Modeler mehrere Codierungsmöglichkeiten und Automatisierungsunterstützung.

- **Control Language for Expression Manipulation (CLEM)** ist eine Sprache zur Analyse und Bearbeitung der Daten, die IBM SPSS Modeler-Streams durchlaufen. Data-Mining-Experten verwenden CLEM in hohem Maße für Streamoperationen, um so einfache Aufgaben durchzuführen, wie das Ableiten von Profit aus Kosten- und Einkommensdaten, oder um so komplexe Aufgaben durchzuführen, wie das Umwandeln von Webprotokolldaten in ein Set von Feldern und Datensätzen mit nützlichen Informationen.
- Die **Scripterstellung** bildet ein leistungsstarkes Tool, mit dem Prozesse in der Benutzerschnittstelle automatisiert werden. Mit Scripts können dieselben Aktionen durchgeführt werden, die die Benutzer mithilfe von Maus und Tastatur durchführen. Sie können auch die Ausgabe angeben und generierte Modelle bearbeiten.

Kapitel 4. Data-Mining

Data-Mining - Übersicht

Mittels unterschiedlicher Verfahren werden beim **Data-Mining** "Informationsnuggets" im Datenhauptteil identifiziert. Beim Data-Mining werden Informationen so extrahiert, dass sie in Bereichen wie Entscheidungsunterstützung, Vorhersagen, Vorhersagen und Schätzungen verwendet werden können. Daten sind häufig sehr umfangreich, jedoch von geringem Wert und in ihrer Rohform von geringem praktischem Nutzen. Wertvoll sind die in den Daten verborgenen Informationen.

Beim Data-Mining beruht der Erfolg auf einer Kombination Ihres Wissens (oder das Ihres Experten) über die Daten und fortgeschrittener, aktiver Analyseverfahren, bei denen der Computer die zugrunde liegenden Beziehungen und Funktionen in den Daten ermittelt. Beim Data-Mining werden aus historischen Daten Modelle erstellt, die anschließend für Vorhersagen, Mustererkennung usw. verwendet werden. Das für die Erstellung dieser Modelle verwendete Verfahren wird **Maschinelernen** bzw. **Modellierung** genannt.

Modellierungsverfahren

IBM SPSS Modeler beinhaltet eine Reihe von Technologien für Maschinelernen und Modellieren, die grob anhand des Typs von Problemen unterteilt werden können, die sie lösen sollen.

- Methoden für Vorhersagemodelle enthalten Entscheidungsbäume, neuronale Netze und statistische Modelle.
- Clustering-Modelle konzentrieren sich auf die Ermittlung ähnlicher Datensätze und auf die Beschriftung der Datensätze anhand der Gruppe, in die sie gehören. Zu den Clustermethoden gehören "Kohonen", "k-Means" und "TwoStep".
- Assoziationsregeln ordnen eine bestimmte Schlussfolgerung (beispielsweise den Kauf eines bestimmten Produkts) einer Menge von Bedingungen (dem Kauf mehrerer anderer Produkte) zu.
- Screening-Modelle können zum Auffinden von Feldern und Datensätzen verwendet werden, die höchstwahrscheinlich bei der Modellierung von Interesse sind oder zur Identifikation von Ausreißern, die keinem bekannten Muster entsprechen. Zu den verfügbaren Methoden gehören die Merkmalauswahl und die Anomalieerkennung.

Datenbearbeitung und -erkennung

IBM SPSS Modeler beinhaltet außerdem viele Funktionen, mit denen Sie Ihr Fachwissen auf die Daten anwenden können:

- **Datenbearbeitung.** Sie können neue Datenelemente erstellen, die aus bestehenden Daten abgeleitet sind, und die Daten in sinnvolle Subsets einteilen. Daten aus verschiedenen Quellen können zusammengeführt und gefiltert werden.
- **Durchsuchen und Visualisierung.** Sie können Aspekte der Daten mithilfe des Data Audit-Knotens anzeigen, um ein ursprüngliches Audit einschließlich Grafiken und Statistiken durchzuführen. Zur erweiterten Visualisierung gehören interaktive Grafiken, die zur Aufnahme in Projektberichte exportiert werden können.
- **Statistik.** Mit Statistiken können Sie vermutete Beziehungen zwischen den Variablen in den Daten überprüfen. Die Statistikfunktionen aus IBM SPSS Statistics können auch in IBM SPSS Modeler verwendet werden.
- **Hypothesentests.** Sie können Modelle für das Verhalten der Daten erstellen und überprüfen.

Normalerweise verwenden Sie diese Funktionen zur Ermittlung eines erfolgversprechenden Sets von Attributen in den Daten. Diese Attribute können anschließend in die Modellierungsverfahren eingespeist werden, die wiederum versuchen, etwaige zugrunde liegende Regeln und Beziehungen zu ermitteln.

Typische Anwendungen

Beispiele für typische Anwendungsgebiete für Data-Mining-Verfahren:

Direktmarketing. Ermitteln Sie, welche demografischen Gruppen die höchste Rücklaufquote aufweisen. Mit diesen Informationen können Sie die Reaktionen auf zukünftige Anschreiben erhöhen.

Kreditscoring. Verwenden Sie die Kreditgeschichte einer Person als Grundlage für Kreditvergabeentscheidungen.

Humanressourcen. Gewinnen Sie einen Einblick in frühere Einstellungspraktiken und ermitteln Sie Entscheidungsrichtlinien zur Rationalisierung des Einstellungsprozesses.

Medizinische Forschung. Erstellen Sie auf medizinischen Erkenntnissen beruhende Entscheidungsregeln für die geeigneten Verfahren.

Marktanalyse. Ermitteln Sie, welche Variablen, wie Standort, Preis und Kundenmerkmale Einfluss auf den Absatz haben.

Qualitätskontrolle. Analysieren Sie Daten aus der Produktherstellung und ermitteln Sie Variablen, die entscheidenden Einfluss auf Produktmängel haben.

Grundsatzuntersuchungen. Verwenden Sie Umfragedaten zur Erarbeitung einer generellen Vorgehensweise, indem Sie anhand von Entscheidungsregeln die bedeutsamsten Variablen auswählen.

Gesundheitsfürsorge. Benutzerumfragen und klinische Daten können kombiniert werden, um Variablen zu ermitteln, die zur Gesundheit beitragen.

Terminologie

Die Ausdrücke **Attribut**, **Feld** und **Variable** beziehen sich auf ein einzelnes Datenelement, das alle betrachteten Fälle gemeinsam haben. Eine Sammlung von Attributwerten, die sich auf einen bestimmten Fall bezieht, heißt **Datensatz**, **Beispiel** oder **Fall**.

Bewerten der Daten

Data-Mining kann kaum erfolgreich sein, wenn die Daten, die Sie verwenden möchten, nicht bestimmten Kriterien entsprechen. In folgenden Abschnitten finden Sie einige Aspekte der Daten und ihrer Anwendung, die Sie beachten sollten.

Stellen Sie sicher, dass die Daten verfügbar sind.

Das mag trivial erscheinen, Sie sollten sich jedoch darüber im Klaren sein, dass möglicherweise zwar Daten vorliegen, jedoch eventuell nicht in einer problemlos verwendbaren Form. Daten können von IBM SPSS Modeler aus Datenbanken (über ODBC) oder aus Dateien importiert werden. Die Daten können jedoch in einem anderen Format auf einem System vorliegen, auf das kein direkter Zugriff möglich ist. Sie müssen heruntergeladen oder in einem geeigneten Format exportiert werden, bevor sie verwendet werden können. Möglicherweise sind sie auf verschiedene Datenbanken und Quellen verteilt und müssen zusammgebracht werden. Möglicherweise sind sie nicht einmal digital vorhanden. Wenn sie nur auf Papier vorliegen, ist eine Dateneingabe erforderlich, bevor Sie mit dem Data-Mining beginnen können.

Überprüfen Sie, ob die Daten die relevanten Attribute abdecken.

Ziel des Data-Mining ist die Ermittlung relevanter Attribute, sodass diese Überprüfung zunächst merkwürdig erscheinen mag. Es ist jedoch sinnvoll zu untersuchen, welche Daten verfügbar sind, und zu versuchen, die vermutlich relevanten Faktoren zu ermitteln, die nicht aufgezeichnet sind. Beim Versuch, den

Absatz an Speiseeis vorherzusagen, liegen Ihnen eventuell umfassende Informationen zu Verkaufsstellen und frühere Absatzdaten vor, möglicherweise jedoch keine Informationen zu Wetter und Temperatur, die mit großer Wahrscheinlichkeit eine signifikante Rolle spielen. Fehlende Attribute bedeuten nicht unbedingt, dass das Data-Mining keine nützlichen Ergebnisse hervorbringt, die Genauigkeit der resultierenden Vorhersagen kann jedoch eingeschränkt sein.

Eine schnelle Möglichkeit zur Einschätzung der Situation besteht in der Durchführung eines umfassenden Audits der vorliegenden Daten. Bevor Sie fortfahren, sollten Sie der Datenquelle einen Data Audit-Knoten hinzufügen und ihn ausführen, um einen vollständigen Bericht zu generieren.

Vorsicht bei verrauschten Daten.

Daten enthalten häufig Fehler oder subjektive (und deshalb variable) Urteile. Diese Phänomene werden als **Rauschen** bezeichnet. Manchmal ist Rauschen in den Daten normal. Es können durchaus zugrunde liegende Regeln vorhanden sein, diese gelten jedoch eventuell nicht für 100 % der Fälle.

Je mehr Rauschen in den Daten vorliegt, desto schwieriger ist es in der Regel, genaue Ergebnisse zu erzielen. Die Methoden zum Maschinenlernen in IBM SPSS Modeler können jedoch mit verrauschten Daten umgehen und wurden erfolgreich bei Datensets eingesetzt, die annähernd 50 % Rauschen enthielten.

Stellen Sie sicher, dass ausreichend Daten vorliegen.

Beim Data-Mining ist nicht unbedingt die Größe des Datensets wichtig. Wesentlich bedeutsamer sind die Repräsentativität des Datensets sowie der Umfang, in dem es mögliche Ergebnisse und Variablenkombinationen abdeckt.

Je mehr Attribute betrachtet werden, desto mehr Datensätze sind in der Regel erforderlich, um eine repräsentative Abdeckung zu erzielen.

Wenn die Daten repräsentativ sind und allgemeine Regeln zugrunde liegen, ist es durchaus möglich, dass eine Datenstichprobe von wenigen tausend (oder sogar wenigen hundert) Datensätzen zu ebenso guten Ergebnissen führen wie eine Million Datensätze. Außerdem erhalten Sie die Ergebnisse schneller.

Kontaktieren Sie diejenigen, die sich mit den Daten auskennen.

In vielen Fällen arbeiten Sie mit Ihren eigenen Daten und ihr Inhalt und ihre Bedeutung sind Ihnen daher bestens bekannt. Wenn Sie jedoch mit Daten arbeiten, die aus einer anderen Abteilung in Ihrem Unternehmen oder von einem Kunden stammen, sollten Ihnen unbedingt Experten zur Verfügung stehen, die die Daten kennen. Sie können Sie in die Identifizierung relevanter Attribute einweisen und Ihnen helfen, die Ergebnisse des Data-Mining zu interpretieren und die echten "Informationsnuggets" von "Katzengold" bzw. Artefakten zu unterscheiden, die durch Anomalien in den Datensets verursacht wurden.

Strategie für Data-Mining

Wie bei den meisten geschäftlichen Aktivitäten ist auch Data-Mining am effektivsten, wenn es auf geplante, systematische Weise durchgeführt wird. Selbst bei führenden Data-Mining-Tools wie IBM SPSS Modeler ist für den größten Teil der Arbeit beim Data-Mining ein sachkundiger Unternehmensanalyst erforderlich, damit der Prozess in zielgerichteten Bahnen abläuft. Bei Ihrer Planung sollten Sie sich folgende Fragen stellen:

- Welches inhaltliche Problem soll gelöst werden?
- Welche Datenquellen sind verfügbar und welche Teile der Daten sind für das aktuelle Problem relevant?
- Welche Art der Vorverarbeitung und Datenbereinigung ist erforderlich, bevor mit dem Data-Mining der Daten begonnen werden kann?
- Welche Data-Mining-Verfahren sollen eingesetzt werden?

- Wie sollen die Ergebnisse der Data-Mining-Analyse evaluiert werden?
- Wie können Sie die aus dem Data-Mining gewonnenen Informationen möglichst optimal nutzen?

Der übliche Data-Mining-Vorgang kann sehr schnell ziemlich kompliziert werden. Sehr viele Dinge müssen ständig bedacht werden: komplexe Geschäftsprobleme, mehrere Datenquellen, unterschiedliche Datenqualität in den einzelnen Datenquellen, diverse Data-Mining-Techniken, verschiedene Methoden zur Messung des Data-Mining-Erfolgs usw.

Um nicht den Überblick zu verlieren, sollten Sie über ein explizit definiertes Prozessmodell für Data-Mining verfügen. Das Prozessmodell hilft Ihnen dabei, die in diesem Abschnitt aufgeführten Fragen zu beantworten, und stellt sicher, dass die wichtigen Punkte beachtet werden. Es dient als grundlegender Data-Mining-Plan, der Ihnen hilft, in der Komplexität der Daten nicht das Ziel aus den Augen zu verlieren.

Für die Verwendung mit SPSS Modeler wird das Data-Mining-Prozessmodell CRISP-DM (Cross-Industry Standard Process for Data Mining) empfohlen. Wie aus dem Namen ersichtlich, wurde dieses Modell als allgemeines Modell entwickelt, das für eine Vielzahl von Branchen und Geschäftsproblemen angewendet werden kann.

Prozessmodell "CRISP-DM"

Das allgemeine Prozessmodell CRISP-DM umfasst sechs Phasen, die die Hauptprobleme beim Data-Mining behandeln. Die sechs Phasen verschmelzen zu einem zyklischen Prozess, mit dem die Einbettung von Data-Mining in Ihre umfassenderen Geschäftspraktiken möglich ist.

Bei den sechs Phasen handelt es sich um folgende:

- **Geschäftsverständnis.** Dies ist wahrscheinlich die wichtigste Phase des Data-Mining. Das Geschäftsverständnis umfasst die Ermittlung von Geschäftszielen, die Einschätzung der Situation, die Bestimmung der Data-Mining-Ziele und die Erstellung eines Projektplans.
- **Datenverständnis.** Daten sind das "Rohmaterial" beim Data-Mining. Diese Phase widmet sich der Notwendigkeit, einen Einblick in die Datenressourcen und die Eigenschaften dieser Ressourcen zu gewinnen. Dazu gehört das Sammeln der Anfangsdaten, das Beschreiben von Daten, das Untersuchen von Daten und das Überprüfen der Datenqualität. Der Data Audit-Knoten, der in der Palette "Ausgabeknoten" verfügbar ist, ist ein unverzichtbares Tool für das Verständnis der Daten.
- **Datenvorbereitung.** Nach der Katalogisierung der Datenressourcen müssen Sie die Daten für das Data-Mining vorbereiten. Zu den Vorbereitungen gehören die Auswahl, Bereinigung, Konstruktion, Integration und Formatierung der Daten.
- **Modellierung.** Dies ist natürlich der Hauptteil des Data-Mining, in dem mit hoch entwickelten Analysemethoden Informationen aus den Daten extrahiert werden. Zu dieser Phase gehört die Auswahl von Modellierungsverfahren, das Generieren von Testdesigns und der Aufbau und die Auswertung von Modellen.
- **Evaluierung.** Nach der Auswahl der Modelle können Sie evaluieren, wie die Data-Mining-Ergebnisse zum Erreichen Ihrer Geschäftsziele beitragen können. Zu den Elementen in dieser Phase gehören die Evaluierung von Ergebnissen, die Überprüfung des Data-Mining-Prozesses und die Bestimmung der nächsten Schritte.
- **Bereitstellung.** Dies ist die Phase, in der Sie erstmals praktischen Nutzen aus ihren Bemühungen ziehen können. Diese Phase konzentriert sich auf die Integration des neu erworbenen Wissens in die alltäglichen Geschäftsabläufe zur Lösung des ursprünglichen Geschäftsproblems. Zu dieser Phase gehören die Planbereitstellung, Überwachung und Anpassung, das Erstellen eines endgültigen Berichts und die Überprüfung des Projekts.

Einige Punkte in diesem Prozessmodell sind besonders wichtig. Zum einen besteht zwar die allgemeine Tendenz, dass der Prozess die Schritte in der in den vorherigen Abschnitten angegebenen Reihenfolge durchläuft, aber es gibt auch eine Reihe von Stellen, an denen die Phasen sich gegenseitig auf nicht lineare Weise beeinflussen. So geht die Datenvorbereitung in der Regel der Modellierung voraus. Die während

der Modellierungsphase getroffenen Entscheidungen und gesammelten Informationen können den Benutzer häufig veranlassen, Teile der Datenvorbereitungsphase zu überdenken, was wiederum neue Modellierungsfragen aufwerfen kann. Bei den beiden Phasen findet eine gegenseitige Rückkopplung statt, bis beide Phasen adäquat gelöst wurden. Ebenso kann die Evaluierungsphase zu einer Neubewertung des ursprünglichen Geschäftsverständnisses führen und Sie könnten zu dem Schluss kommen, dass Sie sich bisher mit der falschen Frage auseinandergesetzt haben. An diesem Punkt können Sie Ihr Geschäftsverständnis überarbeiten und den Rest des Prozesses mit einem besseren Ziel im Auge erneut durchführen.

Der zweite wichtige Punkt ist die iterative Natur des Data-Mining. Es wird selten der Fall sein, dass Sie einfach ein Data-Mining-Projekt planen, ausführen und dann fertig sind und die Daten verwenden. Data-Mining für die Erfüllung der Anforderungen Ihrer Kunden ist keine einmalige Angelegenheit, sondern eine Daueraufgabe. Das aus einem Data-Mining-Zyklus gewonnene Wissen führt fast immer zu neuen Fragen, neuen Problemen und neuen Möglichkeiten zur Ermittlung und Erfüllung der Bedürfnisse Ihrer Kunden. Diese neuen Fragen, Probleme und Möglichkeiten lassen sich normalerweise durch ein erneutes Data-Mining Ihrer Daten angehen. Dieser Prozess aus Data-Mining und Ermittlung neuer Möglichkeiten sollte Ihnen in Fleisch und Blut übergehen und zu einem Grundstein Ihrer geschäftlichen Gesamtstrategie werden.

Diese Einführung kann nur einen kurzen Überblick über das Prozessmodell CRISP-DM bieten. Vollständige Einzelheiten zum Modell finden Sie in den folgenden Ressourcen:

- *CRISP-DM Handbuch*, das zusammen mit anderen Dokumenten im Verzeichnis *\Documentation* auf dem Installationsmedium aufgerufen werden kann.
- CRISP-DM-Hilfesystem, das über das Startmenü oder durch Klicken auf **CRISP-DM-Hilfe** im IBM SPSS Modeler-Hilfemenü aufgerufen werden kann.

Typen von Modellen

IBM SPSS Modeler bietet eine Vielzahl von Modellbildungsmethoden, die aus dem maschinellen Lernen, der künstlichen Intelligenz und der Statistik stammen. Mit den in der Modellierungspalette verfügbaren Methoden können Sie aus Ihren Daten neue Informationen ableiten und Vorhersagemodelle erstellen. Jede Methode hat ihre speziellen Stärken und eignet sich besonders für bestimmte Problemtypen.

Im *IBM SPSS Modeler-Anwendungshandbuch* finden Sie Beispiele für viele dieser Methoden sowie eine allgemeine Einführung in den Modellierungsprozess. Dieses Handbuch ist als Online-Lernprogramm und auch im PDF-Format verfügbar. Weitere Informationen finden Sie in „Anwendungsbeispiele“ auf Seite 4.

Modellierungsmethoden werden in die folgenden Kategorien unterteilt:

- Überwacht
- Assoziation
- Segmentierung

Überwachte Modelle

Überwachte Modelle verwenden den Wert mindestens eines **Eingabefeldes**, um den Wert mindestens eines Ausgabe- oder **Zielfeldes** vorherzusagen. Einige Beispiele dieser Verfahren sind: Entscheidungsbäume (C & R-Baum-, QUEST-, CHAID- und C5.0-Algorithmen), Regression (lineare, logistische, verallgemeinert lineare und Cox-Regressionsalgorithmen), neuronale Netze, Support Vector Machines und Bayes-Netze.

Überwachte Modelle können Unternehmen ein bekanntes Ergebnis vorhersagen. Beispielsweise, ob ein Kunde kaufen wird oder nicht oder ob eine Transaktion mit einem bekannten Betrugsmuster übereinstimmt. Zu den Modellierungstechniken gehören Maschinenlernen, Regelinduktion, Identifikation von Untergruppen, statistische Methoden und die Generierung mehrerer Modelle.

Überwachte Knoten



Mit dem Knoten "Autom. Klassifikationsmerkmal" können Sie eine Reihe verschiedener Modelle für binäre Ergebnisse ("Ja" oder "Nein", "Abwanderung" oder "Keine Abwanderung" usw.) erstellen und vergleichen, um den besten Ansatz für die jeweilige Analyse auszuwählen. Es wird eine Reihe von Modellierungsalgorithmen unterstützt, sodass Sie die gewünschten Methoden, die spezifischen Optionen für die jeweilige Methode und die Kriterien zum Vergleich der Ergebnisse auswählen können. Der Knoten generiert eine Gruppe von Modellen, die auf den angegebenen Optionen beruhen, und erstellt anhand der von Ihnen angegebenen Kriterien eine Rangordnung der besten Kandidaten.



Der Knoten "Autonumerisch" schätzt und vergleicht mit einer Reihe verschiedener Methoden Modelle für die Ergebnisse stetiger numerischer Bereiche. Der Knoten arbeitet auf dieselbe Weise wie der Knoten "Automatisches Klassifikationsmerkmal": Sie können die zu verwendenden Algorithmen auswählen und in einem Modellierungsdurchlauf mit mehreren Optionskombinationen experimentieren. Folgende Algorithmen werden unterstützt: neuronale Netze, C&R-Baum, CHAID, lineare Regression, verallgemeinerte lineare Regression und Support Vector Machines (SVM). Modelle können anhand von Korrelation, relativem Fehler bzw. Anzahl der verwendeten Variablen verglichen werden.



Der Knoten für Klassifizierungs- und Regressionsbäume (C&R-Bäume) generiert einen Entscheidungsbaum, mit dem Sie zukünftige Beobachtungen vorhersagen oder klassifizieren können. Bei dieser Methode wird eine rekursive Partitionierung verwendet, um die Trainingsdatensätze in Segmente aufzuteilen. Dabei wird bei jedem Schritt die Unreinheit verringert und ein Knoten im Baum wird als "rein" betrachtet, wenn 100 % der Fälle in eine bestimmte Kategorie des Zielfelds fallen. Ziel- und Eingabefelder können numerische Bereiche oder kategorial (nominal, ordinal oder Flags) sein. Alle Aufteilungen sind binär (nur zwei Untergruppen).



Der QUEST-Knoten bietet eine binäre Klassifizierungsmethode zum Erstellen von Entscheidungsbäumen, die dazu dient, die für Analysen von großen C&R-Bäumen erforderliche Verarbeitungszeit zu verkürzen. Gleichzeitig soll die in den Klassifizierungsbaummodellen festgestellte Tendenz verringert werden, die darin besteht, dass Eingaben bevorzugt werden, die mehr Aufteilungen erlauben. Eingabefelder können stetig (numerische Bereiche) sein, das Zielfeld muss aber kategorial sein. Alle Aufteilungen sind binär.



Der CHAID-Knoten generiert Entscheidungsbäume unter Verwendung von Chi-Quadrat-Statistiken zur Ermittlung optimaler Aufteilungen. Im Gegensatz zu den Knoten vom Typ "C&R-Baum" und "QUEST" kann CHAID nicht binäre Bäume generieren, d. h. Bäume mit Aufteilungen mit mehr als zwei Verzweigungen. Ziel- und Eingabefelder können in einem numerischen Bereich (stetig) oder kategorial sein. Exhaustive CHAID ist eine Änderung von CHAID, die noch gründlicher vorgeht, indem sie alle möglichen Aufteilungen untersucht, allerdings mehr Rechenzeit beansprucht.



Der C5.0-Knoten erstellt entweder einen Entscheidungsbaum oder ein Regelset. Das Modell teilt die Stichprobe auf der Basis des Felds auf, das auf der jeweiligen Ebene den maximalen Informationsgewinn liefert. Das Zielfeld muss kategorial sein. Es sind mehrere Aufteilungen in mehr als zwei Untergruppen zulässig.



Der Knoten "Entscheidungsliste" kennzeichnet Untergruppen bzw. Segmente, die eine höhere oder geringere Wahrscheinlichkeit für ein bestimmtes binäres Ergebnis aufweisen als die Gesamtpopulation. Sie könnten beispielsweise nach Kunden suchen, deren Abwanderung unwahrscheinlich ist oder die mit großer Wahrscheinlichkeit positiv auf eine Kampagne reagieren. Sie können Ihr Fachwissen in das Modell integrieren, indem Sie eigene, benutzerdefinierte Segmente hinzufügen und eine Vorschau anzeigen, in der alternative Modelle nebeneinander angezeigt werden, um die Ergebnisse zu vergleichen. Entscheidungslistenmodelle bestehen aus einer Liste von Regeln, bei denen jede Regel eine Bedingung und ein Ergebnis aufweist. Regeln werden in der vorgegebenen Reihenfolge angewendet und die erste Regel, die zutrifft, bestimmt das Ergebnis.



Bei linearen Regressionsmodellen wird ein stetiges Ziel auf der Basis linearer Beziehungen zwischen dem Ziel und einem oder mehreren Prädiktoren vorhergesagt.



Der Faktor/PCA-Knoten bietet leistungsstarke Datenreduktionsverfahren zur Verringerung der Komplexität der Daten. Die Hauptkomponentenanalyse (PCA) findet lineare Kombinationen der Eingabefelder, die die Varianz im gesamten Set der Felder am besten erfassen, wenn die Komponenten orthogonal (senkrecht) zueinander sind. Mit der Faktorenanalyse wird versucht, die zugrunde liegenden Faktoren zu bestimmen, die die Korrelationsmuster innerhalb eines Sets beobachteter Felder erklären. Bei beiden Ansätzen besteht das Ziel darin, eine kleinere Zahl abgeleiteter Felder zu finden, mit denen die Informationen im ursprünglichen Set der Felder effektiv zusammengefasst werden können.



Der Merkmalauswahlknoten sichtet die Eingabefelder, um auf der Grundlage einer Reihe von Kriterien (z. B. dem Prozentsatz der fehlenden Werte) zu entscheiden, ob diese entfernt werden sollen. Anschließend erstellt er eine Wichtigkeitsrangfolge der verbleibenden Eingaben in Bezug auf ein angegebenes Ziel. Beispiel: Angenommen, Sie haben ein Dataset mit Hunderten potenzieller Eingaben. Welche davon sind voraussichtlich für die Modellierung von medizinischen Behandlungsergebnissen von Bedeutung?



Bei der Diskriminanzanalyse werden strengere Annahmen als bei der logistischen Regression verwendet, sie kann jedoch eine wertvolle Alternative oder Ergänzung zu einer logistischen Regressionsanalyse sein, wenn diese Annahmen erfüllt sind.



Die logistische Regression ist ein statistisches Verfahren zur Klassifizierung von Datensätzen auf der Grundlage der Werte von Eingabefeldern. Sie ist analog zur linearen Regression, außer dass statt eines numerischen Bereichs ein kategoriales Zielfeld verwendet wird.



Das verallgemeinerte lineare Modell erweitert das allgemeine lineare Modell so, dass die abhängige Variable über eine angegebene Verknüpfungsfunktion in linearem Zusammenhang zu den Faktoren und Kovariaten steht. Außerdem ist es mit diesem Modell möglich, dass die abhängige Variable eine von der Normalverteilung abweichende Verteilung aufweist. Es deckt die Funktionen einer großen Bandbreite an Statistikmodellen ab, darunter lineare Regression, logistische Regression, loglineare Modelle für Häufigkeitsdaten und Überlebensmodelle mit Intervallzensurierung.



Verallgemeinerte lineare gemischte Modelle (GLMM - Generalized Linear Mixed Models) erweitern lineare Modelle so, dass das Ziel nicht normalverteilt zu sein braucht und über eine angegebene Verknüpfungsfunktion in einer linearen Beziehung zu den Faktoren und Kovariaten steht und die Beobachtungen korreliert werden können. Verallgemeinerte lineare gemischte Modelle decken eine breite Palette verschiedener Modelle ab, von einfacher linearer Regression bis hin zu komplexen Mehrebenenmodellen für nicht normalverteilte Longitudinaldaten.



Der Knoten vom Typ "Cox-Regression" ermöglicht Ihnen auch bei zensierten Datensätzen die Erstellung eines Überlebensmodells für Daten über die Zeit bis zum Eintreten des Ereignisses. Das Modell erstellt eine Überlebensfunktion, die die Wahrscheinlichkeit vorhersagt, dass das untersuchte Ereignis für bestimmte Werte der Eingabevariablen zu einem bestimmten Zeitpunkt (t) eingetreten ist.



Der Knoten "Support Vector Machine" (SVM) ermöglicht die Klassifizierung von Daten in eine von zwei Gruppen ohne Überanpassung. SVM eignet sich gut für umfangreiche Datasets, beispielsweise solche mit einer großen Anzahl von Eingabefeldern.



Mithilfe des Bayes-Netzknötens können Sie ein Wahrscheinlichkeitsmodell erstellen, indem Sie beobachtete und aufgezeichnete Hinweise mit "gesundem Menschenverstand" kombinieren, um die Wahrscheinlichkeit ihres Vorkommens zu ermitteln. Der Knoten ist speziell für Netze vom Typ "Tree Augmented Naïve Bayes" (TAN) und "Markov-Decke" gedacht, die in erster Linie zur Klassifizierung verwendet werden.



Mithilfe des Knotens für das lernfähige Antwortmodell (Self-Learning Response Model, SLRM) können Sie ein Modell erstellen, in dem das Modell anhand eines einzelnen neuen Falls oder einer kleinen Anzahl neuer Fälle neu eingeschätzt werden kann, ohne dass das Modell mit allen Daten neu trainiert werden muss.



Der Zeitreihenknoten berechnet Schätzungen für die exponentielle Glättung sowie univariate und multivariate ARIMA-Modelle (ARIMA steht für Autoregressive Integrated Moving Average (autoregressiver, integrierter gleitender Durchschnitt)) für Zeitreihendaten und erstellt Vorhersagen über die zukünftige Leistung. Dieser Zeitreihenknoten ähnelt dem bisherigen Zeitreihenknoten, der in SPSS Modeler Version 18 nicht mehr unterstützt wird. Allerdings ist dieser neuere Zeitreihenknoten für die Nutzung der Leistungsstärke von IBM SPSS Analytic Server und für die Verarbeitung großer Datenmengen konzipiert. Er zeigt das resultierende Modell im Ausgabeviewer an, der in SPSS Modeler Version 17 hinzugefügt wurde.



Der Knoten " k -Nächste Nachbarn" (KNN) verknüpft einen neuen Fall mit der Kategorie oder dem Wert der k Objekte, die ihm im Prädiktorraum am nächsten liegen, wobei k eine Ganzzahl ist. Ähnliche Fälle liegen nah beieinander und Fälle mit geringer Ähnlichkeit sind weit voneinander entfernt.



Der STP-Knoten (Spatio-Temporal Prediction - räumliche temporale Vorhersage) verwendet Daten, die Positionsdaten, Eingabefelder für Vorhersagen (Prädiktoren), ein Zeitfeld und ein Zielfeld enthalten. Die Daten enthalten für jede Position zahlreiche Zeilen, die die Werte der einzelnen Prädiktoren zum Zeitpunkt der Messung darstellen. Nach der Analyse der Daten können diese zur Vorhersage von Zielwerten an einer beliebigen Position innerhalb der in der Analyse verwendeten Shapedaten eingesetzt werden.

Assoziationsmodelle

Assoziationsmodelle finden Muster in Ihren Daten, bei denen mindestens eine Entität (wie Ereignisse, Einkäufe oder Attribute) mindestens einer anderen Entität zugeordnet sind. Die Modelle erstellen Regelsets, die diese Beziehungen definieren. Hier können die Felder innerhalb der Daten sowohl als Eingabe- als auch als Zielfelder fungieren. Sie könnten diese Assoziationen manuell finden, doch mithilfe von Assoziationsregelalgorithmen ist die Suche wesentlich schneller und es können komplexere Muster untersucht werden. Apriori- und Carma-Modelle sind Beispiele für die Verwendung solcher Algorithmen. Ein weiterer Typ eines Assoziationsmodells ist ein Sequenzerkennungsmodell, das sequenzielle Muster in zeitstrukturierten Daten findet.

Assoziationsmodelle sind bei der Vorhersage mehrerer Ergebnisse am nützlichsten, beispielsweise Kunden, die Produkt X gekauft haben, kauften auch Produkt Y und Z. Assoziationsmodelle ordnen einem Set von Bedingungen eine bestimmte Schlussfolgerung zu (wie zum Beispiel die Entscheidung, etwas zu kaufen). Der Vorteil von Algorithmen für Assoziationsregeln im Vergleich zu Algorithmen für Standardentscheidungsbäume (C5.0 und C&R-Baum) liegt darin, dass Zuordnungen zwischen beliebigen Attributen bestehen können. Ein Entscheidungsbaumalgorithmus erstellt Regeln mit nur einer Schlussfolgerung, während Assoziationsalgorithmen viele Regeln zu finden versuchen, von denen jede zu einer anderen Schlussfolgerung kommen kann.

Assoziationsknoten



Der Apriori-Knoten extrahiert ein Regelset aus den Daten und daraus die Regeln mit dem höchsten Informationsgehalt. Apriori bietet fünf verschiedene Methoden zur Auswahl von Regeln und verwendet ein ausgereiftes Indizierungsschema zur effizienten Verarbeitung großer Datensets. Bei großen Problemen ist Apriori in der Regel schneller zu trainieren, es gibt keine willkürliche Begrenzung für die Anzahl der Regeln, die beibehalten werden können, und es können Regeln mit bis zu 32 Vorbedingungen verarbeitet werden. Bei Apriori müssen alle Ein- und Ausgabefelder kategorial sein; dafür bietet es jedoch eine bessere Leistung, da es für diesen Datentyp optimiert ist.



Beim CARMA-Modell wird ein Regelset aus den Daten extrahiert, ohne dass Sie Eingabe- oder Zielfelder angeben müssen. Im Gegensatz zu Apriori bietet der CARMA-Knoten Einstellungen für Regelunterstützung (Unterstützung für Antezedens und Sukzedens) und nicht nur Unterstützung für Antezedens. Die erstellten Regeln können somit für eine größere Palette an Anwendungen verwendet werden, beispielsweise um eine Liste mit Produkten und Dienstleistungen (Antezedenzen) zu finden, deren Nachfolger (Sukzedens) das Element darstellt, das Sie in der Ferienzeit desselben Jahres bewerben möchten.



Der Sequenzknoten erkennt Assoziationsregeln in sequenziellen oder zeitorientierten Daten. Eine Sequenz ist eine Liste mit Elementsets, die in einer vorhersagbaren Reihenfolge auftreten. Beispiel: Ein Kunde, der einen Rasierer und After-Shave-Lotion kauft, kauft möglicherweise beim nächsten Einkauf Rasiercreme. Der Sequenzknoten basiert auf dem CARMA-Assoziationsregelalgorithmus, der eine effiziente bidirektionale Methode zum Suchen von Sequenzen verwendet.



Der Assoziationsregelknoten ähnelt dem Apriori-Knoten. Im Gegensatz zu diesem kann er jedoch Listendaten verarbeiten. Darüber hinaus kann der Assoziationsregelknoten in Verbindung mit IBM SPSS Analytic Server verwendet werden, um große Datenmengen zu verarbeiten und die schnellere parallele Verarbeitung zu nutzen.

Segmentierungsmodelle

Segmentierungsmodelle teilen die Daten in Segmente, oder Cluster, von Datensätzen auf, die ähnliche Muster von Eingabefeldern aufweisen. Da sie nur an den Eingabefeldern interessiert sind, verfügen Segmen-

tierungsmodelle nicht über die Konzepte der Ausgabe- oder Zielfelder. Beispiele für Segmentierungsmodelle sind Kohonen-Netze, K-Means-Clustering, TwoStep-Clustering und Anomalieerkennung.

Segmentierungsmodelle (auch "Clustering-Modelle") sind dann nützlich, wenn das genaue Ergebnis unbekannt ist (zum Beispiel beim Ermitteln neuer Betrugsmuster oder von Interessengruppen in Ihrem Kundenstamm). Clustering-Modelle konzentrieren sich auf die Ermittlung ähnlicher Datensätze und auf die Beschriftung der Datensätze anhand der Gruppe, in die sie gehören. Dies erfolgt ohne den Vorteil bereits zuvor vorhandener Kenntnisse der Gruppen und der zugehörigen Merkmale. Dies unterscheidet Clustering-Modelle von anderen Modellierungsverfahren: Es gibt kein zuvor definiertes Ausgabe- oder Zielfeld für das vorherzusagende Modell. Für diese Modelle gibt es keine richtigen oder falschen Antworten. Ihr Wert wird durch die Möglichkeit bestimmt, interessante Gruppierungen in den Daten zu erfassen und sinnvolle Beschreibungen dieser Gruppierungen zu liefern. Clustering-Modelle werden häufig verwendet, um Cluster oder Segmente zu erstellen, die dann als Eingaben in nachfolgenden Analysen verwendet werden (z. B. zur Segmentierung potenzieller Kunden in homogene Untergruppen).

Segmentierungsknoten



Mit dem Knoten "Autom. Cluster" können Sie Clustering-Modelle, die Gruppen und Datensätze mit ähnlichen Merkmalen identifizieren, schätzen und vergleichen. Die Funktionsweise des Knotens gleicht der von anderen Knoten für automatisierte Modellierung, und Sie können in einem einzigen Modellierungsdurchgang mit mehreren Optionskombinationen experimentieren. Modelle können mithilfe grundlegender Messwerte für Filterung und Rangfolge der Nützlichkeit von Clustermodellen verglichen werden, um ein Maß auf der Basis der Wichtigkeit von bestimmten Feldern zu liefern.



Der K-Means-Knoten teilt das Dataset in unterschiedliche Gruppen (oder Cluster) auf. Bei dieser Methode wird eine festgelegte Anzahl von Clustern definiert, den Clustern werden iterativ Datensätze zugewiesen und die Clusterzentren werden angepasst, bis eine weitere Verfeinerung keine wesentliche Verbesserung des Modells mehr darstellen würde. Statt zu versuchen, ein Ergebnis vorherzusagen, versucht K-Means mithilfe eines als "nicht überwacht Lernen" bezeichneten Verfahrens Muster im Set der Eingabefelder zu entdecken.



Der Kohonen-Knoten erstellt eine Art von neuronalem Netz, das verwendet werden kann, um ein Clustering des Datasets in einzelne Gruppen vorzunehmen. Wenn das Netz voll trainiert ist, sollten ähnliche Datensätze auf der Ausgabekarte eng nebeneinander stehen, während Datensätze, die sich unterscheiden, weit voneinander entfernt sein sollten. Die Zahl der von jeder Einheit im Modellnugget erfassten Beobachtungen gibt Aufschluss über die starken Einheiten. Dadurch wird ein Eindruck von der ungefähren Zahl der Cluster vermittelt.



Der TwoStep-Knoten verwendet eine aus zwei Schritten bestehende Clustering-Methode. Im ersten Schritt wird ein einzelner Durchlauf durch die Daten vorgenommen, bei dem die Eingangsrohdaten zu einem verwaltbaren Set von Subclustern komprimiert werden. Im zweiten Schritt werden die Subcluster mithilfe einer hierarchischen Clustering-Methode nach und nach in immer größere Cluster zusammengeführt. TwoStep hat den Vorteil, dass die optimale Anzahl von Clustern für die Trainingsdaten automatisch geschätzt wird. Mit dem Verfahren können gemischte Feldtypen und große Datasets effizient verarbeitet werden.



Der Knoten "Anomalieerkennung" ermittelt ungewöhnliche Fälle bzw. "Ausreißer", die nicht den Mustern von "normalen" Daten entsprechen. Mit diesem Knoten können Ausreißer ermittelt werden, selbst wenn sie keinem bereits bekannten Muster entsprechen und selbst wenn Sie nicht genau wissen, wonach Sie suchen.

Modelle für datenbankinternes Mining

IBM SPSS Modeler unterstützt die Integration in Data-Mining- und Modellierungstools von Datenbank Anbietern wie Oracle Data Miner, und Microsoft Analysis Services. Sie können Modelle erstellen, scoren und in der Datenbank speichern, ohne dazu die IBM SPSS Modeler-Anwendung verlassen zu müssen. Ausführliche Details finden Sie im Handbuch *IBM SPSS Modeler Datenbankinternes Mining*.

IBM SPSS Statistics-Modelle

Wenn auf Ihrem Computer eine Kopie von IBM SPSS Statistics installiert und lizenziert ist, können Sie auf bestimmte IBM SPSS Statistics-Routinen in IBM SPSS Modeler zugreifen und diese ausführen, um Modelle zu erstellen und zu scoren.

Data-Mining-Beispiele

Um einen Einblick in die praktische Verwendung von Data Mining zu gewinnen, beginnen Sie am besten mit einem Beispiel. Eine Reihe von Anwendungsbeispielen steht im *IBM SPSS Modeler-Anwendungshandbuch* zur Verfügung. Dieses Handbuch bietet eine kurze, gezielte Einführung in bestimmte Modellierungsmethoden und -verfahren. Weitere Informationen finden Sie in „Anwendungsbeispiele“ auf Seite 4.

Kapitel 5. Erstellen von Streams

Streamerstellung - Übersicht

Das Data-Mining mit IBM SPSS Modeler konzentriert sich auf die Ausführung von Daten über eine Reihe von Knoten, die als **Stream** bezeichnet wird. Diese Knotenreihe stellt Operationen dar, die für die Daten durchgeführt werden, während die Verknüpfungen zwischen den Knoten die Richtung des Datenflusses angeben. In der Regel verwenden Sie einen Datenstream, um Daten in IBM SPSS Modeler einzulesen, die Daten einer Reihe von Bearbeitungen zu unterziehen und die Daten anschließend an ein Ziel, z. B. eine Tabelle oder einen Viewer, zu senden.

Nehmen wir z. B. an, dass Sie eine Datenquelle öffnen, ein neues Feld hinzufügen, Datensätze auf der Basis von Werten in dem neuen Feld auswählen und dann die Ergebnisse in einer Tabelle anzeigen wollen. In diesem Fall würde Ihr Datenstream aus vier Knoten bestehen:



Ein Knoten **Variable Datei**, den Sie einrichten, um die Daten aus der Datenquelle zu lesen.



Ein Ableitungsknoten, über den Sie das neue berechnete Feld dem Dataset hinzufügen.



Ein Auswahlknoten, über den Sie Auswahlkriterien einrichten, um Datensätze aus dem Datenstream auszuschließen.



Ein Tabellenknoten, über den Sie die Ergebnisse Ihrer Bearbeitungen auf dem Bildschirm anzeigen.

Erstellen von Datenstreams

Die einzigartige SPSS Modeler-Schnittstelle ermöglicht das visuelle Mining von Daten unter Verwendung von Diagrammen mit Datenstreams. Am einfachsten können Sie einen Datenstream wie folgt erstellen:

- Fügen Sie dem Streamerstellungsbereich Knoten hinzu.
- Verbinden Sie die Knoten, um einen Stream zu bilden.
- Legen Sie die Knoten- bzw. Streamoptionen fest.
- Führen Sie den Stream aus.

Dieser Abschnitt enthält detaillierte Informationen zum Arbeiten mit Knoten, um komplexere Datenstreams zu erstellen. Es werden außerdem Optionen und Einstellungen für Knoten und Streams beschrieben. Beispiele mit Einzelschrittanleitungen zum Erstellen von Streams anhand der Daten aus dem Lieferumfang von SPSS Modeler (im Ordner Demos Ihrer Programminstallation) finden Sie in „Anwendungsbeispiele“ auf Seite 4.

Arbeiten mit Knoten

Knoten werden in IBM SPSS Modeler zum Untersuchen von Daten verwendet. Verschiedene Knoten im Arbeitsbereich stellen verschiedene Objekte und Aktionen dar. Die Palette unten im IBM SPSS Modeler-Fenster enthält alle Knoten, mit denen Streams erstellt werden können.

Es gibt mehrere Knotentypen. **Quellenknoten** bringen Daten in den Stream und befinden sich auf der Registerkarte "Quellen" der Knotenpalette. **Prozessknoten** führen Operationen an einzelnen Datensätzen und Feldern aus und sind auf den Registerkarten "Datensatzoperationen" und "Feldoperationen" der Palette zu finden. **Ausgabeknoten** erzeugen eine Vielzahl von Ausgaben für Daten-, Diagramm- und Modellergebnisse und befinden sich auf den Registerkarten "Diagramme", "Ausgabe" und "Export" der Knotenpalette. **Modellierungsknoten** verwenden statistische Algorithmen, um Modellnuggets zu erstellen, und befinden sich auf der Registerkarte "Modellierung" sowie der Registerkarte "Datenbankmodellierung" (falls aktiviert) der Knotenpalette. Weitere Informationen finden Sie in „Knotenpalette“ auf Seite 15.

Sie verbinden die Knoten, um Streams zu bilden, mit denen Sie bei Ausführung die Beziehungen visualisieren und Schlussfolgerungen ziehen können. Streams sind wie Scripts - Sie können sie speichern und mit anderen Datendateien wieder verwenden.

Ein ausführbarer Knoten, der Streamdaten verarbeitet, wird auch **Endknoten** genannt. Ein Modellierungs- oder Ausgabeknoten ist ein Endknoten, wenn er sich am Ende eines Streams oder Streamzweigs befindet. Sie können an einen Endknoten keine weiteren Knoten anschließen.

Anmerkung: Sie können die Knotenpalette anpassen. Weitere Informationen finden Sie in „Anpassen der Knotenpalette“ auf Seite 232.

Hinzufügen von Knoten zu einem Stream

Es gibt mehrere Möglichkeiten zum Hinzufügen von Knoten zu einem Stream über die Knotenpalette:

- Doppelklicken Sie auf einen Knoten in der Palette. *Hinweis:* Durch das Doppelklicken auf einen Knoten wird dieser automatisch mit dem aktuellen Stream verbunden. Weitere Informationen finden Sie in „Verbinden von Knoten in einem Stream“.
- Ziehen Sie einen Knoten der Palette in den Streamerstellungsbereich und legen Sie ihn dort ab.
- Klicken Sie auf einen Knoten in der Palette und dann in den Streamerstellungsbereich.
- Wählen Sie eine geeignete Option im Menü "Einfügen" von IBM SPSS Modeler aus.

Sobald Sie einen Knoten zum Streamerstellungsbereich hinzugefügt haben, doppelklicken Sie auf den Knoten, um das zugehörige Dialogfeld anzuzeigen. Die verfügbaren Optionen richten sich nach dem Knotentyp, den Sie hinzufügen. Informationen zu bestimmten Steuerelementen im Dialogfeld erhalten Sie, indem Sie auf die zugehörige Schaltfläche **Hilfe** klicken.

Entfernen von Knoten

Um einen Knoten aus dem Datenstream zu entfernen, klicken Sie darauf und drücken Sie die Löschtaste oder klicken Sie mit der rechten Maustaste und wählen Sie im Menü die Option **Löschen**.

Verbinden von Knoten in einem Stream

Knoten, die dem Streamerstellungsbereich hinzugefügt wurden, bilden so lange keinen Datenstream, bis sie verbunden sind. Die Verbindungen zwischen den Knoten zeigen an, in welche Richtung die Daten von einer Operation zur nächsten fließen. Es gibt mehrere Möglichkeiten, Knoten so zu verbinden, dass sie einen Stream bilden: durch Doppelklicken, mithilfe der mittleren Maustaste oder manuell.

Hinzufügen von Knoten durch Doppelklicken und Verbinden der Knoten:

Die einfachste Methode zum Bilden eines Streams besteht darin, auf die Knoten in der Palette doppelzu-klicken. Mit dieser Methode wird der neue Knoten automatisch mit dem im Streamerstellungsbereich ausgewählten Knoten verbunden. Wenn der Erstellungsbereich z. B. einen Datenbankknoten enthält, kön-

nen Sie diesen Knoten auswählen und dann auf den nächsten Knoten in der Palette doppelklicken, z. B. auf einen Ableitungsknoten. Mit dieser Aktion wird der Ableitungsknoten automatisch mit dem bereits vorhandenen Datenbankknoten verbunden. Sie können diesen Prozess so lange wiederholen, bis Sie einen Endknoten, z. B. einen Histogramm- oder Tabellenknoten, erreicht haben. Zu diesem Zeitpunkt werden sämtliche neue Knoten mit dem letzten vorgeordneten Knoten verbunden, der kein Endknoten ist.

Verbinden von Knoten mithilfe der mittleren Maustaste:

Im Streamerstellungsbereich können Sie mithilfe der mittleren Maustaste durch Klicken und Ziehen einen Knoten mit einem anderen verbinden. (Wenn Ihre Maus keine mittlere Taste aufweist, können Sie dies simulieren, indem Sie die Taste Alt drücken, während Sie die Maus von einem Knoten zu einem anderen ziehen.)

Manuelles Verbinden von Knoten:

Falls Ihre Maus keine mittlere Maustaste aufweist und Sie es bevorzugen, die Knoten manuell zu verbinden, können Sie das Popup-Menü eines Knotens verwenden, um diesen mit einem bereits im Erstellungsbereich befindlichen anderen Knoten zu verbinden.

1. Klicken Sie mit der rechten Maustaste auf den Knoten, der die Verbindung starten soll. Dadurch wird das Knotenmenü geöffnet.
2. Klicken Sie im Menü auf **Verbinden**.
3. Sowohl am Ausgangsknoten als auch am Cursor wird ein Verbindungssymbol angezeigt. Klicken Sie auf einen zweiten Knoten im Erstellungsbereich, um die beiden Knoten zu verbinden.

Beim Verbinden von Knoten müssen verschiedene Richtlinien befolgt werden. Sie erhalten eine Fehlermeldung, wenn Sie versuchen, einen der folgenden Verbindungstypen zu erstellen:

- Eine Verbindung, die zu einem Quellenknoten führt
- Eine Verbindung, die von einem Endknoten kommt
- Einen Knoten mit mehr als der maximalen Anzahl Eingabeverbindungen
- Verbindung von zwei Knoten, die bereits miteinander verbunden sind
- Umlauf (Daten kehren zu einem Knoten zurück, von dem sie bereits geflossen sind)

Umgehen von Knoten in einem Stream

Wenn Sie einen Knoten im Datenstream umgehen, werden sämtliche Ein- und Ausgabeverbindungen, die direkt von den Eingabeknoten zu den Ausgabeknoten führen, ersetzt. Wenn der Knoten nicht über Eingabe- und Ausgabeverbindungen verfügt, werden sämtliche Verbindungen gelöscht anstatt umgeleitet.

So verfügen Sie u. U. über einen Stream, der ein neues Feld ableitet, Felder filtert und dann die Ergebnisse in einem Histogramm oder in einer Tabelle untersucht. Wenn Sie dasselbe Diagramm und die Tabelle für Daten *vor* dem Filtern von Feldern anzeigen möchten, können Sie dem Stream entweder neue Histogramm- und Tabellenknoten hinzufügen oder den Filterknoten umgehen. Wenn Sie den Filterknoten umgehen, werden die Verbindungen mit dem Diagramm und der Tabelle direkt vom Ableitungsknoten übergeben. Die Verbindung des Filterknotens mit dem Stream wird aufgehoben.

Umgehen eines Knotens

1. Verwenden Sie im Streamerstellungsbereich die mittlere Maustaste, um auf den Knoten doppelzuklicken, den Sie umgehen möchten. Alternativ können Sie auch die Kombination Alt+Doppelklick verwenden.

Hinweis: Sie können diese Aktion unter Verwendung der Option **Rückgängig** im Menü "Bearbeiten" oder mit der Tastenkombination Strg+Z rückgängig machen.

Inaktivieren von Knoten in einem Stream

Prozessknoten mit einer einzelnen Eingabe innerhalb von Streams können inaktiviert werden, mit dem Ergebnis, dass der Knoten bei der Ausführung des Streams ignoriert wird. Auf diese Weise müssen Sie den Knoten nicht löschen oder umgehen und können seine Verbindung mit den restlichen Knoten bestehen lassen. Die Knoteneinstellungen können nach wie vor angezeigt und bearbeitet werden. Änderungen werden jedoch erst dann wirksam, wenn Sie den Knoten wieder aktivieren.

Angenommen Sie haben einen Stream, der mehrere Felder filtert und anschließend aus dem reduzierten Dataset Modelle erstellt. Wenn Sie dieselben Modelle auch *ohne* Filterung von Feldern erstellen möchten, um festzustellen, ob sie das Ergebnis verbessern, können Sie den Filterknoten inaktivieren. Wenn Sie den Filterknoten inaktivieren, werden die Verbindungen mit dem Modellierungsknoten direkt vom Ableitungsknoten an den Typknoten übergeben.

So inaktivieren Sie einen Knoten:

1. Klicken Sie im Streamerstellungsbereich mit der rechten Maustaste auf den Knoten, den Sie inaktivieren möchten.
2. Klicken Sie im Popup-Menü auf **Knoten inaktivieren**.

Alternativ können Sie auch im Menü "Bearbeiten" auf **Knoten > Knoten inaktivieren** klicken. Wenn Sie den Knoten wieder in den Stream integrieren möchten, wählen Sie auf dieselbe Weise wie oben beschriebenen **Knoten aktivieren**.

Hinweis: Sie können diese Aktion unter Verwendung der Option **Rückgängig** im Menü "Bearbeiten" oder mit der Tastenkombination Strg+Z rückgängig machen.

Hinzufügen von Knoten in bereits vorhandenen Verbindungen

Zwischen zwei verbundenen Knoten können Sie einen neuen Knoten einfügen, indem Sie den Pfeil ziehen, der die beiden Knoten verbindet.

1. Klicken Sie mit der mittleren Maustaste auf den Verbindungspfeil, in den Sie den Knoten einfügen möchten, und ziehen Sie diesen. Alternativ können Sie die Taste Alt drücken und dann mit der Maus ziehen. Damit simulieren Sie die mittlere Maustaste.
2. Ziehen Sie die Verbindung an den Knoten, den Sie einfügen möchten, und lassen Sie die Maustaste los.

Hinweis: Sie können neue Verbindungen aus dem Knoten entfernen und den ursprünglichen Knoten wiederherstellen, indem Sie den Knoten **umgehen**.

Löschen von Verbindungen zwischen Knoten

So löschen Sie die Verbindung zwischen zwei Knoten:

1. Klicken Sie mit der rechten Maustaste auf den Verbindungspfeil.
2. Klicken Sie im Menü auf **Verbindung löschen**.

Führen Sie zum Löschen aller Verbindungen zu und von einem Knoten eine der folgenden Aktionen aus:

- Wählen Sie den Knoten aus und drücken Sie F3.
- Wählen Sie einen Knoten aus und klicken Sie im Hauptmenü auf:

Bearbeiten > Knoten > Trennen

Festlegen von Optionen für Knoten

Sobald Sie Knoten erstellt und miteinander verbunden haben, gibt es verschiedene Optionen zum Anpassen der Knoten. Klicken Sie mit der rechten Maustaste auf einen Knoten und wählen Sie eine der Menüoptionen aus.

- Klicken Sie auf **Bearbeiten**, um das Dialogfeld für den ausgewählten Knoten zu öffnen.
- Klicken Sie auf **Verbinden**, um einen Knoten manuell mit einem anderen zu verbinden.

- Klicken Sie auf **Trennen**, um alle Verknüpfungen zum und vom Knoten zu löschen.
- Klicken Sie auf **Umbenennen und mit Anmerkung versehen**, um die Registerkarte "Anmerkungen" im Bearbeitungsdialogfeld aufzurufen.
- Klicken Sie auf **Neuer Kommentar**, um einem Knoten einen Kommentar hinzuzufügen. Weitere Informationen finden Sie in „Hinzufügen von Kommentaren und Anmerkungen zu Knoten und Streams“ auf Seite 60.
- Klicken Sie auf **Knoten inaktivieren**, um den Knoten während der Verarbeitung "auszublenden". Um den Knoten wieder "einzublenden", klicken Sie auf **Knoten aktivieren**. Weitere Informationen finden Sie in „Inaktivieren von Knoten in einem Stream“ auf Seite 40.
- Klicken Sie auf **Ausschneiden** oder **Löschen**, um die ausgewählten Knoten aus dem Streamerstellungsbereich zu entfernen. *Hinweis:* Mit der Option **Ausschneiden** können Sie Knoten einfügen, nicht jedoch mit **Löschen**.
- Klicken Sie auf **Knoten kopieren**, um eine Kopie des Knotens ohne Verbindungen zu erstellen. Diese kann einem neuen oder bereits vorhandenen Stream hinzugefügt werden.
- Klicken Sie auf **Knoten laden**, um einen zuvor gespeicherten Knoten zu öffnen und seine Optionen in den aktuell ausgewählten Knoten zu laden. Die Knoten müssen denselben Typ aufweisen.
- Klicken Sie auf **Knoten abrufen**, um einen Knoten von einer verbundenen IBM SPSS Collaboration and Deployment Services Repository-Instanz abzurufen.
- Klicken Sie auf **Knoten speichern**, um die Knotendetails in einer Datei zu speichern. Sie können Knotendetails nur in einen anderen Knoten desselben Typs laden.
- Klicken Sie auf **Knoten temporär speichern**, um den ausgewählten Knoten in einer verbundenen IBM SPSS Collaboration and Deployment Services Repository-Instanz zu speichern.
- Klicken Sie auf **Cache**, um das Menü zu erweitern, das Optionen für das Caching des ausgewählten Knotens enthält.
- Klicken Sie auf **Datenzuordnung**, um das Menü zu erweitern, das Optionen zum Zuordnen von Daten zu einer neuen Quelle oder zum Festlegen von erforderlichen Feldern enthält.
- Klicken Sie auf **Superknoten erstellen**, um das Menü zu erweitern, das Optionen zum Erstellen eines Superknotens im aktuellen Stream enthält.
- Klicken Sie auf **Benutzereingabeknoten generieren**, um den ausgewählten Knoten zu ersetzen. Die von diesem Knoten erstellten Beispiele enthalten dieselben Felder wie der aktuelle Knoten.
- Klicken Sie auf **Ab hier ausführen**, um alle Endknoten auszuführen, die dem ausgewählten Knoten nachgeordnet sind.

Caching-Optionen für Knoten

Um die Streamausführung zu optimieren, können Sie einen *Cache* für Knoten einrichten, die keine Endknoten sind. Wenn Sie einen Cache für einen Knoten einrichten, wird der Cache mit den Daten gefüllt, die bei der nächsten Ausführung des Datenstreams durch den Knoten laufen. Ab diesem Zeitpunkt werden die Daten aus dem Cache (der sich in einem temporären Verzeichnis auf dem Datenträger befindet) und nicht mehr aus der Datenquelle eingelesen.

Caching ist am nützlichsten im Anschluss an zeitaufwendige Operationen, wie Sortieren, Zusammenführen oder Aggregation. Angenommen, es ist beispielsweise ein Quellenknoten vorhanden, der zum Lesen von Umsatzdaten aus einer Datenbank eingerichtet ist, und ein Aggregatknoten, der den Umsatz nach Standort zusammenfasst. Sie können einen Cache an dem Aggregatknoten einrichten anstatt an dem Quellenknoten, weil Sie möchten, dass der Cache die aggregierten Daten und nicht das gesamte Dataset speichert.

Anmerkung: Ein Caching bei Quellenknoten, bei denen einfach eine Kopie der ursprünglichen Dateien, die in IBM SPSS Modeler eingelesen wurden, gespeichert wird, bewirkt meist keine Leistungsverbesserung.

Knoten mit aktiviertem Caching werden mit einem kleinen Dokumentsymbol in der oberen rechten Ecke angezeigt. Wenn die Daten am Knoten zwischengespeichert werden, ist das Dokumentsymbol grün.

So aktivieren Sie einen Cache:

1. Klicken Sie im Streamerstellungsbereich mit der rechten Maustaste auf den Knoten und klicken Sie im Menü auf die Option **Cache**.
2. Klicken Sie im Untermenü auf die Option **Aktivieren**.
3. Sie können den Cache inaktivieren, indem Sie mit der rechten Maustaste auf den Knoten klicken und im Untermenü auf **Inaktivieren** klicken.

Caching von Knoten in einer Datenbank

Bei in einer Datenbank ausgeführten Streams muss nicht das Dateisystem als Cache verwendet werden, sondern die Daten können mitten im Stream in einer temporären Tabelle in der Datenbank zwischengespeichert werden. Bei einer Kombination mit SQL-Optimierung kann dies zu einer erheblichen Leistungssteigerung führen. So kann beispielsweise die Ausgabe aus einem Stream, der mehrere Tabellen zusammenfasst, um eine Data-Mining-Ansicht zu erstellen, im Cache gespeichert und bei Bedarf wieder verwendet werden. Durch die automatische Generierung von SQL für alle nachgeordneten Knoten lässt sich die Leistung weiter erhöhen.

Um das Caching von Datenbanken nutzen zu können, müssen sowohl SQL-Optimierung als auch Datenbankcaching aktiviert sein. Die Optimierungseinstellungen auf dem Server haben Vorrang vor jenen auf dem Client. Weitere Informationen finden Sie in „Festlegen von Optimierungsoptionen für Streams“ auf Seite 47.

Klicken Sie bei aktiviertem Datenbankcaching einfach mit der rechten Maustaste auf einen Knoten (keinen Endknoten), um die Daten an dieser Stelle im Cache zu speichern, und der Cache wird bei der nächsten Ausführung des Streams automatisch unmittelbar in der Datenbank erstellt. Wenn Datenbankcaching oder SQL-Optimierung nicht aktiviert ist, wird der Cache stattdessen in das Dateisystem geschrieben.

Anmerkung: Die folgenden Datenbanken unterstützen temporäre Tabellen zu Caching-Zwecken: Db2, Oracle, SQL Server und Teradata. Andere Datenbanken, beispielsweise Netezza, verwenden für das Datenbankcaching eine normale Tabelle. Der SQL-Code kann für bestimmte Datenbanken angepasst werden. Wenden Sie sich an den Kundendienst, wenn Sie Unterstützung benötigen.

Leeren eines Cache

Ein weißes Dokumentsymbol an einem Knoten weist darauf hin, dass sein Cache leer ist. Wenn der Cache voll ist, wird das Dokumentsymbol grün. Wenn Sie den Inhalt des Caches ersetzen möchten, müssen Sie den Cache zunächst leeren und den Datenstream dann erneut ausführen, um ihn erneut zu füllen.

1. Klicken Sie im Streamerstellungsbereich mit der rechten Maustaste auf den Knoten und klicken Sie im Menü auf die Option **Cache**.
2. Wählen Sie im Untermenü die Option **Leeren** aus.

Speichern eines Cache

Sie können den Inhalt eines Cache als Datendatei von IBM SPSS Statistics (*.sav) speichern. Sie können die Datei entweder als Cache erneut laden oder Sie können einen Knoten einrichten, der die Cachedatei als Datenquelle verwendet. Sie können auch einen Cache laden, den Sie aus einem anderen Projekt gespeichert haben.

1. Klicken Sie im Streamerstellungsbereich mit der rechten Maustaste auf den Knoten und klicken Sie im Menü auf die Option **Cache**.
2. Wählen Sie im Untermenü die Option **Cache speichern** aus.
3. Blättern Sie im Dialogfeld "Cache speichern" an den Speicherort, an dem die Cachedatei gespeichert werden soll.
4. Geben Sie im Textfeld "Dateiname" einen Namen ein.

5. Stellen Sie sicher, dass in der Liste "Dateityp" die Option ***.sav** ausgewählt ist, und klicken Sie auf **Speichern**.

Laden eines Cache

Wenn Sie eine Cachedatei gespeichert haben, bevor Sie diese aus dem Knoten entfernen, können Sie sie erneut laden.

1. Klicken Sie im Streamerstellungsbereich mit der rechten Maustaste auf den Knoten und klicken Sie im Menü auf die Option **Cache**.
2. Wählen Sie im Untermenü die Option **Cache laden** aus.
3. Blättern Sie im Dialogfeld "Cache laden" an den Speicherort der Cachedatei, wählen Sie diese aus und klicken Sie auf **Laden**.

Vorschau von Daten in Knoten

Damit sichergestellt ist, dass Daten beim Erstellen eines Streams wie erwartet geändert werden, können Sie alle Ihre Daten bei jedem wesentlichen Schritt durch einen Tabellenknoten laufen lassen. Um sich das zu ersparen, können Sie von jedem Knoten eine Vorschau mit einem Muster der Daten generieren, die erstellt werden, und damit den Zeitaufwand für die Erstellung jedes Knotens verringern.

Für Knoten, die einem Modellnugget vorgeordnet sind, zeigt die Vorschau die Eingabefelder an. Für ein dem Nugget vorgeordnetes Modellnugget oder Knoten, die dem Nugget vorgeordnet sind (ausgenommen Endknoten), zeigt die Vorschau die Eingabefelder und die generierten Felder an.

Die Standardanzahl der angezeigten Zeilen ist 10. Sie können diese Einstellung jedoch in den Streameigenschaften ändern. Weitere Informationen finden Sie in „Festlegen allgemeiner Optionen für Streams“ auf Seite 44.

Mithilfe des Menüs **Generieren** können Sie mehrere Arten von Knoten erstellen.

Anmerkung: Bei der Vorschau der von diesem Knoten generierten Daten werden alle Eigenschaftsänderungen auf diesen Knoten angewendet und können nicht abgebrochen werden (dieses Verhalten entspricht dem Klicken auf **Anwenden**).

Sperrungen von Knoten

Um andere Benutzer daran zu hindern, die Einstellungen eines oder mehrerer Knoten in einem Stream zu ändern, können Sie den oder die Knoten im besonderen Knotentyp namens Superknoten verkapseln und dann den Superknoten durch Zuweisen eines Kennwortschutzes sperren.

Arbeiten mit Streams

Sobald Sie die Quellen-, Prozess- und Endknoten im Streamerstellungsbereich verbunden haben, haben Sie einen Stream erstellt. Als Knotensammlung können Streams gespeichert, mit Anmerkungen versehen und zu Projekten hinzugefügt werden. Sie können für Streams auch zahlreiche Optionen wie Optimierung, Datums- und Uhrzeiteinstellungen, Parameter und Scripts festlegen. Diese Eigenschaften werden in den nachfolgenden Themen beschrieben.

In IBM SPSS Modeler können Sie innerhalb derselben IBM SPSS Modeler-Sitzung mehrere Datenstreams verwenden und ändern. Die rechte Seite des Hauptfensters enthält den Managerbereich, mit dem Sie durch die aktuell geöffneten Streams, Ausgaben und Modelle navigieren können. Wenn der Managerbereich nicht sichtbar ist, wählen Sie im Ansichtsmenü die Option **Manager** aus und klicken Sie dann auf die Registerkarte **Streams**.

Auf dieser Registerkarte können Sie:

- Auf Streams zugreifen.
- Streams speichern.

- Streams im aktuellen Projekt speichern.
- Streams schließen.
- Neue Streams öffnen.
- Streams aus einem Repository von IBM SPSS Collaboration and Deployment Services (falls an Ihrem Standort verfügbar) speichern und abrufen. Weitere Informationen finden Sie in „Informationen zu IBM SPSS Collaboration and Deployment Services Repository“ auf Seite 191.

Klicken Sie mit der rechten Maustaste auf die Registerkarte "Streams", um diese Optionen aufzurufen.

Festlegen von Optionen für Streams

Sie können eine Reihe von Optionen angeben, die auf den aktuellen Stream angewendet werden sollen. Sie können diese Optionen auch als Standard zur Anwendung auf alle Ihre Streams speichern. Es handelt sich hier um die folgenden Optionen:

- **Allgemein.** Verschiedene Optionen, wie beispielsweise die im Stream zu verwendenden Symbole und die zu verwendende Textcodierung. Weitere Informationen finden Sie in „Festlegen allgemeiner Optionen für Streams“.
- **Datum/Uhrzeit.** Optionen, die sich auf das Format von Datums- und Zeitausdrücken beziehen. Weitere Informationen finden Sie in „Festlegen von Datums- und Zeitoptionen für Streams“ auf Seite 46.
- **Zahlenformate.** Optionen, die das Format numerischer Ausdrücke festlegen. Weitere Informationen finden Sie in „Festlegen von Zahlenformatoptionen für Streams“ auf Seite 47.
- **Optimierung.** Optionen zur Optimierung der Streamleistung. Weitere Informationen finden Sie in „Festlegen von Optimierungsoptionen für Streams“ auf Seite 47.
- **Protokollierung und Status.** Optionen zur Festlegung von SQL-Protokollierung und Datensatzstatus. Weitere Informationen finden Sie in „Festlegen von Optionen für SQL-Protokollierung und Datensatzstatus für Streams“ auf Seite 49.
- **Layout.** Optionen, die sich auf das Layout des Streams im Erstellungsbereich beziehen. Weitere Informationen finden Sie in „Festlegen von Layoutoptionen für Streams“ auf Seite 49.
- **Analytic Server.** Optionen, die sich auf die Verwendung von Analytic Server mit SPSS Modeler beziehen. Weitere Informationen finden Sie in „Analytic Server-Streameigenschaften“ auf Seite 50.
- **Räumlich.** Optionen, die sich auf die Formatierung georäumlicher Daten für die Verwendung im Stream beziehen. Weitere Informationen finden Sie in „Festlegen georäumlicher Optionen für Streams“ auf Seite 51.

So legen Sie Streamoptionen fest

1. Klicken Sie im Menü "Datei" auf die Option **Streameigenschaften** (oder wählen Sie den Stream auf der Registerkarte "Streams" im Managerbereich aus, klicken Sie mit der rechten Maustaste und wählen Sie dann im Popup-Menü die Option **Streameigenschaften** aus).
2. Klicken Sie auf die Registerkarte **Optionen**.

Klicken Sie als Alternative im Menü "Tools" auf:

Streameigenschaften > Optionen

Festlegen allgemeiner Optionen für Streams: Die allgemeinen Optionen sind ein Set verschiedener Optionen, die für verschiedene Aspekte des aktuellen Streams gelten.

Der Abschnitt **Basis** enthält die folgenden Basisoptionen:

- **Dezimalzeichen.** Wählen Sie entweder ein Komma (,) oder einen Punkt (.) als Dezimaltrennzeichen aus.
- **Symbol für Zifferngruppierung.** Als Zahlenanzeigeformat wählen Sie das Symbol aus, das zur Gruppierung von Werten verwendet wird (z. B. der Punkt in 3.000,00). Zu den Optionen gehören "Keine",

"Punkt", "Komma", "Leerzeichen" und "Durch Ländereinstellung definiert" (in diesem Fall wird der Standardwert für die aktuelle Ländereinstellung verwendet).

- **Codierung.** Geben Sie die Streamstandardmethode für die Textcodierung an. (*Hinweis:* Gilt nur für den Quellenknoten "Variable Datei" und den Exportknoten "Flatfile". Andere Knoten verwenden diese Einstellung nicht; die meisten Datendateien verfügen über eingebettete Codierungsinformationen.) Sie können entweder den Systemstandardwert oder "UTF-8" auswählen. Der Systemstandardwert wird in der Windows-Systemsteuerung bzw. bei Ausführung im verteilten Modus auf dem Server-Computer angegeben. Weitere Informationen finden Sie in „Unicode-Unterstützung in IBM SPSS Modeler“ auf Seite 257.
- **Regelsetauswertung.** Legt fest, wie Regelsetmodelle ausgewertet werden. Standardmäßig verwenden Regelsets die Option **Voting**, um Vorhersagen aus einzelnen Regeln zu kombinieren und die endgültige Vorhersage zu bestimmen. Um sicherzustellen, dass Regelsets standardmäßig die Regel "Erster Treffer" verwenden, wählen Sie **Erster Treffer** aus. Beachten Sie, dass diese Option nicht für Entscheidungslistenmodelle gilt, die die Regel "Erster Treffer" wie im Algorithmus definiert immer verwenden.

Maximale Anzahl an Zeilen in Datenvorschau. Geben Sie die Anzahl der Zeilen an, die angezeigt werden sollen, wenn eine Vorschau der Daten für einen Knoten angefordert wird. Weitere Informationen finden Sie in „Vorschau von Daten in Knoten“ auf Seite 43.

Maximale Mitglieder für nominale Felder. Wählen Sie diese Option aus, um eine maximale Anzahl Mitglieder für nominale (festgelegte) Felder anzugeben, nach der sich der Datentyp des Feldes in **Ohne Typ** ändert. Diese Option ist hilfreich, wenn Sie mit großen nominalen Feldern arbeiten. *Hinweis:* Wenn das Messniveau eines Feldes auf **Ohne Typ** gesetzt wird, wird seine Rolle automatisch auf **Keine** gesetzt. Dies bedeutet, dass die Felder nicht zur Modellierung verfügbar sind.

Satzgröße für Kohonen- und K-Means-Modellierung begrenzen. Wählen Sie diese Option aus, um eine maximale Anzahl Mitglieder für nominale Felder anzugeben, die bei der Kohonen-Netz- und K-Means-Modellierung verwendet werden. Die Standardsatzgröße beträgt 20. Danach wird das Feld ignoriert und es wird eine Warnung angezeigt, in der Informationen zu dem betreffenden Feld angegeben werden.

Beachten Sie, dass aus Kompatibilitätsgründen diese Option auch für den alten Knoten "Neuronales Netz" gilt, der in IBM SPSS Modeler Version 14 ersetzt wurde; einige alte Streams verfügen möglicherweise immer noch über diesen Knoten.

Quellenknoten bei Ausführung aktualisieren. Wählen Sie diese Option aus, um bei der Ausführung des aktuellen Streams alle Quellenknoten automatisch zu aktualisieren. Diese Aktion entspricht dem Klicken auf die Schaltfläche **Aktualisieren** auf einem Quellenknoten, mit der Ausnahme, dass mit dieser Option automatisch alle Quellenknoten (außer die Benutzereingabeknoten) für den aktuellen Stream aktualisiert werden.

Anmerkung: Wenn Sie diese Option auswählen, werden die Caches von nachgeordneten Knoten geleert, auch wenn sich die Daten nicht geändert haben. Wenn Sie die Option **Aktuellen Stream ausführen** aus der Symbolleiste verwenden, werden die Caches jedoch nur einmal pro Streamausführung geleert, sodass Sie nachgeordnete Caches weiterhin als temporären Speicher für eine einzelne Ausführung verwenden können. Nehmen wir z. B. an, dass Sie einen Cache in die Mitte des Streams nach einer komplexen Ableitungsoperation gesetzt haben und Sie mehrere Grafiken und Berichte auf diesen Ableitungsknoten folgend angefügt haben. Wenn Sie den Stream ausführen, wird der Cache am Ableitungsknoten geleert und neu gefüllt, jedoch nur für die erste Grafik oder den ersten Bericht. Nachfolgende Endknoten lesen die Daten aus dem Cache des Ableitungsknotens. *Hinweis:* Wenn Sie sich für die einzelne Ausführung der Endknoten entscheiden (sofern Sie mehrere Endknoten haben), wird anstelle der Option **Aktuellen Stream ausführen** bei jeder Ausführung eines Endknotens der Cache geleert.

Feld- und Wertbeschriftungen in Ausgabe anzeigen. Zeigt Feld- und Wertbeschriftungen in Tabellen, Diagrammen und anderen Ausgabeformen an. Wenn keine Beschriftungen vorhanden sind, werden stattdessen die Feldnamen und Datenwerte angezeigt. Die Beschriftungen sind standardmäßig ausgeschaltet;

Sie können die Beschriftungen jedoch individuell an anderer Stelle in IBM SPSS Modeler ein- und ausschalten. Sie können mit einer Schaltfläche zum Ein-/Ausschalten in der Symbolleiste auch auswählen, dass die Beschriftungen im Ausgabefenster angezeigt werden.



Abbildung 11. Symbolleistensymbol zum Ein-/Ausschalten von Feld- und Wertbeschriftungen

Ausführungszeiten anzeigen. Zeigt nach der Ausführung des Streams einzelne Ausführungszeiten für Streamknoten auf der Registerkarte "Ausführungszeiten" an. Weitere Informationen finden Sie in „Anzeigen der Knotenausführungszeiten“ auf Seite 52.

Der Abschnitt **Automatische Knotenerstellung** enthält die folgenden Optionen zum automatischen Erstellen von Knoten in einzelnen Streams. Diese Optionen steuern, ob die Modellierungsnuggets im Streamerstellungsbereich eingefügt werden sollen, wenn neue Nuggets generiert werden. Standardmäßig gelten diese Optionen nur für Streams, die in Version 16 oder höher erstellt wurden. Wenn Sie in IBM SPSS Modeler Version 16 oder höher einen Stream öffnen, der in Version 15 oder niedriger erstellt wurde, und einen Modellierungsknoten ausführen, wird das Nugget nicht wie in früheren Releases in den Streamerstellungsbereich gestellt. Wenn Sie einen neuen Stream über IBM SPSS Modeler Version 16 oder eine spätere Version erstellen und einen Modellierungsknoten ausführen, wird das generierte Nugget in den Streamerstellungsbereich gestellt. Dies ist beabsichtigt, weil z. B. die Option **Modellanwendungsknoten für neue Modellausgabe erstellen** vermutlich Streams mit einer Version vor Version 16, die im Stapelmodus ausgeführt werden, in IBM SPSS Collaboration and Deployment Services und in anderen Umgebungen, in denen die Clientbenutzerschnittstelle von IBM SPSS Modeler Server nicht vorhanden ist, unterbrechen würde.

- **Modellanwendungsknoten für neue Modellausgabe erstellen.** Erstellt automatisch Modellanwendungsknoten für die neue Modellausgabe. Wenn Sie diese Option auswählen, können Sie in **Modellaktualisierungslinks erstellen** auch auswählen, ob die Links als aktiviert oder inaktiviert festgelegt oder nicht erstellt werden sollen.

Wenn ein neuer Modellanwender oder Quellenknoten erstellt wird, steuern die Linkoptionen in den Dropdown-Menüs, ob die Aktualisierungslinks zwischen dem Erstellerknoten und dem neuen Knoten erstellt werden und ggf., in welchem Modus sie sich befinden. Wenn Links erstellt werden, sollen sie vermutlich aktiviert sein, allerdings geben diese Optionen dem Benutzer eine uneingeschränkte Kontrolle.

- **Quellenknoten aus Quellenerstellungsprogrammen erstellen.** Erstellt automatisch Quellenknoten aus den Quellenerstellungsprogrammen. Ähnlich wie in der vorherigen Option können Sie bei Auswahl dieser Option im Dropdown-Menü **Quellenaktualisierungslinks erstellen** auch angeben, ob die Aktualisierungslinks als aktiviert oder inaktiviert festgelegt oder nicht erstellt werden sollen.

Als Standard speichern. Die angegebenen Optionen gelten nur für den aktuellen Stream. Klicken Sie auf diese Schaltfläche, um diese Optionen als Standard für alle Streams festzulegen.

Festlegen von Datums- und Zeitoptionen für Streams: Diese Optionen geben das für verschiedene Datums- und Zeitausdrücke im aktuellen Stream zu verwendende Format an.

Datum/Uhrzeit importieren als. Wählen Sie, ob der Speichertyp "Datum/Uhrzeit" für Datums-/Uhrzeitfelder verwendet werden soll oder ob sie als Zeichenfolgevariablen importiert werden sollen.

Datumsformat. Wählen Sie ein Datumsformat aus, das für die Datumsspeicherfelder verwendet werden soll oder wenn Zeichenfolgen von den CLEM-Datumsfunktionen als Datumsangaben interpretiert werden.

Zeitformat. Wählen Sie ein Zeitformat aus, das für die Zeitspeicherfelder verwendet werden soll oder wenn Zeichenfolgen von den CLEM-Zeitfunktionen als Zeitangaben interpretiert werden.

Tage/Minuten übertragen. Wählen Sie für Zeitformate aus, ob negative Zeitunterschiede so interpretiert werden, dass sie sich auf den vorherigen Tag bzw. die vorherige Stunde beziehen.

Datumsbasis (1. Jan.). Wählen Sie die Basisjahre (immer 1. Jan.), die von den CLEM-Datumsfunktionen verwendet werden müssen, die mit einem einzelnen Datum arbeiten.

Zweistellige Datumsangaben ab. Geben Sie das Schwellenjahr an, ab dem Jahrhundertziffern für Jahre hinzugefügt werden, die mit lediglich zwei Ziffern angegeben sind. Wenn Sie beispielsweise 1930 als Schwellenjahr festlegen, wird angenommen, dass das Datum 05/11/02 im Jahr 2002 liegt. Für Daten nach 30 wird das 20. Jahrhundert verwendet, d. h., bei 05/11/73 wird angenommen, dass es sich um ein Datum aus dem Jahr 1973 handelt.

Zeitzone Wählen Sie aus, wie die Zeitzone für die Verwendung mit dem CLEM-Ausdruck `datetime_now` ausgewählt wird.

- Wenn Sie **Server** auswählen, hängt die Zeitzone von den folgenden Elementen ab:
 - Wenn der aktuelle Stream eine Analytic Server-Datenquelle verwendet, verwendet der Ausdruck `datetime_now` die Uhrzeit von Analytic Server; der Server verwendet standardmäßig UTC-Zeit.
 - Wenn der aktuelle Stream einen Datenbankquellenknoten verwendet, verwenden die unterstützten Datenbanken SQL-Pushpack und der Ausdruck `datetime_now` verwendet die Uhrzeit der Datenbank.
 - Bei allen anderen Streams wird die Uhrzeit von SPSS Modeler Server für die Zeitzone verwendet.
- Wenn Sie **Modeler-Client** auswählen, spiegelt die Zeitzone die Zeitzonendetails des Computers wider, auf dem SPSS Modeler installiert ist.
- Alternativ können Sie einen beliebigen UTC-Wert für die Zeitzone auswählen.

Als Standard speichern. Die angegebenen Optionen gelten nur für den aktuellen Stream. Klicken Sie auf diese Schaltfläche, um diese Optionen als Standard für alle Streams festzulegen.

Festlegen von Zahlenformatoptionen für Streams: Diese Optionen geben das für verschiedene numerische Ausdrücke im aktuellen Stream zu verwendende Format an.

Zahlenanzeigeformat. Sie können aus den Anzeigeformaten Standard (####.###), Wissenschaftlich (#.###E+###) und Währung (\$###.##) wählen.

Dezimalstellen (Standard, Wissenschaftlich, Währung). Gibt bei Zahlenanzeigeformaten an, wie viele Dezimalstellen bei der Anzeige oder beim Drucken reeller Zahlen verwendet werden sollen. Diese Option wird getrennt für jedes Anzeigeformat angegeben.

Berechnungen in. Wählen Sie **Radiant** oder **Grad** als Maßeinheit für trigonometrische CLEM-Ausdrücke. Weitere Informationen finden Sie in „Trigonometrische Funktionen“ auf Seite 168.

Als Standard speichern. Die angegebenen Optionen gelten nur für den aktuellen Stream. Klicken Sie auf diese Schaltfläche, um diese Optionen als Standard für alle Streams festzulegen.

Festlegen von Optimierungsoptionen für Streams: Sie können die Optimierungseinstellungen verwenden, um die Streamleistung zu optimieren. Beachten Sie, dass die Leistungs- und Optimierungseinstellungen auf IBM SPSS Modeler Server (sofern verwendet) alle entsprechenden Einstellungen auf dem Client außer Kraft setzen. Wenn diese Einstellungen im Server inaktiviert sind, kann der Client sie nicht aktivieren. Sind sie auf dem Server aktiviert, kann der Client sie jedoch inaktivieren.

Anmerkung: Datenbankmodellierung und SQL-Optimierung erfordern, dass auf dem IBM SPSS Modeler-Computer IBM SPSS Modeler Server-Konnektivität aktiviert ist. Wenn diese Einstellung aktiviert ist, können Sie auf Datenbankalgorithmen zugreifen, SQL direkt aus IBM SPSS Modeler per Pushback übertragen und auf IBM SPSS Modeler Server zugreifen. Wählen Sie zur Überprüfung des aktuellen Lizenzstatus Folgendes im Menü von IBM SPSS Modeler aus.

Wenn Konnektivität aktiviert ist, wird auf der Registerkarte "Lizenzstatus" die Option **Serveraktivierung** angezeigt.

Weitere Informationen finden Sie in „Verbinden mit IBM SPSS Modeler Server“ auf Seite 10.

Anmerkung: Ob SQL-Pushback und Optimierung unterstützt werden, hängt vom verwendeten Datenbanktyp ab. Neueste Informationen zu Datenbanken und ODBC-Treibern, die für die Verwendung mit IBM SPSS Modeler unterstützt werden und getestet wurden, finden Sie auf der unternehmensweiten Support-Site unter <http://www.ibm.com/support>.

Streamneuerstellung aktivieren. Wählen Sie diese Option aus, um die Streamneuerstellung in IBM SPSS Modeler zu aktivieren. Es stehen vier Neuerstellungsarten zur Verfügung, von denen Sie eine oder mehrere auswählen können. Die Streamneuerstellung ordnet die Knoten in einem Stream für eine effizientere Operation im Hintergrund neu an, ohne die Streamsemantik zu ändern.

- **SQL-Generierung optimieren.** Mit dieser Option können die Knoten im Stream so neu angeordnet werden, dass ein Pushback weiterer Operationen mithilfe der SQL-Generierung für die Ausführung in der Datenbank möglich ist. Wenn die Optimierungsfunktion einen Knoten findet, der nicht in SQL wiedergegeben werden kann, sucht sie nachgeordnete Knoten, die in SQL wiedergegeben und ohne Beeinträchtigung der Streamsemantik vor den problematischen Knoten verschoben werden können. Nicht nur, dass die Datenbank die Operationen effizienter als IBM SPSS Modeler durchführen kann, solche Pushbacks dienen auch dazu, die Größe des Datensets, das an IBM SPSS Modeler zur Verarbeitung zurückgegeben wird, zu reduzieren. So wird wiederum der Netzverkehr reduziert und die Geschwindigkeit der Streamoperationen erhöht. Das Kontrollkästchen **SQL generieren** muss aktiviert sein, damit die SQL-Optimierung funktionieren kann.
- **CLEM-Ausdruck optimieren.** Mit dieser Option kann das Optimierungsprogramm nach CLEM-Ausdrücken suchen, die vor der Ausführung des Streams vorverarbeitet werden können, damit die Verarbeitungsgeschwindigkeit erhöht wird. Ein einfaches Beispiel dafür ist, dass das Optimierungsprogramm bei einem Ausdruck wie $\log(\text{Gehalt})$ den tatsächlichen Gehaltswert berechnen und diesen dann für die Verarbeitung übergeben würde. Dies kann die Leistung von SQL-Pushback und IBM SPSS Modeler Server erhöhen.
- **Syntaxausführung optimieren.** Diese Methode der Streamneuerstellung erhöht die Effizienz von Operationen, die mehr als einen Knoten mit IBM SPSS Statistics-Syntax umfassen. Die Optimierung wird durch die Zusammenfassung der Syntaxbefehle in einer einzigen Operation erreicht, anstatt jeden einzelnen Syntaxbefehl als separate Operation auszuführen.
- **Andere Ausführung optimieren.** Diese Methode der Streamneuerstellung erhöht die Effizienz von Operationen, die nicht an die Datenbank delegiert werden können. Eine Optimierung wird dadurch erreicht, dass die Datenmenge im Stream so früh wie möglich reduziert wird. Während die Datenintegrität gewahrt bleibt, wird der Stream neu erstellt, um die Operationen näher zur Datenquelle zu bringen. So werden die Daten abwärts für kostenintensive Operationen, z. B. Joins, reduziert.

Parallele Verarbeitung aktivieren. Bei Ausführung auf einem Computer mit mehreren Prozessoren ermöglicht diese Option dem System, die Last auf diese Prozessoren zu verteilen, was zu einer schnelleren Leistung führen kann. Durch die Verwendung mehrerer Knoten oder der folgenden Einzelknoten können Sie von der Parallelverarbeitung profitieren: C5.0, Zusammenführen (nach Schlüssel), Sortieren, Klasse (Rangbildungsmethode und N-Perzentil-Methode) sowie Aggregieren (unter Verwendung mindestens eines Schlüsselfelds).

SQL exportieren. Wählen Sie diese Option aus, um die SQL-Generierung zu aktivieren. Dabei wird zugelassen, dass Streamoperationen per Pushback an die Datenbank zurückübertragen werden, indem SQL-Code zur Generierung der Ausführungsprozesse verwendet wird, was zu einer Leistungssteigerung führen kann. Zur weiteren Steigerung der Leistung kann außerdem **SQL-Generierung optimieren** ausgewählt werden. Dadurch wird die Anzahl der Operationen, die per Pushback an die Datenbank zu-

rückübertragen werden, maximiert. Wenn Operationen für einen Knoten per Pushback an die Datenbank zurückübertragen wurden, wird der Knoten bei der Ausführung des Streams violett gekennzeichnet.

- **Datenbankcaching.** Bei Streams, die in der Datenbank auszuführendes SQL generieren, muss nicht das Dateisystem als Cache verwendet werden, sondern die Daten können mitten im Stream in einer temporären Tabelle in der Datenbank zwischengespeichert werden. Bei einer Kombination mit SQL-Optimierung kann dies zu einer erheblichen Leistungssteigerung führen. So kann beispielsweise die Ausgabe aus einem Stream, der mehrere Tabellen zusammenfasst, um eine Data-Mining-Ansicht zu erstellen, im Cache gespeichert und bei Bedarf wieder verwendet werden. Klicken Sie bei aktiviertem Datenbankcaching einfach mit der rechten Maustaste auf einen Knoten (keinen Endknoten), um die Daten an dieser Stelle im Cache zu speichern, und der Cache wird bei der nächsten Ausführung des Streams automatisch unmittelbar in der Datenbank erstellt. Dadurch kann SQL für nachgeordnete Knoten generiert werden, was zu weiteren Leistungsverbesserungen führt. (Alternativ kann diese Option bei Bedarf inaktiviert werden, beispielsweise wenn es aufgrund von Richtlinien oder Berechtigungen nicht möglich ist, Daten in die Datenbank zu schreiben. Wenn Datenbankcaching oder SQL-Optimierung nicht aktiviert ist, wird der Cache stattdessen in das Dateisystem geschrieben. Weitere Informationen finden Sie in „Caching-Optionen für Knoten“ auf Seite 41.
- **Unproblematische Umwandlung verwenden.** Mit dieser Option können Daten von Zeichenfolgen in Zahlen oder von Zahlen in Zeichenfolgen konvertiert werden, sofern sie in einem geeigneten Format gespeichert sind. Wenn die Daten beispielsweise als Zeichenfolge in der Datenbank gespeichert sind, tatsächlich jedoch eine sinnvolle Zahl enthalten, können die Daten beim Eintreten des Pushbacks zur Verwendung konvertiert werden.

Anmerkung: Aufgrund kleinerer Unterschiede in der SQL-Implementierung weichen die Ergebnisse der in einer Datenbank ausgeführten Streams möglicherweise leicht von den Ergebnissen bei Ausführung in IBM SPSS Modeler ab. Aus ähnlichen Gründen können diese Unterschiede auch je nach Datenbankanbieter variieren.

Als Standard speichern. Die angegebenen Optionen gelten nur für den aktuellen Stream. Klicken Sie auf diese Schaltfläche, um diese Optionen als Standard für alle Streams festzulegen.

Festlegen von Optionen für SQL-Protokollierung und Datensatzstatus für Streams: Zu diesen Einstellungen gehören verschiedene Optionen zur Festlegung der Anzeige der vom Stream generierten SQL-Anweisungen sowie der Anzeige für die Anzahl der vom Stream verarbeiteten Datensätze.

SQL wird während der Streamausführung im Nachrichtenprotokoll angezeigt. Gibt an, ob das während der Ausführung des Streams generierte SQL an das Nachrichtenprotokoll übergeben wird.

Einzelheiten zur SQL-Generierung während der Streamvorbereitung im Nachrichtenprotokoll anzeigen. Legt während der Streamvorschau fest, ob eine Vorschau des SQL, das generiert werden würde, an das Nachrichtenprotokoll übergeben wird.

SQL anzeigen. Gibt an, ob ein im Protokoll angezeigtes SQL systemeigene SQL-Funktionen oder Standard-ODBC-Funktionen in der Form {fn FUNC(...)}, wie von SPSS Modeler generiert, enthalten soll. Ersteres richtet sich nach den ODBC-Treiberfunktionen, die u. U. nicht implementiert werden.

SQL zur Verbesserung der Lesbarkeit neu formatieren. Gibt an, ob im Protokoll angezeigtes SQL zur Verbesserung der Lesbarkeit formatiert werden soll.

Status für Datensätze anzeigen. Gibt an, ob Datensätze gemeldet werden sollen, wenn sie an Endknoten ankommen. Geben Sie eine Zahl an, um den Status alle N Datensätze zu aktualisieren.

Als Standard speichern. Die angegebenen Optionen gelten nur für den aktuellen Stream. Klicken Sie auf diese Schaltfläche, um diese Optionen als Standard für alle Streams festzulegen.

Festlegen von Layoutoptionen für Streams: Diese Einstellungen bieten eine Reihe von Optionen, die sich auf Anzeige und Verwendung des Streamerstellungsbereichs beziehen.

Minimale Breite des Streamerstellungsbereichs. Geben Sie die minimale Breite des Streamerstellungsbereichs in Pixeln an.

Minimale Höhe des Streamerstellungsbereichs. Geben Sie die minimale Höhe des Streamerstellungsbereichs in Pixeln an.

Streamblättersgeschwindigkeit. Geben Sie die Blättersgeschwindigkeit für den Streamerstellungsbereich an, um die Blättersgeschwindigkeit des Streamerstellungsbereichs zu steuern, wenn ein Knoten von einer Position im Erstellungsbereich an eine andere gezogen wird. Je höher die Zahl, desto höher die Blättersgeschwindigkeit.

Maximum für Symbolnamen. Geben Sie eine Begrenzung in der Anzahl Zeichen für die Namen von Knoten im Streamerstellungsbereich ein.

Symbolgröße. Wählen Sie eine Option aus, um die gesamte Streamansicht auf eine Anzahl von Größenwerten zwischen 8 % und 200 % der Standardsymbolgröße zu skalieren.

Rasterzellengröße. Wählen Sie eine Rasterzellengröße in der Liste aus. Diese Zahl wird für die Ausrichtung von Knoten im Streamerstellungsbereich mithilfe eines unsichtbaren Rasters verwendet. Die Standardrasterzellengröße lautet 0,25.

Am Raster ausrichten. Wählen Sie diese Option aus, um Symbole an einem unsichtbaren Rastermuster auszurichten (Standardauswahl).

Generierte Symbolplatzierung. Wählen Sie im Erstellungsbereich die gewünschte Position für Symbole der Knoten, die aus Modellnuggets generiert wurden. Standard ist oben links.

Als Standard speichern. Die angegebenen Optionen gelten nur für den aktuellen Stream. Klicken Sie auf diese Schaltfläche, um diese Optionen als Standard für alle Streams festzulegen.

Analytic Server-Streamereigenschaften: Diese Einstellungen bieten eine Reihe von Optionen für die Arbeit mit Analytic Server.

Maximale Anzahl der Datensätze, die außerhalb von Analytic Server zu verarbeiten sind

Geben Sie die maximale Anzahl der Datensätze an, die aus einer Analytic Server-Datenquelle auf den SPSS Modeler-Server importiert werden sollen.

Nachricht, wenn ein Knoten nicht in Analytic Server verarbeitet werden kann

Diese Einstellung bestimmt, was geschieht, wenn ein Stream, der an Analytic Server übergeben würde, einen Knoten enthält, der in Analytic Server nicht verarbeitet werden kann. Geben Sie an, ob eine Warnung abgesetzt und die Verarbeitung des Streams fortgesetzt werden soll oder ob ein Fehler ausgelöst und die Verarbeitung gestoppt werden soll.

Speichereinstellungen für aufgeteilte Modelle

Modell nach Verweis in Analytic Server speichern, wenn die Modellgröße (MB) übersteigt

Modellnuggets werden in der Regel als Teil des Streams gespeichert. Aufgeteilte Modelle mit vielen Aufteilungen können große Nuggets erzeugen und das Vor- und Zurückschieben des Nuggets zwischen dem Stream und Analytic Server kann sich auf die Leistung auswirken. Wenn ein aufgeteiltes Modell die angegebene Größe überschreitet, wird es daher in Analytic Server gespeichert und das Nugget in SPSS Modeler enthält einen Verweis auf das Modell.

Standardordner, in dem Modelle nach Verweis in Analytic Server gespeichert werden, wenn die Ausführung beendet ist

Geben Sie den Standardpfad an, unter dem Sie aufgeteilte Modelle in Analytic Server speichern wollen. Der Pfad sollte mit einem gültigen Analytic Server-Projektnamen beginnen.

Ordner zum Speichern hochgestufter Modelle

Geben Sie den Standardpfad an, unter dem Sie "hochgestufte" Modelle speichern wollen. Ein hochgestuftes Modell wird nicht bereinigt, wenn die SPSS Modeler-Sitzung vorüber ist.

Festlegen georäumlicher Optionen für Streams: Jedem georäumlichen Feld, ob Form, Koordinate oder einzelner Achsenwert (z. B. x oder y, Breitengrad oder Längengrad) ist ein Koordinatensystem zugeordnet. Dieses Koordinatensystem legt Attribute wie den Ursprung (0,0) und die den Werten zugeordneten Einheiten fest.

Es steht eine Anzahl von Koordinatensystemen zur Verfügung und es gibt die beiden Typen "geografisch" und "projiziert". Alle räumlichen Funktionen in SPSS Modeler können nur mit einem projizierten Koordinatensystem verwendet werden.

Bei Koordinatensystemen wird vorausgesetzt, dass für das Zusammenführen oder Anhängen von Daten aus zwei getrennten Quellen georäumlicher Daten die Quellen dasselbe Koordinatensystem verwenden müssen. Daher müssen Sie eine Koordinateneinstellung für alle georäumlichen Daten angeben, die im Stream verwendet werden.

Daten werden in den folgenden Situationen automatisch umprojiziert, sodass sie das ausgewählte Streamkoordinatensystem verwenden:

- Für räumliche Funktionen (wie "area", "closeto" oder "within") wird der an die Funktion übergebene Parameter automatisch umprojiziert. Die ursprünglichen Zeilendaten bleiben allerdings unverändert.
- Beim Verwenden der Erstellung oder des Scorings von (Nugget-)Knoten bei STP (Spatio-Temporal Prediction - räumliche temporale Vorhersage) wird das Speicherortfeld automatisch umprojiziert. Beim Scoring ist der Speicherort, der aus dem Nugget kommt, der ursprüngliche Speicherort.
- Beim Verwenden des Kartenvisualisierungsknotens.

Streamkoordinatensystem. Nur verfügbar, wenn Sie das Kontrollkästchen auswählen. Klicken Sie auf **Ändern**, um eine Liste verfügbarer projizierter Koordinatensysteme anzuzeigen und das für den aktuellen Stream zu verwendende auszuwählen.

Als Standard speichern. Das Koordinatensystem, das Sie auswählen, gilt nur für den aktuellen Stream. Klicken Sie auf diese Schaltfläche, wenn Sie das System als Standardwert für alle Streams auswählen wollen.

Auswählen georäumlicher Koordinatensysteme: Alle räumlichen Funktionen in SPSS Modeler können nur mit einem projizierten Koordinatensystem verwendet werden.

Das Dialogfeld zum Auswählen des Streamkoordinatensystems enthält eine Liste aller projizierten Koordinatensysteme, die Sie für alle georäumlichen Daten auswählen können, die in einem Stream verwendet werden.

Die folgenden Informationen werden für jedes Koordinatensystem aufgelistet.

- **WKID** Die WKID (Well Known ID), die für jedes Koordinatensystem eindeutig ist.
- **Name** Der Name des Koordinatensystems.
- **Einheiten** Die Maßeinheit, die dem Koordinatensystem zugeordnet ist.

Zusätzlich zur Liste aller Koordinatensysteme verfügt das Dialogfeld über eine Steuerung für das **Filtern**. Wenn Sie den gesamten Namen oder einen Teil des Namens des gewünschten Koordinatensystems kennen, geben Sie ihn im Feld **Name** unten im Dialogfeld ein. Die Liste von Koordinatensystemen, aus denen Sie auswählen können, wird automatisch gefiltert, sodass nur die Systeme mit Namen angezeigt werden, die den von Ihnen eingegebenen Text enthalten.

Anzeigen von Nachrichten zu Streamoperationen

Nachrichten zu Streamoperationen wie Ausführung, Optimierung und vergangene Zeit für die Modellerstellung und -evaluierung können problemlos mithilfe der Registerkarte "Nachrichten" im Dialogfeld "Streameigenschaften" angezeigt werden. In dieser Tabelle werden auch Fehlernachrichten gespeichert.

Anzeigen von Streamnachrichten

1. Klicken Sie im Menü "Datei" auf die Option **Streameigenschaften** (oder wählen Sie den Stream auf der Registerkarte "Streams" im Managerbereich aus, klicken Sie mit der rechten Maustaste und wählen Sie dann im Popup-Menü die Option **Streameigenschaften** aus).
2. Klicken Sie auf die Registerkarte **Nachrichten**.

Klicken Sie als Alternative im Menü "Tools" auf:

Streameigenschaften > Nachrichten

Zusätzlich zu den Nachrichten zu Streamoperationen werden hier Fehlernachrichten gespeichert. Wenn die Streamausführung aufgrund eines Fehlers beendet wird, wird in diesem Dialogfeld die Registerkarte "Nachrichten" mit der sichtbaren Fehlernachricht geöffnet. Außerdem wird der Knoten mit den Fehlern im Streamerstellungsbereich rot markiert.

Falls die Optionen für die SQL-Optimierung und Protokollierung im Dialogfeld "Benutzeroptionen" aktiviert sind, werden auch Informationen zum generierten SQL angezeigt. Weitere Informationen finden Sie in „Festlegen von Optimierungsoptionen für Streams“ auf Seite 47.

Sie können hier für einen Stream gemeldete Nachrichten speichern, indem Sie in der Dropdown-Liste der Schaltfläche "Speichern" (links, direkt unter dem Register "Nachrichten") auf **Nachrichten speichern** klicken.

Sie können alle Nachrichten für einen bestimmten Stream löschen, indem Sie in der Dropdown-Liste **Alle Nachrichten löschen** auswählen.

Die CPU-Zeit bezeichnet die Zeit, während derer der Serverprozess die CPU belegt. Die abgelaufene Zeit ist die Gesamtzeit zwischen Beginn und Ende der Ausführung und umfasst daher z. B. auch die Übertragung von Dateien sowie das Rendering von Ausgaben. Die CPU-Zeit kann höher als die abgelaufene Zeit sein, wenn ein Stream mehrere CPUs verwendet (parallele Ausführung). Wenn ein Stream komplett auf die als Datenquelle verwendete Datenbank zurückgreift, ist die CPU-Zeit null.

Anzeigen der Knotenausführungszeiten

Auf der Registerkarte **Nachrichten** können Sie auswählen, dass die Registerkarte **Ausführungszeiten** angezeigt werden soll, auf der die einzelnen Ausführungszeiten für alle Knoten im Stream angezeigt werden, die in IBM SPSS Modeler Server ausgeführt werden. Die Ausführungszeiten sind für Streams, die in anderen Bereichen, beispielsweise R oder Analytic Server, ausgeführt werden, möglicherweise nicht genau.

Anmerkung: Damit diese Funktion ausgeführt werden kann, muss auf der Registerkarte **Optionen** in der Einstellung **Allgemein** das Kontrollkästchen **Ausführungszeiten anzeigen** aktiviert sein.

In der Tabelle der Knotenausführungszeiten gibt es die unten angegebenen Spalten. Klicken Sie auf eine Spaltenüberschrift, um die Einträge in aufsteigender bzw. absteigender Reihenfolge zu sortieren (z. B. um anzuzeigen, welche Knoten die längsten Ausführungszeiten haben).

Endknoten. Die ID der Verzweigung, zu der der Knoten gehört. Die ID ist der Name des Endknotens am Ende der Verzweigung.

Knotenbeschriftung. Der Name des Knotens, auf den sich die Ausführungszeit bezieht.

Knoten-ID. Die eindeutige ID des Knotens, auf den sich die Ausführungszeit bezieht. Diese ID wird bei der Knotenerstellung vom System generiert.

Ausführungszeit(en). Die Zeitdauer der Ausführung dieses Knotens in Sekunden.

Festlegen der Stream- und Sitzungsparameter

Parameter können für die Verwendung in CLEM-Ausdrücken und Scripts definiert werden. Es handelt sich dabei im Grunde um benutzerdefinierte Variablen, die mit dem aktuellen Stream, der aktuellen Sitzung bzw. dem aktuellen Superknoten gespeichert und persistent gemacht werden. Ein Zugriff auf diese Variablen ist über die Benutzerschnittstelle und über Scripts möglich. Wenn Sie also einen Stream speichern, werden alle für diesen Stream festgelegten Parameter ebenfalls gespeichert. (Dies unterscheidet sie von lokalen Scriptvariablen, die nur in dem Script verwendet werden können, in dem sie deklariert sind.) Parameter werden oft beim Scripting verwendet, um das Verhalten des Scripts zu steuern, indem Informationen zu Feldern und Werten angegeben werden, die nicht im Script fest codiert werden müssen.

Der Bereich eines Parameters hängt davon ab, wo er festgelegt wird:

- Streamparameter können in einem Stream-Script oder im Dialogfeld "Streameigenschaften" festgelegt werden und sie stehen für alle Knoten im Stream zur Verfügung. Sie werden in der Parameterliste im Expression Builder angezeigt.
- Sitzungsparameter können in einem Standalone-Script oder im Dialogfeld "Sitzungsparameter" festgelegt werden. Sie stehen allen Streams zur Verfügung, die in der aktuellen Sitzung verwendet werden (alle Streams auf der Registerkarte "Streams" im Managerbereich).

Parameter können auch für Superknoten festgelegt werden. In diesem Fall sind sie nur für Knoten sichtbar, die in dem betreffenden Superknoten gekapselt sind.

So können Sie die Stream- und Sitzungsparameter über die Benutzerschnittstelle festlegen:

1. Zum Festlegen der Streamparameter klicken Sie im Hauptmenü auf:
Tools > Streameigenschaften > Parameter
2. Zum Festlegen der Sitzungsparameter klicken Sie im Menü "Tools" auf die Option **Sitzungsparameter festlegen**.

Eingabeaufforderung? Markieren Sie dieses Kontrollkästchen, wenn der Benutzer während der Laufzeit zur Eingabe eines Werts für diesen Parameter aufgefordert werden soll.

Name. Hier werden die Parameternamen aufgelistet. Sie können einen neuen Parameter erstellen, indem Sie in diesem Feld einen Namen eingeben. Um beispielsweise einen Parameter für die Mindesttemperatur zu erstellen, könnten Sie `minvalue` eingeben. Verwenden Sie nicht das Präfix `$P-`, das Parameter in CLEM-Ausdrücken kennzeichnet. Dieser Name wird auch zur Anzeige im CLEM Expression Builder verwendet.

Langname. Listet den beschreibenden Namen für die einzelnen erstellten Parameter auf.

Speichertyp. Wählen Sie einen Speichertyp aus der Liste aus. Der Speichertyp gibt an, wie die Datenwerte im Parameter gespeichert werden. Wenn Sie z. B. mit Werten arbeiten, die führende Nullen enthalten und die Sie beibehalten möchten (wie 008), sollten Sie **Zeichenfolge** als Speichertyp wählen. Andernfalls werden die Nullen aus dem Wert entfernt. Verfügbare Speichertypen sind "Zeichenfolge", "Ganzzahl", "Reelle Zahl", "Uhrzeit", "Datum" und "Zeitmarke". Beachten Sie, dass bei Datumsparametern die Werte gemäß der im nächsten Absatz erläuterten ISO-Standardnotation eingegeben werden müssen.

Wert. Listet den aktuellen Wert für die einzelnen Parameter auf. Ändern Sie den Parameter wie gewünscht. Datumsparameter müssen in ISO-Standardnotation angegeben werden (d. h. in der Form JJJJ-MM-TT). Datumsangaben in anderen Formaten sind nicht zulässig.

Typ (optional). Wenn Sie den Stream für eine externe Anwendung bereitstellen möchten, wählen Sie aus der Liste ein Messniveau aus. Andernfalls sollten Sie die Spalte *Typ* so belassen, wie sie ist. Wenn Sie Wertbeschränkungen für den Parameter festlegen möchten, z. B. die Ober- und Untergrenze für einen numerischen Bereich, wählen Sie **Angeben** aus der Liste aus.

Die Optionen "Langname", "Speichertyp" und "Typ" können für Parameter nur über die Benutzerschnittstelle festgelegt werden. Die Festlegung dieser Optionen mithilfe von Scripts ist nicht möglich.

Klicken Sie auf die Pfeile rechts, um den ausgewählten Parameter in der Liste verfügbarer Parameter weiter nach oben oder weiter nach unten zu verschieben. Verwenden Sie die Schaltfläche zum Löschen (mit einem X markiert), um den ausgewählten Parameter zu entfernen.

Festlegen von Eingabeaufforderungen für Parameterwerte während der Laufzeit

Wenn Sie in manchen Streams zu verschiedenen Gelegenheiten für denselben Parameter unterschiedliche Werte eingeben müssen, können Sie Laufzeiteingabeaufforderungen für einen oder mehrere Stream- oder Sitzungsparameterwerte festlegen.

Parameter. (Optional) Geben Sie einen Wert für den Parameter ein oder belassen Sie den Standardwert, sofern vorhanden.

Diese Eingabeaufforderungen inaktivieren. Aktivieren Sie dieses Kontrollkästchen, wenn diese Eingabeaufforderungen beim Ausführen des Streams nicht angezeigt werden sollen. Sie können sie wieder anzeigen, indem Sie das Kontrollkästchen **Eingabeaufforderung?** im Dialogfeld für Stream- oder Sitzungseigenschaften, in dem die Parameter definiert wurden, aktivieren. Weitere Informationen finden Sie in „Festlegen der Stream- und Sitzungsparameter“ auf Seite 53.

Festlegen von Wertbeschränkungen für einen Parametertyp

Sie können Wertbeschränkungen für einen Parameter während einer Streambereitstellung für eine externe Anwendung, die Datenmodellierungsstreams liest, zur Verfügung stellen. In diesem Dialogfeld können Sie die Werte angeben, die einem externen Benutzer, der den Stream ausführt, zur Verfügung stehen. Je nach Datentyp variieren die Wertbeschränkungen im Dialogfeld dynamisch. Die hier angezeigten Optionen stimmen mit den Optionen überein, die für Werte vom Typknoten zur Verfügung stehen.

Typ. Zeigt das aktuell ausgewählte Messniveau an. Sie können diesen Wert so ändern, wie der Parameter in IBM SPSS Modeler verwendet werden soll.

Speichertyp. Zeigt den Speichertyp an, sofern dieser bekannt ist. Die Speichertypen werden nicht von dem für die Arbeit in IBM SPSS Modeler ausgewählten Messniveau (stetig, nominal oder Flag) beeinflusst. Sie können den Speichertyp auf der Hauptregisterkarte "Parameter" ändern.

Die untere Hälfte des Dialogfelds ändert sich dynamisch in Abhängigkeit vom im Feld **Typ** ausgewählten Messniveau.

Stetige Messniveaus

Minimum. Geben Sie eine Untergrenze für die Parameterwerte ein.

Maximum. Geben Sie eine Obergrenze für die Parameterwerte ein.

Beschriftungen. Sie können Beschriftungen für jeden Wert eines Bereichsfelds angeben. Klicken Sie auf die Schaltfläche **Beschriftungen**, um ein gesondertes Dialogfeld zur Angabe der Wertbeschriftungen zu öffnen.

Nominale Messniveaus

Werte. Mit dieser Option können Sie Werte für einen Parameter angeben, der als nominales Feld verwendet wird. Die Werte werden im IBM SPSS Modeler-Stream nicht erzwungen, werden jedoch in einer Dropdown-Liste für externe Bereitstellungsanwendungen verwendet. Mit den Pfeilschaltflächen und der Löschschaftfläche können Sie bestehende Werte bearbeiten sowie Werte neu sortieren und löschen.

Flagmessniveaus

Wahr. Geben Sie einen Flagwert für den Parameter an, wenn die Bedingung erfüllt ist.

Falsch. Geben Sie einen Flagwert für den Parameter an, wenn die Bedingung nicht erfüllt ist.

Beschriftungen. Sie können Beschriftungen für die Werte eines Flagfelds angeben.

Bereitstellungsoptionen für den Stream

Auf der Registerkarte **Bereitstellung** im Dialogfeld für die Streameigenschaften können Sie Optionen zum Bereitstellen des Streams in IBM SPSS Collaboration and Deployment Services zur Modellaktualisierung, für automatische Aufgabenplanung oder weitere Verwendung durch IBM Analytical Decision Management angeben. Alle Streams benötigen eine spezielle Scoring-Verzweigung, bevor sie bereitgestellt werden können. Weitere Informationen finden Sie in „Speichern und Bereitstellen von Repository-Objekten“ auf Seite 192.

Ausführung unter Verwendung von Schleifen für Streams

Auf der Registerkarte "Ausführung" im Dialogfeld "Streameigenschaften" können Sie Bedingungen für die Verwendung von Schleifen festlegen, um sich wiederholende Aufgaben für den aktuellen Stream zu automatisieren.

Nachdem Sie diese Bedingungen festgelegt haben, können Sie diese Registerkarte als Einführung in das Scripting verwenden, da das Scriptfenster mit grundlegenden Scripting-Angaben für Ihren Stream ausgefüllt wird, die Sie dann ändern können, um sie vielleicht als Basis für die Erstellung besserer Scripts zu verwenden. Weitere Informationen finden Sie in „Globale Funktionen“ auf Seite 187.

So legen Sie die Verwendung von Schleifen für einen Stream fest:

1. Klicken Sie im Menü "Datei" auf die Option **Streameigenschaften** (oder wählen Sie den Stream auf der Registerkarte "Streams" im Managerbereich aus, klicken Sie mit der rechten Maustaste und wählen Sie dann im Popup-Menü die Option **Streameigenschaften** aus).
2. Klicken Sie auf die Registerkarte **Ausführung**.
3. Wählen Sie den Ausführungsmodus **Verwendung von Schleifen/Bedingte Ausführung** aus.
4. Klicken Sie auf die Registerkarte **Verwendung von Schleifen**.

Klicken Sie als Alternative im Menü "Tools" auf:

Streameigenschaften > Ausführung

Als weitere Alternative klicken Sie mit der rechten Maustaste auf den Knoten und klicken Sie im Kontextmenü auf Folgendes:

Verwendung von Schleifen/Bedingte Ausführung > Einstellung für die Verwendung von Schleifen bearbeiten

Iteration. Sie können diesen Zeilennummernwert nicht bearbeiten, Sie können eine Iteration jedoch mit den Schaltflächen auf der rechten Seite der Tabelle hinzufügen, löschen oder nach oben oder unten verschieben.

Tabellenkopfzeilen. Diese geben den Iterationsschlüssel und alle Iterationsvariablen an, sofern vorhanden, die Sie beim Definieren der Schleife erstellt haben.

Anzeigen von globalen Werten für Streams

Mit der Registerkarte "Globalwerte" im Dialogfeld "Streameigenschaften" können Sie die globalen, für den aktuellen Stream festgelegten Werte anzeigen. Globalwerte werden mit einem Knoten vom Typ "Globalwerte" erstellt, um Statistiken zu ermitteln, wie Mittelwert, Summe oder Standardabweichung für ausgewählte Felder.

Sobald der Knoten "Globalwerte" ausgeführt wird, stehen diese Werte für eine Vielzahl von Anwendungen in Streamoperationen zur Verfügung. Weitere Informationen finden Sie in „Globale Funktionen“ auf Seite 187.

So zeigen Sie Globalwerte für einen Stream an:

1. Klicken Sie im Menü "Datei" auf die Option **Streameigenschaften** (oder wählen Sie den Stream auf der Registerkarte "Streams" im Managerbereich aus, klicken Sie mit der rechten Maustaste und wählen Sie dann im Popup-Menü die Option **Streameigenschaften** aus).
2. Klicken Sie auf die Registerkarte **Globalwerte**.

Klicken Sie als Alternative im Menü "Tools" auf:

Streameigenschaften > Globalwerte

Verfügbare Globalwerte. Die verfügbaren Globalwerte werden in dieser Tabelle aufgelistet. Hier können Sie keine Globalwerte bearbeiten, aber alle Globalwerte für einen Stream unter Verwendung der Schaltfläche "Alle Werte löschen" rechts in der Tabelle löschen.

Suchen nach Knoten in einem Stream

Sie können Knoten in einem Stream suchen, indem Sie eine Reihe von Suchkriterien wie Knotenname, Kategorie und ID angeben. Diese Funktion ist besonders nützlich für komplexe Streams, die eine große Anzahl Knoten enthalten.

So suchen Sie Knoten in einem Stream:

1. Klicken Sie im Menü "Datei" auf die Option **Streameigenschaften** (oder wählen Sie den Stream auf der Registerkarte "Streams" im Managerbereich aus, klicken Sie mit der rechten Maustaste und wählen Sie dann im Popup-Menü die Option **Streameigenschaften** aus).
2. Klicken Sie auf die Registerkarte **Suchen**.

Klicken Sie als Alternative im Menü "Tools" auf:

Streameigenschaften > Suchen

Sie können mehrere Optionen zur Einschränkung der Suche angeben. Einzig die Suche nach Knoten-ID (mithilfe des Felds **ID ist gleich**) schließt die anderen Optionen aus.

Knotenbeschriftung enthält. Markieren Sie dieses Kontrollkästchen und geben Sie eine Knotenbeschriftung ganz oder teilweise für die Suche nach einem bestimmten Knoten ein. Suchvorgänge unterscheiden Groß-/Kleinschreibung nicht und mehrere Wörter werden als ein einziges Textstück betrachtet.

Knotenkategorie. Aktivieren Sie dieses Kontrollkästchen und klicken Sie in der Liste auf eine Kategorie, um nach einem bestimmten Knotentyp zu suchen. **Prozessknoten** bedeutet einen Knoten aus der Registerkarte "Datensatzoperationen" oder "Feldoperationen" der Knotenpalette; **Modellknoten anwenden** bezieht sich auf ein Modellnugget.

Schlüsselwörter enthalten. Aktivieren Sie dieses Kontrollkästchen und geben Sie eines oder mehrere vollständige Schlüsselwörter ein, um Knoten zu suchen, für die diese Schlüsselwörter auf der Registerkarte "Anmerkungen" im Knotendialogfeld eingegeben wurden. Der eingegebene Schlüsselworttext muss eine exakte Übereinstimmung sein. Trennen Sie mehrere Schlüsselwörter durch Semikolons, um nach Alternativen

tiven zu suchen (z. B. werden mit proton;neutron alle Knoten mit einem dieser Schlüsselwörter gefunden). Weitere Informationen finden Sie in „Anmerkungen“ auf Seite 64.

Anmerkung enthält. Aktivieren Sie dieses Kontrollkästchen und geben Sie eines oder mehrere Wörter ein, um Knoten zu suchen, die diesen Text im Haupttextbereich der Registerkarte "Anmerkungen" im Knotendialogfeld enthalten. Suchvorgänge unterscheiden Groß-/Kleinschreibung nicht und mehrere Wörter werden als ein einziges Textstück betrachtet. Weitere Informationen finden Sie in „Anmerkungen“ auf Seite 64.

Feld generieren mit Name. Markieren Sie dieses Kontrollkästchen und geben Sie den Namen eines generierten Felds ein (zum Beispiel \$C-Drug). Mithilfe dieser Option können Sie Modellierungsknoten suchen, die ein bestimmtes Feld generieren. Geben Sie nur einen Feldnamen ein, der eine exakte Übereinstimmung sein muss.

ID ist gleich. Markieren Sie dieses Kontrollkästchen und geben Sie eine Knoten-ID ein, um einen bestimmten Knoten mit dieser Kennung zu finden. (Durch Auswählen dieser Option werden alle vorherigen Optionen inaktiviert.) Knoten-IDs werden beim Erstellen eines Knotens vom System zugewiesen. Sie können als Verweis auf den Knoten zur Scripterstellung oder Automatisierung verwendet werden. Geben Sie nur eine Knoten-ID ein, die eine exakte Übereinstimmung sein muss. Weitere Informationen finden Sie in „Anmerkungen“ auf Seite 64.

In Superknoten suchen. Dieses Kontrollkästchen ist standardmäßig markiert, d. h., die Suche erfolgt an Knoten innerhalb und außerhalb von Superknoten. Entfernen Sie die Markierung aus dem Kontrollkästchen, wenn Sie die Suche nur an Knoten außerhalb von Superknoten auf der oberen Streamebene ausführen möchten.

Suchen. Wenn Sie alle gewünschten Optionen festgelegt haben, klicken Sie auf diese Schaltfläche, um den Suchvorgang zu starten.

Knoten, die den angegebenen Optionen entsprechen, werden im unteren Bereich des Dialogfelds aufgelistet. Wählen Sie einen Knoten in der Liste aus, um ihn im Streamerstellungsbereich hervorzuheben.

Umbenennen von Streams

Auf der Registerkarte "Anmerkungen" im Dialogfeld "Streameigenschaften" können Sie beschreibende Anmerkungen für einen Stream hinzufügen und einen benutzerdefinierten Namen für den Stream erstellen. Diese Optionen sind besonders hilfreich, wenn Sie Berichte für Streams generieren, die dem Projektbereich hinzugefügt werden. Weitere Informationen finden Sie in „Anmerkungen“ auf Seite 64.

Streambeschreibungen

Für jeden erstellten Stream generiert IBM SPSS Modeler eine Streambeschreibung, die Informationen zum Inhalt des Streams enthält. Diese Beschreibungen sind hilfreich, wenn Sie die Funktion eines Streams interessiert, Sie jedoch IBM SPSS Modeler nicht installiert haben (z. B. wenn Sie über IBM SPSS Collaboration and Deployment Services auf einen Stream zugreifen).

Die Streambeschreibung wird als HTML-Dokument mit verschiedenen Abschnitten angezeigt.

Allgemeine Streaminformationen

Dieser Abschnitt enthält den Streamnamen sowie Details über seine Erstellung und die letzte Speicherung.

Beschreibung und Kommentare

Dieser Abschnitt enthält Folgendes:

- Streamanmerkungen (siehe „Anmerkungen“ auf Seite 64)

- Nicht mit bestimmten Knoten verknüpfte Kommentare
- Kommentare, die sowohl in der Modellerstellungs- als auch der Scoring-Verzweigung des Streams mit Knoten verknüpft sind

Scoring-Informationen

Dieser Abschnitt enthält unter mehreren Überschriften Informationen zur Scoring-Verzweigung des Streams.

- **Kommentare.** Enthält nur Kommentare zu Knoten in der Scoring-Verzweigung.
- **Eingaben.** Listet die Eingabefelder zusammen mit ihren Speichertypen auf (z. B. "Zeichenfolge", "Ganze Zahl", "Reelle Zahl" usw.).
- **Ausgaben.** Listet die Ausgabefelder inklusive der zusätzlichen, vom Modellierungsknoten erstellten Felder zusammen mit ihren Speichertypen auf.
- **Parameter.** Listet alle Parameter auf, die sich auf die Scoring-Verzweigung des Streams beziehen und die bei jedem Modellscoring angezeigt oder bearbeitet werden können. Diese Parameter werden ermittelt, wenn Sie auf die Schaltfläche **Scoring-Parameter** auf der Registerkarte **Bereitstellung** des Dialogfelds für die Streameigenschaften klicken.
- **Modellknoten.** Zeigt den Namen und den Typ des Modells an (z. B. "Neuronales Netz", "C&R-Baum" usw.). Dieses Modellnugget ist für das Feld **Modellknoten** auf der Registerkarte **Bereitstellung** des Dialogfelds für die Streameigenschaften ausgewählt.
- **Modelldetails.** Zeigt Details zum unter der vorherigen Überschrift identifizierten Modellnugget an. Wenn möglich werden für das Modell auch die Bedeutsamkeit der Prädiktoren und Evaluierungsdigramme angegeben.

Modellinformationen

Enthält Informationen zur Modellierungsverzweigung des Streams.

- **Kommentare.** Listet alle Kommentare oder Anmerkungen zu den Knoten in der Modellierungsverzweigung auf.
- **Eingaben.** Listet die Eingabefelder inklusive ihrer Rolle in der Modellierungsverzweigung auf (in Form des Werts für die Feldrolle, z. B. "Eingabe", "Ziel", "Aufteilung" usw.).
- **Parameter.** Listet alle Parameter auf, die sich auf die Modellierungsverzweigung des Streams beziehen und die bei jeder Aktualisierung des Modells angezeigt oder bearbeitet werden können. Diese Parameter werden ermittelt, wenn Sie auf die Schaltfläche **Modellerstellungsparameter** auf der Registerkarte **Bereitstellung** des Dialogfelds für die Streameigenschaften klicken.
- **Modellierungsknoten.** Zeigt Name und Typ des Modellknotens an, der zur Erstellung oder Aktualisierung des Modells verwendet wird.

Vorschau der Streambeschreibungen

Sie können den Inhalt einer Streambeschreibung in einem Web-Browser anzeigen, indem Sie im Dialogfeld für die Streameigenschaften auf eine Option klicken. Der Inhalt der Beschreibung hängt von den auf der Registerkarte "Bereitstellung" des Dialogfelds angegebenen Optionen ab. Weitere Informationen finden Sie in „Bereitstellungsoptionen für den Stream“ auf Seite 205.

So zeigen Sie eine Streambeschreibung an:

1. Klicken Sie im IBM SPSS Modeler-Hauptmenü auf:
Tools > Streameigenschaften > Bereitstellung
2. Geben Sie Bereitstellungstyp, den zugehörigen Scoring-Knoten sowie eventuelle Scoring-Parameter an.
3. Für den Bereitstellungstyp "Modellaktualisierung" können Sie optional Folgendes auswählen:
 - Modellierungsknoten und eventuelle Modellerstellungsparameter
 - Modellnugget in der Scoring-Verzweigung des Streams
4. Klicken Sie auf die Schaltfläche **Vorschau der Streambeschreibung anzeigen**.

Exportieren von Streambeschreibungen

Sie können den Inhalt der Streambeschreibung in eine HTML-Datei exportieren.

So exportieren Sie eine Streambeschreibung:

1. Klicken Sie im Hauptmenü auf Folgendes:
Datei > Streambeschreibung exportieren
2. Geben Sie einen Namen für die HTML-Datei ein und klicken Sie auf **Speichern**.

Ausführen von Streams

Sobald Sie die gewünschten Optionen für Streams angegeben und die gewünschten Knoten verbunden haben, können Sie den Stream ausführen, indem Sie die Daten durch die Knoten im Stream laufen lassen. Es gibt verschiedene Möglichkeiten, einen Stream in IBM SPSS Modeler auszuführen. Sie verfügen über folgende Möglichkeiten:

- Klicken Sie im Menü "Tools" auf **Ausführen**.
- Klicken Sie auf eine der Schaltflächen **Ausführen...** in der Symbolleiste. Mit diesen Schaltflächen können Sie den gesamten Stream ausführen oder lediglich den ausgewählten Endknoten. Weitere Informationen finden Sie in „IBM SPSS Modeler-Symbolleiste“ auf Seite 18.
- Um einen einzelnen Datenstream auszuführen, klicken Sie mit der rechten Maustaste auf einen Endknoten und dann im Popup-Menü auf **Ausführen**.
- Um einen Teil eines Datenstreams auszuführen, klicken Sie mit der rechten Maustaste auf einen Knoten, der kein Endknoten ist, und klicken Sie dann im Popup-Menü auf **Ab hier ausführen**. Dadurch werden nur die Operationen nach dem ausgewählten Knoten ausgeführt.

Um die Ausführung eines gerade laufenden Streams anzuhalten, können Sie auf die rote Schaltfläche "Stopp" in der Symbolleiste oder im Menü "Tools" auf **Ausführung anhalten** klicken.

Wenn ein Stream mehr als drei Sekunden für die Ausführung benötigt, wird das Dialogfeld "Ausführungs-Feedback" eingeblendet, in dem der Fortschritt angezeigt wird.

Bei einigen Knoten gibt es weitere Anzeigen, die zusätzliche Informationen zur Streamausführung bieten. Diese werden durch Auswahl der entsprechenden Zeile im Dialogfeld angezeigt. Die erste Zeile wird automatisch ausgewählt.

Arbeiten mit Modellen

Wenn ein Stream einen Modellierungsknoten enthält (d. h. einen Knoten aus der Registerkarte "Modellierung" oder "Datenbankmodellierung" der Knotenpalette), wird bei Ausführung des Streams ein **Modellnugget** erstellt. Ein Modellnugget ist ein Behälter für ein **Modell**, d. h. das Set von Regeln, Formeln und Gleichungen, das Ihnen Vorhersagen zu Ihren Quelldaten ermöglicht und das den Kern von Vorhersageanalysen ausmacht.



Abbildung 12. Modellnugget

Wenn Sie einen Modellierungsknoten erfolgreich ausführen, wird ein entsprechendes Modellnugget auf den Streamerstellungsbereich platziert, wo er durch ein goldfarbenes Rautensymbol repräsentiert wird (daher der Name "Nugget"). Sie können das Nugget öffnen und durch seinen Inhalt navigieren, um Details über das Modell zu sehen. Zur Anzeige der Vorhersagen verbinden Sie einen oder mehrere Modellknoten und führen diese aus. Die Ausgabe davon präsentiert die Vorhersagen in einer lesbaren Form.

Ein typischer Modellierungsstream besteht aus zwei Verzweigungen. Die **Modellierungsverzweigung** enthält den Modellierungsknoten zusammen mit der Quelle und vorangehenden Verarbeitungsknoten. Die **Modellierungsverzweigung** wird beim Ausführen des Modellierungsknotens erstellt und enthält das Modellnugget sowie den oder die Endknoten, die Sie zur Anzeige der Vorhersagen verwenden.

Weitere Informationen finden Sie im Handbuch *IBM SPSS Modeler Modellierungsknoten*.

Hinzufügen von Kommentaren und Anmerkungen zu Knoten und Streams

Sie müssen einen Stream eventuell für andere Personen in Ihrer Organisation beschreiben. Als Hilfe können Sie Streams, Knoten und Modellnuggets erklärende Kommentare hinzufügen.

Andere können diese Kommentare dann am Bildschirm ansehen oder Sie können eine Abbildung des Streams samt Kommentaren ausdrucken.

Sie können alle Kommentare für einen Stream oder Superknoten auflisten, die Reihenfolge der Kommentare in der Liste ändern, den Kommentartext bearbeiten und die Vorder- oder Hintergrundfarbe eines Kommentars ändern. Weitere Informationen finden Sie in „Auflisten von Streamkommentaren“ auf Seite 63.

Sie können auch Knoten in der Form von Textanmerkungen zu Streams, Knoten und Nuggets mithilfe der Registerkarte "Anmerkungen" eines Dialogfelds mit Streameigenschaften, eines Knotendialogfelds oder eines Modellnuggetfensters hinzufügen. Diese Anmerkungen sind nur sichtbar, wenn die Registerkarte "Anmerkungen" geöffnet ist, aber die Streamanmerkungen lassen sich auch als Bildschirmkommentare anzeigen. Weitere Informationen finden Sie in „Anmerkungen“ auf Seite 64.

Kommentare

Kommentare können die Form von Textfeldern haben, in die Sie eine beliebige Textmenge eingeben können, und Sie können beliebig viele Kommentare hinzufügen. Ein Kommentar kann freistehend (mit keinerlei Streamobjekten verbunden) oder mit einem oder mehreren Knoten oder Modellnuggets im Stream verbunden sein. Freistehende Kommentare werden in der Regel verwendet, um den Gesamtzweck des Streams zu beschreiben; verbundene Kommentare beschreiben den Knoten oder das Nugget, mit dem sie verbunden sind. Mit Knoten und Nuggets können mehrere Kommentare verbunden sein, und der Stream kann eine beliebige Anzahl freistehender Kommentare aufweisen.

Hinweis: Sie können Streamanmerkungen auch als Bildschirmkommentare anzeigen, allerdings lassen sich diese nicht an Knoten oder Nuggets anhängen. Weitere Informationen finden Sie in „Umwandeln von Anmerkungen in Kommentare“ auf Seite 64.

Das Aussehen des Textfelds ändert sich entsprechend dem aktuellen Modus des Kommentars (oder der als Kommentar angezeigten Anmerkung). Siehe dazu die folgende Tabelle.

Tabelle 3. Modi des Kommentar- und Anmerkungstextfelds



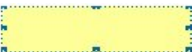

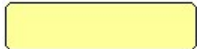

Kommentartextfeld	Anmerkungstextfeld	Modus	Bedeutung	Erzielt durch...
		Bearbeiten	Kommentar ist zur Bearbeitung geöffnet.	Erstellen eines neuen Kommentars bzw. einer neuen Anmerkung oder Doppelklick auf einen bestehenden Kommentar.
		Zuletzt ausgewählt	Kommentar lässt sich verschieben oder löschen oder seine Größe kann geändert werden.	Klick auf den Streamhintergrund nach der Bearbeitung oder Einmalklick auf einen bestehenden Kommentar bzw. eine bestehende Anmerkung.

Tabelle 3. Modi des Kommentar- und Anmerkungstextfelds (Forts.)

Kommentartextfeld	Anmerkungstextfeld	Modus	Bedeutung	Erzielt durch...
		Ansicht	Bearbeitung ist abgeschlossen.	Klick auf einen anderen Knoten, einen anderen Kommentar oder eine andere Anmerkung nach der Bearbeitung.

Wenn Sie einen neuen freistehenden Kommentar erstellen, wird dieser anfangs in der oberen linken Ecke des Streamerstellungsbereichs angezeigt.

Wenn Sie einen Kommentar mit einem Knoten oder ein Nugget verknüpfen, wird der Kommentar anfangs über dem Streamobjekt angezeigt, mit dem er verbunden ist.

Das Textfeld ist weiß, um zu zeigen, dass sich Text eingeben lässt. Wenn Sie den Text eingegeben haben, klicken Sie außerhalb des Textfelds. Der Kommentarhintergrund wird gelb, um zu zeigen, dass der Texteintrag abgeschlossen ist. Der Kommentar bleibt ausgewählt, damit Sie ihn verschieben, löschen oder seine Größe ändern können.

Bei erneutem Klicken ändert sich die Umrandung in durchgehende Linien und zeigt damit, dass die Bearbeitung abgeschlossen ist.

Durch Doppelklicken auf einen Kommentar wechselt das Textfeld in den Bearbeitungsmodus und der Hintergrund wird weiß, d. h. der Kommentartext kann bearbeitet werden.

Außerdem können Sie Kommentare mit Superknoten verknüpfen.

Operationen mit Kommentaren: Sie können eine Vielzahl von Operationen mit Kommentaren ausführen. Sie verfügen über folgende Möglichkeiten:

- Hinzufügen eines freistehenden Kommentars
- Verknüpfen eines Kommentars mit einem Knoten oder Nugget
- Bearbeiten eines Kommentars
- Größenänderung eines Kommentars
- Verschieben eines Kommentars
- Lösen der Verbindung eines Kommentars
- Löschen eines Kommentars
- Ein-/Ausblenden aller Kommentare für einen Stream

So fügen Sie einen freistehenden Kommentar hinzu:

1. Stellen Sie sicher, dass im Stream keine Auswahl getroffen ist.
2. Führen Sie eine der folgenden Aktionen aus:
 - Klicken Sie im Hauptmenü auf Folgendes:
Einfügen > Neuer Kommentar
 - Klicken Sie mit der rechten Maustaste auf den Streamhintergrund und klicken Sie im Popup-Menü auf **Neuer Kommentar**.
 - Klicken Sie in der Symbolleiste auf die Schaltfläche **Neuer Kommentar**.
3. Geben Sie den Kommentartext ein (oder fügen Sie Text aus der Zwischenablage ein).
4. Klicken Sie auf einen Knoten im Stream, um den Kommentar zu speichern.

So verknüpfen Sie einen Kommentar mit einem Knoten oder Nugget:

1. Wählen Sie einen oder mehrere Knoten oder Nuggets im Streamerstellungsbereich aus.
2. Führen Sie eine der folgenden Aktionen aus:
 - Klicken Sie im Hauptmenü auf Folgendes:
Einfügen > Neuer Kommentar
 - Klicken Sie mit der rechten Maustaste auf den Streamhintergrund und klicken Sie im Popup-Menü auf **Neuer Kommentar**.
 - Klicken Sie in der Symbolleiste auf die Schaltfläche **Neuer Kommentar**.
3. Geben Sie den Kommentartext ein.
4. Klicken Sie auf einen anderen Knoten im Stream, um den Kommentar zu speichern.
Sie haben außerdem folgende Möglichkeiten:
5. Fügen Sie einen freistehenden Kommentar ein (siehe vorherigen Abschnitt).
6. Führen Sie eine der folgenden Aktionen aus:
 - Wählen Sie den Kommentar aus, drücken Sie F2 und wählen Sie dann den Knoten oder das Nugget aus.
 - Wählen Sie den Knoten oder das Nugget aus, drücken Sie F2 und wählen Sie dann den Kommentar aus.
 - (Nur Mäuse mit drei Tasten) Bewegen Sie den Mauszeiger über den Kommentar, halten Sie die mittlere Taste gedrückt, ziehen Sie den Mauszeiger auf den Knoten oder das Nugget und lassen Sie die Maustaste los.

So verknüpfen Sie einen Kommentar mit einem zusätzlichen Knoten oder Nugget:

Wenn ein Kommentar bereits mit einem Knoten oder Nugget verknüpft ist oder sich auf Streamebene befindet und Sie diesen mit einem weiteren Knoten oder Nugget verknüpfen möchten, führen Sie eine der folgenden Aktionen aus:

- Wählen Sie den Kommentar aus, drücken Sie F2 und wählen Sie dann den Knoten oder das Nugget aus.
- Wählen Sie den Knoten oder das Nugget aus, drücken Sie F2 und wählen Sie dann den Kommentar aus.
- (Nur Mäuse mit drei Tasten) Bewegen Sie den Mauszeiger über den Kommentar, halten Sie die mittlere Taste gedrückt, ziehen Sie den Mauszeiger auf den Knoten oder das Nugget und lassen Sie die Maustaste los.

So bearbeiten Sie einen bestehenden Kommentar:

1. Führen Sie eine der folgenden Aktionen aus:
 - Doppelklicken Sie auf das Kommentartextfeld.
 - Wählen Sie das Textfeld aus und drücken Sie die Eingabetaste.
 - Klicken Sie mit der rechten Maustaste auf das Textfeld, um sein Menü zu öffnen, und klicken Sie auf **Bearbeiten**.
2. Bearbeiten Sie den Kommentartext. Sie können die Windows-Standardtastenkombinationen zur Bearbeitung verwenden, z. B. Strg+C, um Text zu kopieren. Andere Optionen während der Bearbeitung werden im Popup-Menü für den Kommentar aufgelistet.
3. Klicken Sie einmal außerhalb des Textfelds, um die Größenänderungssteuerungen anzuzeigen, und noch einmal, um den Kommentar abzuschließen.

Ändern der Größe eines Kommentartextfelds

1. Wählen Sie den Kommentar aus, um die Größenänderungssteuerungen anzuzeigen.
2. Klicken Sie auf eine Steuerung und ziehen Sie sie, um die Größe des Felds zu ändern.
3. Klicken Sie außerhalb des Textfelds, um die Änderung zu speichern.

So verschieben Sie einen bestehenden Kommentar:

Wenn Sie einen Kommentar ohne seine verknüpften Objekte (falls vorhanden) verschieben möchten, führen Sie eine der folgenden Aktionen aus:

- Bewegen Sie den Mauszeiger über den Kommentar, halten Sie die linke Maustaste gedrückt und ziehen Sie den Kommentar an die neue Position.
- Wählen Sie den Kommentar aus, halten Sie die Taste Alt gedrückt und verschieben Sie den Kommentar mithilfe der Pfeiltasten.

Wenn Sie einen Kommentar zusammen mit etwaigen verknüpften Knoten oder Nuggets verschieben möchten:

1. Wählen Sie alle Objekte aus, die Sie verschieben möchten.
2. Führen Sie eine der folgenden Aktionen aus:
 - Verschieben Sie den Mauszeiger über eines der Objekte, halten Sie die linke Maustaste gedrückt und ziehen Sie die Objekte an die neue Position.
 - Wählen Sie eines der Objekte aus, halten Sie die Taste Alt gedrückt und verschieben Sie die Objekte mithilfe der Pfeiltasten.

So lösen Sie einen Kommentar von einem Knoten oder Nugget:

1. Wählen Sie einen oder mehrere Kommentare aus, die Sie lösen möchten.
2. Führen Sie eine der folgenden Aktionen aus:
 - Drücken Sie die Taste F3.
 - Klicken Sie mit der rechten Maustaste auf einen ausgewählten Kommentar und klicken Sie in dessen Menü auf **Trennen**".

So löschen Sie einen Kommentar:

1. Wählen Sie einen oder mehrere Kommentare aus, die Sie löschen möchten.
2. Führen Sie eine der folgenden Aktionen aus:
 - Drücken Sie die Lösch Taste.
 - Klicken Sie mit der rechten Maustaste auf einen ausgewählten Kommentar und klicken Sie in dessen Menü auf **Löschen**.

Wenn der Kommentar mit einem Knoten oder Nugget verknüpft war, wird die Verbindungslinie ebenfalls gelöscht.

Wenn der Kommentar ursprünglich eine Stream- oder Superknotenannotation war, die in einen freistehenden Kommentar umgewandelt wurde, wird er aus dem Erstellungsbereich gelöscht, sein Text bleibt jedoch auf der Registerkarte "Anmerkungen" für den Stream bzw. Superknoten erhalten.

So blenden Sie alle Kommentare für einen Stream ein oder aus:

1. Führen Sie eine der folgenden Aktionen aus:
 - Klicken Sie im Hauptmenü auf Folgendes:
Ansicht > Kommentare
 - Klicken Sie in der Symbolleiste auf die Schaltfläche **Kommentare ein-/ausblenden**.

Auflisten von Streamkommentaren: Sie können eine Liste aller Kommentare anzeigen, die für einen bestimmten Stream oder Superknoten erstellt wurden.

Diese Liste bietet die folgenden Möglichkeiten:

- Ändern der Reihenfolge der Kommentare
- Bearbeiten des Kommentartexts

- Ändern der Vorder- oder Hintergrundfarbe eines Kommentars

Auflisten von Kommentaren

Führen Sie zum Auflisten der Kommentare für einen Stream eine der folgenden Aktionen aus:

- Klicken Sie im Hauptmenü auf Folgendes:
Tools > Streameigenschaften > Kommentare
- Klicken Sie mit der rechten Maustaste auf einen Stream im Managerbereich und klicken Sie auf **Streameigenschaften** und dann auf **Kommentare**.
- Klicken Sie mit der rechten Maustaste in einen Streamhintergrund im Erstellungsbereich und klicken Sie auf **Streameigenschaften** und dann auf **Kommentare**.

Text. Der Text des Kommentars. Doppelklicken Sie auf den Text, um das Feld in ein editierbares Textfeld zu verwandeln.

Verknüpfungen. Der Name des Knotens, mit dem der Kommentar verbunden ist. Wenn dieses Feld leer ist, gilt der Kommentar für den Stream.

Positionierungsschaltflächen. Diese verschieben einen ausgewählten Kommentar in der Liste nach oben oder unten.

Kommentarfarben. Um die Vorder- oder Hintergrundfarbe eines Kommentars zu ändern, wählen Sie den Kommentar aus, markieren das Kontrollkästchen **Benutzerdefinierte Farben** und wählen dann eine Farbe aus der Liste **Hintergrund** und/oder **Vordergrund**. Klicken Sie auf **Anwenden** und dann in den Streamhintergrund, um die Wirkung der Änderung zu sehen. Klicken Sie auf **OK**, um die Änderung zu speichern.

Umwandeln von Anmerkungen in Kommentare: Anmerkungen zu Streams oder Superknoten können in Kommentare umgewandelt werden.

Bei Streams wird die Anmerkung im Streamerstellungsbereich in einen freistehenden Kommentar umgewandelt (d. h. ohne Verknüpfung zu Knoten).

Wenn eine Superknotenanmerkung in einen Kommentar umgewandelt wird, wird dieser im Erstellungsbereich nicht mit dem Superknoten verknüpft, aber er ist sichtbar, sobald Sie den Superknoten vergrößern.

So wandeln Sie eine Streamanmerkung in einen Kommentar um:

1. Klicken Sie im Menü "Tools" auf **Streameigenschaften**. (Als Alternative können Sie mit der rechten Maustaste auf einen Stream im Managerbereich und dann auf **Streameigenschaften** klicken.)
2. Klicken Sie auf die Registerkarte **Anmerkungen**.
3. Aktivieren Sie das Kontrollkästchen **Anmerkung als Kommentar anzeigen**.
4. Klicken Sie auf **OK**.

So wandeln Sie eine Superknotenanmerkung in einen Kommentar um:

1. Doppelklicken Sie auf das Symbol des Superknotens im Erstellungsbereich.
2. Klicken Sie auf die Registerkarte **Anmerkungen**.
3. Aktivieren Sie das Kontrollkästchen **Anmerkung als Kommentar anzeigen**.
4. Klicken Sie auf **OK**.

Anmerkungen

Knoten, Streams und Modelle können auf verschiedene Art und Weise mit Anmerkungen versehen werden. Sie können beschreibende Anmerkungen hinzufügen und einen benutzerdefinierten Namen angeben.

Diese Optionen sind besonders hilfreich, wenn Sie Berichte für Streams generieren, die dem Projektbereich hinzugefügt werden. Für Knoten und Modellnuggets können Sie auch QuickInfo-Text hinzufügen, um zwischen ähnlichen Knoten im Streamerstellungsbereich zu unterscheiden.

Hinzufügen von Anmerkungen

Wenn Sie einen Knoten oder ein Modellnugget bearbeiten, wird ein Dialogfeld mit Registerkarten geöffnet, das eine Registerkarte "Anmerkungen" enthält, mit der Sie eine Vielzahl von Anmerkungen festlegen können. Sie können die Registerkarte "Anmerkungen" auch direkt öffnen.

1. Um Anmerkungen für einen Knoten oder ein Nugget zu erstellen, klicken Sie mit der rechten Maustaste auf den Knoten bzw. das Nugget im Streamerstellungsbereich und dann auf **Umbenennen und mit Anmerkung versehen**. Das Bearbeitungsdialogfeld wird so angezeigt, dass die Registerkarte "Anmerkungen" sichtbar ist.
2. Um einen Stream mit Anmerkungen zu versehen, klicken Sie im Menü "Tools" auf **Streameigenschaften**. (Als Alternative können Sie mit der rechten Maustaste auf einen Stream im Managerbereich und dann auf **Streameigenschaften** klicken.) Klicken Sie auf die Registerkarte "Anmerkungen".

Name. Wählen Sie **Benutzerdefiniert** aus, um den automatisch generierten Namen zu ändern oder um einen eindeutigen Namen für den Knoten zu erstellen, wie er im Streamerstellungsbereich angezeigt wird.

QuickInfo-Text. (Nur für Knoten und Modellnuggets) Geben Sie Text ein, der im Streamerstellungsbereich als QuickInfo verwendet wird. Dies ist besonders hilfreich, wenn Sie mit einer größeren Anzahl ähnlicher Knoten arbeiten.

Schlüsselwörter. Geben Sie Schlüsselwörter an, die in Projektberichten und bei der Suche nach Knoten in einem Stream oder beim Verfolgen von Objekten verwendet werden sollen, die im Repository gespeichert sind (siehe „Informationen zu IBM SPSS Collaboration and Deployment Services Repository“ auf Seite 191). Mehrere Schlüsselwörter können durch Semikolons voneinander getrennt werden, z. B. Einkommen; Feldfruchttyp; Subventionshöhe. Leerzeichen am Anfang und am Ende von Schlüsselwörtern werden ausgelassen; Einkommen; Feldfruchttyp führt z. B. zu demselben Ergebnis wie Einkommen;Feldfruchttyp. (Die Leerzeichen innerhalb von Schlüsselwörtern werden jedoch nicht ausgelassen. Zum Beispiel sind Höhe der Subventionen mit einem Leerzeichen vor "der" und Höhe der Subventionen mit zwei Leerzeichen vor "der" nicht dasselbe.)

Der Haupttextbereich kann verwendet werden, um lange Anmerkungen zu den Operationen des Knotens oder Entscheidungen im Knoten einzugeben. Wenn Sie z. B. Streams gemeinsam nutzen und erneut verwenden, ist es hilfreich, Anmerkungen zu Entscheidungen zu machen, z. B. Verwerfen eines Felds mit zahlreichen Leerzeichen unter Verwendung eines Filterknotens. Wenn Sie den Knoten mit Anmerkungen versehen, werden diese Informationen zusammen mit dem Knoten gespeichert. Sie können diese Anmerkungen auch in einem Projektbericht berücksichtigen, der im Projektbereich erstellt wurde. Weitere Informationen finden Sie in „Einführung in Projekte“ auf Seite 215.

Anmerkung als Kommentar anzeigen. (Nur für Stream- und Superknotenankmerkungen) Markieren Sie dieses Kontrollkästchen, um die Anmerkung in einen freistehenden Kommentar umzuwandeln, der im Streamerstellungsbereich sichtbar ist. Weitere Informationen finden Sie in „Hinzufügen von Kommentaren und Anmerkungen zu Knoten und Streams“ auf Seite 60.

ID. Zeigt eine eindeutige ID an, die zur Referenzierung des Knotens für Scripterstellung oder Automatisierung verwendet werden kann. Dieser Wert wird bei der Erstellung des Knotens automatisch generiert und ändert sich nicht. Beachten Sie außerdem, dass zur Vermeidung einer Verwechslung mit dem Buchstaben "O" in Knoten-IDs keine Nullen verwendet werden. Mit der Kopierschaltfläche auf der rechten Seite können Sie die ID kopieren und nach Bedarf in Scripts oder anderswo einfügen.

Speichern von Datenstreams

Nachdem Sie einen Stream erstellt haben, können Sie ihn für eine zukünftige Wiederverwendung speichern.

So speichern Sie einen Stream

1. Klicken Sie im Menü "Datei" auf **Stream speichern** oder **Stream speichern unter**.
2. Blättern Sie im Dialogfeld "Speichern" zu dem Ordner, in dem die Streamdatei gespeichert werden soll.
3. Geben Sie im Textfeld "Dateiname" einen Namen für den Stream ein.
4. Wählen Sie **Zu Projekt hinzufügen**, wenn Sie den gespeicherten Stream zum aktuellen Projekt hinzufügen möchten.

Wenn Sie auf **Speichern** klicken, wird der Stream mit der Erweiterung `*.str` im angegebenen Verzeichnis gespeichert.

Automatische Sicherungsdateien. Jedes Mal, wenn ein Stream gespeichert wird, wird die zuvor gespeicherte Version der Datei automatisch als Sicherung aufbewahrt und ein Bindestrich zum Dateinamen hinzugefügt (z. B. `meinStream.str-`). Um die gesicherte Version wiederherzustellen, löschen Sie einfach den Bindestrich und öffnen Sie die Datei erneut.

Speichern von Statusangaben

Zusätzlich zu Streams können Sie **Statusangaben** speichern, u. a. das aktuell angezeigte Streamdiagramm und alle Modellnuggets, die Sie erstellt haben (auf der Registerkarte "Modelle" im Managerbereich aufgelistet).

So speichern Sie einen Status:

1. Klicken Sie im Menü "Datei" auf:
Status > Status speichern oder **Status speichern unter**
2. Blättern Sie im Dialogfeld "Speichern" zu dem Ordner, in dem die Statusdatei gespeichert werden soll.

Wenn Sie auf **Speichern** klicken, wird der Status mit der Erweiterung `*.cst` im angegebenen Verzeichnis gespeichert.

Speichern von Knoten

Sie können auch einen einzelnen Knoten speichern, indem Sie mit der rechten Maustaste auf den Knoten im Streamerstellungsbereich klicken und im Popup-Menü auf **Knoten speichern** klicken. Verwenden Sie die Dateinamenerweiterung `*.nod`.

Speichern mehrerer Datenstromobjekte

Wenn Sie IBM SPSS Modeler verlassen, obwohl mehrere Objekte, wie Streams, Projekte oder Modellnuggets, nicht gespeichert sind, werden Sie zum Speichern aufgefordert, bevor Sie die Software vollständig beenden. Wenn Sie Elemente speichern möchten, wird ein Dialogfeld mit Optionen zum Speichern der einzelnen Objekte angezeigt.

1. Aktivieren Sie einfach die Kontrollkästchen für die Objekte, die Sie speichern möchten.
2. Klicken Sie auf **OK**, um die Objekte am gewünschten Speicherort zu speichern.

Es wird dann für jedes Objekt ein Standarddialogfeld zum Speichern aufgerufen. Nachdem das Speichern beendet ist, wird die Anwendung geschlossen.

Speichern der Ausgabe

Tabellen, Diagramme und Berichte, die aus IBM SPSS Modeler-Ausgabeknoten erstellt wurden, können nun im Ausgabeobjektformat (`*.cou`) gespeichert werden.

1. Um die zu speichernde Ausgabe anzuzeigen, klicken Sie in den Menüs des Ausgabefensters auf:

Datei > Speichern

2. Geben Sie einen Namen und einen Speicherort für die Ausgabedatei an.
3. Wählen Sie optional **Datei zu Projekt hinzufügen** im Dialogfeld "Speichern", um die Datei in das aktuelle Projekt aufzunehmen. Weitere Informationen finden Sie in „Einführung in Projekte“ auf Seite 215.

Als Alternative können Sie mit der rechten Maustaste auf ein beliebiges Ausgabeobjekt im Managerbereich klicken und im Popup-Menü die Option **Speichern** auswählen.

Verschlüsseln und Entschlüsseln von Informationen

Beim Speichern von Streams, Knoten, Projekten, Ausgabedateien oder Modellnuggets können Sie das betreffende Element verschlüsseln, um es vor Verwendung durch Unbefugte zu schützen. Dazu wählen Sie beim Speichern eine zusätzliche Option aus und legen ein Kennwort für das zu speichernde Element fest. Diese Verschlüsselung kann für alle Objekte festgelegt werden, die Sie speichern, und bietet Ihnen zusätzliche Sicherheit; es ist nicht dasselbe wie die SSL-Verschlüsselung, die beim Austausch von Dateien zwischen IBM SPSS Modeler und IBM SPSS Modeler Server verwendet wird.

Wenn Sie versuchen, ein verschlüsseltes Element zu öffnen, werden Sie zur Eingabe des Kennworts aufgefordert. Nach Eingabe des richtigen Kennworts wird das Element automatisch entschlüsselt und wie gewohnt geöffnet.

So verschlüsseln Sie ein Objekt:

1. Klicken Sie im Dialogfeld "Speichern" für das zu verschlüsselnde Element auf **Optionen**. Das Dialogfeld "Verschlüsselungsoptionen" wird geöffnet.
2. Wählen Sie die Option **Diese Datei verschlüsseln** aus.
3. Für noch größere Sicherheit können Sie optional **Kennwort maskieren** auswählen. Dadurch werden alle Ihre Eingaben nur als eine Reihe von Punkten angezeigt.
4. Geben Sie das Kennwort ein. *Warnung:* Wenn Sie das Kennwort vergessen, kann die Datei bzw. das Modell nicht mehr geöffnet werden.
5. Wenn Sie **Kennwort maskieren** ausgewählt haben, müssen Sie das Kennwort erneut eingeben, um sicherzustellen, dass die erste Eingabe korrekt war.
6. Klicken Sie auf **OK**, um zum Dialogfeld "Speichern" zurückzukehren.

Hinweis: Wenn Sie eine Kopie eines durch Verschlüsselung geschützten Elements speichern, wird das neue Element automatisch in verschlüsselter Form unter Verwendung des ursprünglichen Kennworts gespeichert, es sei denn, Sie ändern die Einstellungen im Dialogfeld "Verschlüsselungsoptionen".

Laden von Dateien

Sie können eine Vielzahl gespeicherter Objekte in IBM SPSS Modeler erneut laden:

- Streams (.str)
- Statusangaben (.cst)
- Modelle (.gm)
- Modellpalette (.gen)
- Knoten (.nod)
- Ausgabe (.cou)
- Projekte (.cpj)

Öffnen neuer Dateien

Streams können direkt im Menü "Datei" geladen werden.

- Klicken Sie im Menü "Datei" auf **Stream öffnen**.

Alle anderen Dateitypen können mithilfe der Untermenüoptionen im Menü "Datei" geöffnet werden. Um z. B. ein Modell zu laden, klicken Sie im Menü "Datei" auf:

Modelle > Modell öffnen oder Modellpalette laden

Öffnen zuletzt verwendeter Dateien

Um die zuletzt verwendeten Dateien schnell zu laden, können Sie die Optionen unten im Menü "Datei" verwenden.

Wählen Sie **Zuletzt verwendete Streams**, **Zuletzt verwendete Projekte** oder **Zuletzt verwendete Statusangaben**, um eine Liste mit den zuletzt verwendeten Dateien zu öffnen.

Zuordnen von Datenstreams

Mit dem Zuordnungstool können Sie eine neue Datenquelle mit einem bereits vorhandenen Stream verbinden. Mit dem Zuordnungstool können Sie nicht nur die Verbindung einrichten, sondern auch festlegen, wie Felder in der neuen Quelle die im bereits vorhandenen Stream ersetzen. Anstatt einen kompletten Datenstream für eine neue Datenquelle erneut zu erstellen, können Sie einfach eine Verbindung mit einem bereits vorhandenen Stream herstellen.

Mit dem Datenzuordnungstool können Sie zwei Streamfragmente zusammenfügen und sicherstellen, dass alle (wichtigen) Feldnamen zugeordnet werden. Das Zuordnen von Daten führt hauptsächlich zur Erstellung eines neuen Filterknotens, der die entsprechenden Felder zuordnet, indem er sie umbenennt.

Es gibt zwei gleichwertige Möglichkeiten zum Zuordnen von Daten:

Auswählen des Ersatzknotens. Diese Methode beginnt mit dem zu ersetzenden Knoten. Klicken Sie zuerst mit der rechten Maustaste auf den zu ersetzenden Knoten. Wählen Sie dann über die Option **Datenzuordnung > Ersatzknoten auswählen** im Popup-Menü den Knoten aus, der den ersten Knoten ersetzen soll.

Zuordnen zu. Diese Methode beginnt mit dem Knoten, der im Stream eingefügt werden soll. Klicken Sie zuerst mit der rechten Maustaste auf den einzufügenden Knoten. Wählen Sie dann über die Option **Datenzuordnung > Zuordnen zu** im Popup-Menü den Knoten aus, dem der erste Knoten hinzugefügt werden soll. Diese Methode ist besonders geeignet für die Zuordnung zu einem Endknoten. *Hinweis:* Für Zusammenführungs- oder Anhangknoten sind keine Zuordnungen möglich. Stattdessen sollte der Stream auf die normale Art und Weise mit dem Zusammenführungsknoten verbunden werden.

Datenzuordnung ist eng in die Streamerstellung integriert. Wenn Sie versuchen, einen Knoten zu verbinden, der bereits über eine Verbindung verfügt, erhalten Sie die Option die Verbindung oder Zuordnung dieses Knotens zu ersetzen.

Zuordnen von Daten zu einer Vorlage

Um die Datenquelle für einen Vorlagenstream durch einen neuen Quellenknoten zu ersetzen, der Ihre eigenen Daten in IBM SPSS Modeler einbringt, sollten Sie aus dem Popup-Menü "Datenzuordnung" die Option **Ersatzknoten auswählen** verwenden. Diese Option steht für alle Knoten mit Ausnahme von Zusammenführungs-, Aggregat- und allen Endknoten zur Verfügung. Wenn Sie das Datenzuordnungstool zum Durchführen dieser Aktion verwenden, wird sichergestellt, dass die Felder zwischen den bereits vorhandenen Streamoperationen und der neuen Datenquelle ordnungsgemäß zugeordnet werden. Die folgenden Schritte bieten einen Überblick über den Datenzuordnungsprozess.

Schritt 1: Geben Sie die erforderlichen Felder im ursprünglichen Quellenknoten an. Damit die Streamoperationen ordnungsgemäß ausgeführt werden können, müssen die erforderlichen Felder angegeben werden. Weitere Informationen finden Sie in „Festlegen von wichtigen Feldern“ auf Seite 69.

Schritt 2: Fügen Sie dem Streamerstellungsbereich eine neue Datenquelle hinzu. Bringen Sie mithilfe eines der Quellenknoten die neuen Ersatzdaten ein.

Schritt 3: Ersetzen Sie den Vorlagenquellenknoten. Wählen Sie im Popup-Menü des Vorlagenquellenknotens die Option "Datenzuordnung" aus, klicken Sie auf **Ersatzknoten auswählen** und wählen Sie dann den Quellenknoten für die Ersatzdaten aus.

Schritt 4: Überprüfen Sie die zugeordneten Felder. Prüfen Sie in dem Dialogfeld, das nun geöffnet wird, ob die Software die Felder aus der Ersatzdatenquelle dem Stream ordnungsgemäß zuordnet. Sämtliche nicht zugeordneten, aber erforderlichen Felder werden rot dargestellt. Diese Felder werden in Streamoperationen verwendet und müssen durch ein ähnliches Feld in der neuen Datenquelle ersetzt werden, damit die nachgeordneten Operationen ordnungsgemäß ausgeführt werden können. Weitere Informationen finden Sie in „Prüfen von zugeordneten Feldern“ auf Seite 70.

Nachdem Sie das Dialogfeld verwendet haben, um sicherzustellen, dass alle erforderlichen Felder ordnungsgemäß zugeordnet sind, wird die alte Datenquelle getrennt und die neue Datenquelle mithilfe des Filterknotens namens *Zuordnen* mit dem Stream verbunden. Dieser Filterknoten leitet die tatsächliche Zuordnung der Felder im Stream. Ein Filterknoten namens *Zuordnung aufheben* ist ebenfalls im Streamerstellungsbereich vorhanden. Der Filterknoten *Zuordnung aufheben* kann dazu verwendet werden, die Feldnamenzuordnung umzukehren, indem er dem Stream hinzugefügt wird. Die Zuordnung der bereits zugeordneten Felder wird dadurch aufgehoben. Beachten Sie jedoch, dass sämtliche nachgeordneten Endknoten für eine erneute Auswahl der Felder und Überlagerungen bearbeitet werden müssen.

Zuordnen zwischen Streams

Ähnlich wie beim Verbinden von Knoten müssen bei dieser Datenzuordnungsmethode keine erforderlichen Felder vorab festgelegt werden. Mit dieser Methode verbinden Sie einen Stream mit einem anderen, indem Sie im Popup-Menü "Datenzuordnung" die Option **Zuordnen zu** auswählen. Diese Art der Datenzuordnung ist für das Zuordnen zu Endknoten und zum Kopieren und Einfügen zwischen Streams sinnvoll. *Hinweis:* Mit der Option **Zuordnen zu** können Sie keine Zuordnung zu Zusammenführungs-, Anhang- und anderen Quellenknoten herstellen.

So ordnen Sie Daten zwischen Streams zu:

1. Klicken Sie mit der rechten Maustaste auf den Knoten, den Sie zum Verbinden mit dem neuen Stream verwenden möchten.
2. Klicken Sie im Menü auf:
Datenzuordnung > Zuordnen zu
3. Verwenden Sie den Cursor, um einen Zielknoten im Ziel-Stream auszuwählen.
4. Stellen Sie sicher, dass die Felder in dem Dialogfeld, das jetzt angezeigt wird, ordnungsgemäß zugeordnet sind. Klicken Sie dann auf **OK**.

Festlegen von wichtigen Feldern

Bei der Zuordnung zu einem bestehenden Stream werden wichtige Felder in der Regel vom Streamautor festgelegt. Diese wichtigen Felder zeigen an, ob ein bestimmtes Feld in nachgeordneten Operationen verwendet wird. Der bereits vorhandene Stream erstellt u. U. ein Modell, das ein Feld namens *Verlust* verwendet. In diesem Stream ist *Verlust* ein wichtiges Feld, da eine Modellerstellung ohne dieses Feld nicht möglich ist. Auf ähnliche Weise sind Felder, die für Bearbeitungsknoten, z. B. für einen Ableitungsknoten, verwendet werden, zum Ableiten des neuen Felds erforderlich. Wenn Sie solche Felder ausdrücklich als wichtig markieren, können Sie sicherstellen, dass ihnen im neuen Quellenknoten die richtigen Felder zugeordnet werden. Wenn keine wichtigen Felder zugeordnet sind, erhalten Sie eine Fehlermeldung. Wenn Sie feststellen, dass einige Bearbeitungs- oder Ausgabeknoten nicht erforderlich sind, können Sie die Knoten aus dem Stream löschen und die entsprechenden Felder aus der Liste der wichtigen Felder entfernen.

So legen Sie die wichtigen Felder fest:

1. Klicken Sie mit der rechten Maustaste auf den Quellenknoten des Vorlagenstreams, der ersetzt werden soll.

2. Klicken Sie im Menü auf:

Datenzuordnung > Wichtige Felder festlegen

3. Mit der Feldauswahlschaltfläche können Sie der Liste Felder hinzufügen oder entfernen. Um die Feldauswahl zu aktivieren, klicken Sie auf das Symbol rechts neben der Feldliste.

Prüfen von zugeordneten Feldern

Sobald Sie den Punkt ausgewählt haben, an dem ein Datenstream bzw. eine Datenquelle einem anderen zugeordnet wird, wird ein Dialogfeld angezeigt, in dem Sie Felder für die Zuordnung auswählen oder sicherstellen können, dass die Systemstandardzuordnung richtig ist. Wenn für den Stream oder die Datenquelle wichtige Felder festgelegt und nicht zugeordnet wurden, werden diese Felder rot dargestellt. Sämtliche nicht zugeordneten Felder aus der Datenquelle durchlaufen den Filterknoten unverändert. Beachten Sie jedoch, dass Sie auch Felder zuordnen können, die nicht "wichtig" sind.

Original. Listet alle Felder in der Vorlage oder im bereits vorhandenen Stream auf - alle Felder, die weiter unten im Stream vorhanden sind. Diesen Feldern werden Felder aus der neuen Datenquelle zugeordnet.

Zugeordnet. Listet die Felder auf, die für die Zuordnung zu Vorlagenfeldern ausgewählt wurden. Dabei handelt es sich um die Felder, deren Namen eventuell geändert werden müssen, um den in Streamoperationen verwendeten ursprünglichen Feldern zu entsprechen. Klicken Sie in die Tabellenzelle für ein Feld, um eine Liste mit verfügbaren Feldern zu aktivieren.

Wenn Sie sich nicht sicher sind, welche Felder zugeordnet werden sollen, sollten Sie die Quelldaten vor der Zuordnung detailliert untersuchen. Sie können z. B. die Registerkarte "Typen" im Quellenknoten verwenden, um eine Zusammenfassung der Quelldaten zu prüfen.

Tipps und Tastenkombinationen

Arbeiten Sie schnell und problemlos, indem Sie sich mit den folgenden Tastenkombinationen und Tipps vertraut machen:

- **Schnelles Erstellen von Streams durch Doppelklicken.** Doppelklicken Sie einfach auf einen Knoten in der Palette, um ihn zum aktuellen Stream hinzuzufügen und zu verbinden.
- **Verwenden von Tastenkombinationen zur Auswahl von nachgeordneten Knoten.** Drücken Sie Strg+Q und Strg+W, um alle nachgeordneten Knoten aus- bzw. abzuwählen.
- **Verwenden von Tastenkombinationen zum Verbinden und Trennen von Knoten.** Wenn ein Knoten im Erstellungsbereich ausgewählt ist, drücken Sie F2, um eine Verbindung zu beginnen, drücken Sie die Tabulatortaste, um zum gewünschten Knoten zu gelangen, und Umschalt+Leertaste, um die Verbindung fertig zu stellen. Drücken Sie F3, um alle Ein- und Ausgaben vom ausgewählten Knoten zu trennen.
- **Anpassen der Registerkarte "Knotenpalette" mit Ihren Lieblingsknoten.** Klicken Sie im Menü "Tools" auf **Paletten verwalten**, um ein Dialogfeld zu öffnen, in dem Sie die in der Knotenpalette angezeigten Knoten hinzufügen, entfernen bzw. verschieben können.
- **Umbenennen von Knoten und Hinzufügen von QuickInfo.** Jedes Knotendialogfeld enthält eine Registerkarte "Anmerkungen", auf der Sie einen benutzerdefinierten Namen für Knoten im Erstellungsbereich angeben können sowie QuickInfo hinzufügen können, wodurch Ihr Stream einfacher organisiert wird. Sie können auch lange Anmerkungen zur Fortschrittsverfolgung einfügen, Prozessdetails speichern und sämtliche Geschäftsentscheidungen notieren, die erforderlich sind oder umgesetzt wurden.
- **Automatisches Einfügen von Werten in einen CLEM-Ausdruck.** Mit dem Expression Builder, den Sie in einer Vielzahl von Dialogfeldern aufrufen können (z. B. Dialogfelder für Ableitungs- und Filterknoten), können Sie automatisch Feldwerte in einen CLEM-Ausdruck einfügen. Klicken Sie auf die Werteschaltfläche im Expression Builder, um eine Auswahl aus den bereits vorhandenen Feldwerten zu treffen.



Abbildung 13. Werteschaltfläche

- **Schnelles Suchen nach Dateien.** Verwenden Sie beim Suchen von Dateien in einem Dialogfeld "Öffnen" die Liste "Datei" (klicken Sie auf das gelbe Rautensymbol oben im Dialogfeld neben dem Feld "Suchen in"), um zuvor verwendete Verzeichnisse sowie IBM SPSS Modeler-Standardverzeichnisse aufzurufen. Verwenden Sie die Vor- und Zurückschaltflächen, um durch die Verzeichnisse zu blättern.
- **Minimieren überflüssiger Elemente im Ausgabefenster.** Sie können die Ausgabe schnell schließen und entfernen, indem Sie die Schaltfläche mit dem roten X verwenden, die Sie in allen Ausgabefenstern in der rechten oberen Ecke finden. So können Sie nur die vielversprechenden und interessanten Ergebnisse auf der Registerkarte "Ausgaben" des Managerbereichs beibehalten.

Für die Software steht eine vollständige Palette mit Tastenkombinationen zur Verfügung. Weitere Informationen finden Sie in „Tastatureingabehilfen“ auf Seite 245.

Wussten Sie, dass Sie ...

- ... mithilfe der Maus eine Gruppe von Knoten in den Streamerstellungsbereich ziehen und auswählen können?
- ... Knoten von einem Stream zu einem anderen kopieren und einfügen können?
- ... in jedem Dialogfeld und Ausgabefenster die Hilfe aufrufen können?
- ... Hilfe zu CRISP-DM (Cross-Industry Standard Process for Data Mining) aufrufen können? (Klicken Sie dazu im Hilfemenü auf **CRISP-DM-Hilfe**.)

Kapitel 6. Arbeiten mit Daten

Sie können die Daten eines Datenstroms untersuchen, indem Sie mit der rechten Maustaste auf einen Datenknoten klicken und **Daten anzeigen** auswählen. Im geöffneten Fenster haben Sie folgende Möglichkeiten:

- Mit der Registerkarte **Diagramm** können Sie erweiterte Datenvisualisierungen erstellen, um Ihre Daten aus verschiedenen Perspektiven zu untersuchen und Muster, Verbindungen sowie Beziehungen in Ihren Daten zu ermitteln.
- Auf der Registerkarte **Tabellenkalkulation** wird eine schreibgeschützte Ansicht Ihrer Daten im Tabellenformat angezeigt.
- Auf der Registerkarte **Data Audit** werden die Häufigkeit und Statistikdaten für jede Spalte angezeigt.
- Sie können auf der Registerkarte **Dashboard** auf **Layout-Design starten** klicken und ein Layout erstellen, in dem auf einer einzigen Seite mehrere Diagramme angezeigt werden können. Sie können Layouts als Vorlagen speichern und Ihre gespeicherten Diagramme per Drag-and-Drop zu Positionen in Ihrem Layout verschieben.
- Mit der Registerkarte **Vorgaben** können Sie Benutzerschnittstellenvorgaben wie die Sprache und das Erscheinungsbild festlegen.

Beachten Sie, dass diese Funktion standardmäßig Port 28900 verwendet. Wenn Sie einen anderen Port verwenden müssen, ändern Sie den Wert der Konfigurationseinstellung `data_view_port_number` in Ihrer Datei `options.cfg`.

Erstellen von Diagrammen

Auf der Registerkarte **Diagramme** können Sie aus vordefinierten Galeriediagrammen oder aus einzelnen Bestandteilen (z. B. Achsen und Balken) Diagramme erstellen. Sie erstellen ein Diagramm, indem Sie aus den bereitgestellten Diagrammtypoptionen einen Galeriediagrammtyp auswählen.

Während Sie das Diagramm erstellen, wird im Erstellungsbereich eine Vorschau des Diagramms angezeigt. In der Vorschau werden die tatsächlichen Variablenbeschriftungen und die Messniveaus verwendet, die für Ihre tatsächlichen Daten repräsentativ sind.

Für neue Benutzer bietet sich die Verwendung der Galerie an. Informationen zur Verwendung der Galerie finden Sie in „Erstellen eines Diagramms aus der Galerie“ auf Seite 74.

Starten der Diagrammerstellung

- Klicken Sie mit der rechten Maustaste auf den Datenknoten, mit dem Sie arbeiten wollen, und wählen Sie dann **Daten anzeigen** aus.

Layout und Begriffe

Eingangsanzeige

Nach dem Start der Diagrammerstellung können Sie einen Diagrammtyp oder Spalten im aktiven Dataset auswählen. Beim Hinzufügen von Spalten zur Visualisierung werden die Optionen für Diagrammtypen aktualisiert, um empfohlene Diagrammtypen für die ausgewählten Spalten anzuzeigen.

Erstellungsbereich

Der Erstellungsbereich im Dialogfeld "Diagrammerstellung" ist die Fläche, auf der Sie das Diagramm erstellen.

Diagrammtypen

Listet die verfügbaren Diagrammtypen auf. Die Grafikelemente sind Elemente im Diagramm, die Daten darstellen. Hierbei handelt es sich um die Balken, Punkte, Linien usw.

Bereich "Details"

Im Bereich "Details" werden die grundlegenden Diagrammbausteine bereitgestellt.

Diagrammeinstellungen

Stellt Optionen für die Auswahl der Variablen bereit, die zum Erstellen des Diagramms, der Verteilungsmethode, der Titel- und Untertitelfelder usw. verwendet werden. Die im Bereich "Details" bereitgestellten Optionen unterscheiden sich je nach dem ausgewählten Diagrammtyp. Detaillierte Informationen zu den Optionen, die für die einzelnen Diagrammtypen verfügbar sind, finden Sie in „Diagrammtypen“.

Aktionen

Stellt Optionen für das Herunterladen von Diagrammkonfigurationsdateien, Herunterladen von Diagrammen als Bilddateien, Zurücksetzen von Diagrammen und Festlegen der globalen Diagrammvorgaben bereit.

Erstellen eines Diagramms aus der Galerie

Die einfachste Methode, ein Diagramm zu erstellen, ist die Verwendung der Galerie. Im Folgenden finden Sie allgemeine Schritte zum Erstellen eines Diagramms aus der Galerie.

1. Wählen Sie im Abschnitt **Diagrammtypen** eine Diagrammkategorie aus. Im Diagrammerstellungsbereich wird eine Vorschauversion des ausgewählten Diagrammtyps angezeigt.
2. Wenn im Erstellungsbereich bereits ein Diagramm angezeigt wird, ersetzt das neue Diagramm das Achsenset und die Grafikelemente des Diagramms.
 - a. Je nach dem ausgewählten Diagrammtyp werden die verfügbaren Variablen unter einer Anzahl unterschiedlicher Überschriften im Bereich "Details" dargestellt (z. B. **Kategorie** für Balkendiagramme oder **X-Achse** und **Y-Achse** für Liniendiagramme). Wählen Sie die entsprechenden Variablen für den ausgewählten Diagrammtyp aus.

Diagrammtypen

Die Galerie enthält eine Sammlung der am häufigsten verwendeten Diagramme. Dazu gehören:

Streudiagramme und Punktdiagramme

1-D-, einfache, gruppierte, überlagerte und 3-D-Streudiagramme; Übersichtspunktdiagramme, 1-D-Punktdiagramme und Verbundliniendiagramme. Weitere Informationen finden Sie in „Streudiagramme und Punktdiagramme“ auf Seite 76.

Liniendiagramme

Einfach. Weitere Informationen finden Sie in „Liniendiagramme“ auf Seite 77.

Diagramme mit mehreren Zeitreihen

Einfach. Weitere Informationen finden Sie in „Diagramme mit mehreren Zeitreihen“ auf Seite 78.

Histogramme

Einfach, gestapelt und Häufigkeitspolygone. Weitere Informationen finden Sie in „Histogrammdiagramme“ auf Seite 79.

Bevölkerungspyramidendiagramme

Einfach. Weitere Informationen finden Sie in Bevölkerungspyramidendiagramme.

Q-Q-Diagramme

Einfach. Weitere Informationen finden Sie in „Q-Q-Diagramme“ auf Seite 75.

Kreisdiagramme

Einfach. Weitere Informationen finden Sie in „Kreisdiagramme“ auf Seite 80.

Balkendiagramme

Einfach, gestapelt und gruppiert. Weitere Informationen finden Sie in „Balkendiagramme“ auf Seite 81.

Paralleldiagramme

Einfach. Weitere Informationen finden Sie in „Paralleldiagramme“ auf Seite 82.

Beziehungsdiagramme

Einfach. Weitere Informationen finden Sie in „Beziehungsdiagramme“ auf Seite 83.

Boxplots

Einfach und gruppiert. Weitere Informationen finden Sie in „Boxplots“ auf Seite 84.

Treemap-Diagramme

Einfach und Sunburst. Weitere Informationen finden Sie in Treemap-Diagramme.

Landkartendiagramme

Einfach. Weitere Informationen finden Sie in „Landkartendiagramme“ auf Seite 86.

Heat-Maps

Einfach. Weitere Informationen finden Sie in „Heat-Map-Diagramme“ auf Seite 88.

t-SNE-Diagramme

Einfach. Weitere Informationen finden Sie in „t-SNE-Diagramme“ auf Seite 90.

Wortwolken

Einfach. Weitere Informationen finden Sie in „Wortwolkendiagramme“ auf Seite 90.

Fehlerbalkendiagramme

Einfach. Weitere Informationen finden Sie in „Fehlerbalkendiagramme“ auf Seite 91.

3-D-Diagramme

Einfach. Weitere Informationen finden Sie in „3-D-Diagramme“ auf Seite 92.

Streudiagrammmatrixdiagramme

Einfach. Weitere Informationen finden Sie in „Streudiagrammmatrixdiagramme“ auf Seite 93.

Candlestick-Diagramme

Einfach. Weitere Informationen finden Sie in „Candlestick-Diagramme“ auf Seite 94.

Diagramme mit doppelter Y-Achse

Einfach. Weitere Informationen finden Sie in Diagramme mit doppelter Y-Achse.

Benutzerdefinierte Diagramme

Weitere Informationen finden Sie in „Benutzerdefinierte Diagramme“ auf Seite 97.

Q-Q-Diagramme

Q-Q-Diagramme vergleichen zwei Wahrscheinlichkeitsverteilungen, indem ihre Quantile gegeneinander dargestellt werden. Mit einem Q-Q-Diagramm werden die Formen von Verteilungen verglichen. Hierdurch wird eine grafische Ansicht der Ähnlichkeit oder Unterschiedlichkeit von Eigenschaften wie Lage, Maßstab und Schiefe in den beiden Verteilungen bereitgestellt.

Erstellen eines einfachen Q-Q-Kurvendiagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Q-Q-Diagramm**.



Der Erstellungsbereich wird aktualisiert, um eine Q-Q-Kurvendiagrammvorlage anzuzeigen.

2. Wählen Sie eine Variable als Variable für die **X-Achse** aus.

Weitere Funktionen

X-Achse

Listet Datasetvariablen auf, die für die x -Achse des Diagramms verfügbar sind.

Verteilung

Die Dropdown-Liste stellt alle verfügbaren Verteilungsmethoden bereit.

Diagrammtyp

Wählen Sie entweder ein Q-Q-Diagramm (Quantil-Quantil) oder ein P-P-Diagramm (Prozent-Prozent) aus.

Verteilung aus den Daten schätzen

Bei Aktivierung werden Datenparameter für die ausgewählte **Verteilung** automatisch geschätzt. Bei Inaktivierung werden die Verteilungswerte für **Form** und **Skalierung** angezeigt.

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Streudiagramme und Punktdiagramme

Mit dem Grafikelement für Punktdiagramme können mehrere große Diagrammkategorien erstellt werden:

- **Streudiagramme.** Diese dienen zur grafischen Darstellung multivariater Daten. Sie können damit leichter potenzielle Beziehungen zwischen metrischen Variablen erkennen. Bei einfachen Streudiagrammen werden zwei Variablen in einem zweidimensionalen Koordinatensystem dargestellt. Bei 3-D-Streudiagrammen werden drei Variablen in einem dreidimensionalen Koordinatensystem dargestellt. Wenn Sie weitere Variablen darstellen möchten, können Sie dies mit überlagerten Streudiagrammen und Streudiagrammmatrizen (SPLOMs) versuchen. Bei überlagerten Streudiagrammen werden überlagerte Paare von x - y -Variablen angezeigt, wobei sich die einzelnen Paare durch Farbe oder Form unterscheiden. Bei einer Streudiagrammmatrix wird eine Matrix mit 2-D-Streudiagrammen erstellt, wobei jede Variable anhand jeder anderen Variablen in der Streudiagrammmatrix geplottet wird.
- **Punktdiagramme.** Wie bei Histogrammen wird in diesen Diagrammen die Verteilung einer einzelnen metrischen Variablen angezeigt. Die Daten werden klassiert (Binning), aber anstatt dass in jeder Klasse genau ein Wert vorliegt (wie bei Anzahl/Häufigkeit), werden alle Punkte in jeder Klasse angezeigt und gestapelt. Diese Diagramme werden zuweilen als Dichtediagramme bezeichnet.
- **Übersichtspunktdiagramme.** Diese Diagramme weisen eine große Ähnlichkeit mit Balkendiagrammen auf, mit der Ausnahme, dass dort, wo sich die Oberkante der Balken befinden würden, Punkte angezeigt werden. Da Übersichtspunktdiagramme Balkendiagrammen so ähnlich sind, finden Sie unter „Balkendiagramme“ auf Seite 81 Informationen zu ihrer Erstellung entnehmen.
- **Verbundliniendiagramme.** Hierbei handelt es sich um einen speziellen Typ von Übersichtspunktdiagrammen. Die Punkte werden gruppiert und die Punkte in den einzelnen Kategorien werden durch eine Linie verbunden. Das Verbundliniendiagramm dient zum Vergleich einer Statistik über die verschiedenen kategorialen Variablen.

Erstellen eines einfachen Streudiagrammes

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Streudiagramm**.

Scatter plot



Der Erstellungsbereich wird aktualisiert, um eine Streudiagrammvorlage anzuzeigen.

2. Wählen Sie eine metrische Variable als Variable für die **X-Achse** aus.

3. Wählen Sie eine metrische Variable als Variable für die **Y-Achse** aus. Sie brauchen keine Statistik anzugeben, da Streudiagramme normalerweise Rohwerte anzeigen.

Weitere Funktionen

X-Achse

Listet Datasetvariablen auf, die für die x -Achse des Diagramms verfügbar sind.

Y-Achse

Listet Datasetvariablen auf, die für die y -Achse des Diagramms verfügbar sind.

Anpassungslinie

Bei Anpassungslinien werden die Datenpunkte an eine Linie angepasst, die normalerweise nicht durch alle Datenpunkte verläuft. Die Anpassungslinie zeigt den Trend der Daten auf. Einige Anpassungslinien beruhen auf Regression. Andere beruhen auf iterativen gewichteten kleinsten Quadraten. Wählen Sie in der Dropdown-Liste eine Anpassungslinienoption aus.

Farbzuordnungstabelle

Listet die verfügbaren Variablen der Farbzuordnungstabelle auf. Diese Variablen verwenden den Farbverlauf basierend auf dem Wertebereich in der angegebenen Spalte zur Darstellung in den Diagrammpunkten. Farbzuordnungstabellen sind auch als Choroplethenkarten bekannt.

Größenzuordnungstabelle

Listet die verfügbaren Variablen der Größenzuordnungstabelle auf. Diese Variablen verwenden unterschiedliche Größen für die Darstellung in den Zeichenpunkten.

Formzuordnungstabelle

Listet die verfügbaren Variablen der Formzuordnungstabelle auf. Diese Variablen verwenden unterschiedliche Formen für die Darstellung in den Zeichenpunkten.

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Liniendiagramme

Ein Liniendiagramm stellt eine Reihe von Datenpunkten in einem Diagramm dar und verbindet sie mit Linien. Ein Liniendiagramm ist besonders nützlich, wenn Trendkurven mit geringfügigen Unterschieden oder mit sich überkreuzenden Datenlinien angezeigt werden. Sie können ein Liniendiagramm zur Zusammenfassung kategorialer Variablen verwenden. In diesem Fall ähnelt es einem Balkendiagramm (siehe „Balkendiagramme“ auf Seite 81). Liniendiagramme sind auch für Zeitreihendaten nützlich.

Erstellen eines einfachen Liniendiagramms mit Zeitreihen

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Linie**.

Line •



Der Erstellungsbereich wird aktualisiert, um eine Liniendiagrammvorlage anzuzeigen.

2. Wählen Sie eine Datumsvariable als Variable für die **X-Achse** aus.
3. Wählen Sie eine metrische Variable als Variable für die **Y-Achse** aus. Hierbei handelt es sich um die Variable, deren Werte im Laufe der Zeit erfasst wurden.

Weitere Funktionen

X-Achse

Listet Datasetvariablen auf, die für die x -Achse des Diagramms verfügbar sind.

Y-Achse

Listet Datasetvariablen auf, die für die y -Achse des Diagramms verfügbar sind.

Fläche Bei Aktivierung wird die Fläche unterhalb der Linie in einer anderen Farbe gefüllt.

Aufteilen nach

Wählen Sie eine kategoriale Variable aus, die eine Tabelle mit Diagrammen erstellt, die für jede Kategorie in der Variablen "Aufteilen nach" eine Zelle enthält. Wie bei der Gruppierung werden dem Diagramm bei der Aufteilung nach Variablen vor allem weitere Dimensionen hinzugefügt, indem Informationen für jede Variablenkategorie angezeigt werden. Weitere Informationen finden Sie in Variablen "Aufteilen nach".

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Diagramme mit mehreren Zeitreihen

Diagramme mit mehreren Zeitreihen ähneln Liniendiagrammen, mit der Ausnahme, dass Sie auf der y -Achse mehrere Felder darstellen können.

Erstellen eines einfachen Diagramms mit mehreren Zeitreihen

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Mehrere Zeitreihen**.

Multi-series



Der Erstellungsbereich wird aktualisiert, um eine Vorlage eines Diagramms mit mehreren Zeitreihen anzuzeigen.

2. Wählen Sie eine Variable als Variable für die **X-Achse** aus.
3. Wählen Sie mindestens zwei metrische Variablen als Variablen für die **Y-Achse** aus.

Weitere Funktionen

X-Achse

Listet Datasetvariablen auf, die für die x -Achse des Diagramms verfügbar sind.

Y-Achse

Listet Datasetvariablen auf, die für die y -Achsen des Diagramms verfügbar sind.

Wählen Sie den Diagrammtyp (Linie oder Balken) in der Dropdown-Liste aus.

Klicken Sie auf **Weitere Spalte hinzufügen**, um weitere Spalten in das Diagramm einzuschließen.

Daten normalisieren

Bei Aktivierung transformiert diese Einstellung Daten zu einer Normalverteilung, mit der Sie Daten aus mehreren Datensets oder mehreren Spalten vergleichen können. Diese Einstellung erstellt 100 % Stapelung für Anzahlwerte und wandelt Statistiken in Prozente um.

Glatt Bei Aktivierung werden die geraden Linien des Diagramms etwas gebogener dargestellt.

Wendepunkt anzeigen

Bei Aktivierung sind die Wendepunkte des Diagramms sichtbar.

Sekundäre Y-Achse

Listet Datasetvariablen auf, die für die sekundären y -Achsen des Diagramms verfügbar sind.

Wählen Sie den Diagrammtyp (Linie oder Balken) in der Dropdown-Liste aus.

Klicken Sie auf **Weitere Spalte hinzufügen**, um weitere Spalten in das Diagramm einzuschließen.

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Histogrammdiagramme

Ein Histogramm ähnelt hinsichtlich der Darstellung einem Balkendiagramm. Anstatt Kategorien zu vergleichen oder nach Trends im Zeitverlauf zu suchen, stellt jeder Balken dar, wie Daten in einer einzelnen Kategorie verteilt sind. Jeder Balken stellt einen fortlaufenden Datenbereich oder die Anzahl Häufigkeiten für einen bestimmten Datenpunkt dar.

Histogramme dienen zur Anzeige der Verteilung einer einzelnen metrischen Variablen. Die Daten werden klassiert und mithilfe einer Häufigkeits- oder Prozentsatzstatistik ausgewertet. Eine Histogrammvariante ist das Häufigkeitspolygon, das wie ein typisches Histogramm ist, mit der Ausnahme, dass anstelle des Grafikelements "Balken" das Grafikelement "Fläche" verwendet wird.

Eine weitere Histogrammvariante ist die Bevölkerungspyramide. Ihr Name leitet sich von ihrem häufigsten Verwendungszweck ab: Auswertung von Bevölkerungsdaten. Bei der Verwendung mit Bevölkerungsdaten wird sie nach Geschlecht aufgeteilt und bietet so zwei horizontale Histogramme mit Altersdaten links und rechts von einer Mittelachse. In Ländern mit einer jungen Bevölkerung ähnelt die Form des sich ergebenden Diagramms einer Pyramide.

Erstellen eines Histogrammdiagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Histogramm**.

Histogramm •



Der Erstellungsbereich wird aktualisiert, um eine Histogrammdiagrammvorlage anzuzeigen.

2. Wählen Sie eine metrische Variable als Variable für die **X-Achse** aus.

Anmerkung: Als Statistik für Histogramme kommen "Histogramm" und "Histogrammprozent" infrage. Bei diesen Statistiken werden die Daten klassiert (d. h. in Klassen eingeteilt) und für jede Klasse wird die Anzahl berechnet.

Weitere Funktionen

X-Achse

Listet Datasetvariablen auf, die für die x -Achse des Diagramms verfügbar sind.

Aufteilen nach

Wählen Sie eine kategoriale Variable aus, die eine Tabelle mit Diagrammen erstellt, die für jede Kategorie in der Variablen "Aufteilen nach" eine Zelle enthält. Wie bei der Gruppierung werden dem Diagramm bei der Aufteilung nach Variablen vor allem weitere Dimensionen hinzugefügt, indem Informationen für jede Variablenkategorie angezeigt werden. Weitere Informationen finden Sie in Variablen "Aufteilen nach".

Verteilungskurve einblenden

Bei Aktivierung wird die Verteilungsanpassungskurve im Diagramm angezeigt.

Klassenbreite

Der Schieberegler steuert die Größe des Intervalls, das zum Aufteilen der Daten in Gruppen verwendet wird.

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Kreisdiagramme

Kreisdiagramme dienen zum Vergleichen von Anteilen. Sie können beispielsweise ein Kreisdiagramm verwenden, um zu zeigen, dass in einem Kurs ein höherer Frauenanteil vorliegt.

Erstellen eines einfachen Kreisdiagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Kreis**.



Der Erstellungsbereich wird aktualisiert, um eine Kreisdiagrammvorlage anzuzeigen.

2. Wählen Sie eine kategoriale (nominale oder ordinale) Variable in der Liste **Kategorie** aus. Die Kategorien in dieser Variablen bestimmen die Anzahl der Ausschnitte im Kreisdiagramm.
3. Wählen Sie eine statistische Zusammenfassungsfunktion für das Grafikelement aus. Bei Kreisdiagrammen wird normalerweise eine anzahlbasierte Statistik oder eine Summe gewünscht. Das Ergebnis der Statistik bestimmt die Größe der einzelnen Ausschnitte.

Weitere Funktionen

Kategorie

Wählen Sie eine kategorische (nominale oder ordinale) Variable aus, die die Anzahl der Ausschnitte im Kreisdiagramm bestimmt.

Kreisdiagrammtyp

Die folgenden Stile sind verfügbar:

Normalverteilung

Die Kreissegmente werden als normale Ausschnitte angezeigt.

Ring Die Kreissegmente werden als Ring angezeigt. Dieser Stil ist auch als Ringdiagramm bekannt.

Rose Im Gegensatz zum normalen Kreisdiagramm, das einen gemeinsamen Radius verwendet, variieren die Kreissegmentgrößen in Abhängigkeit von ihrem Wert.

Zusammenfassung

Wählen Sie eine statistische Zusammenfassungsfunktion für das Grafikelement aus. Bei Kreisdiagrammen wird normalerweise eine anzahlbasierte Statistik oder eine Summe gewünscht. Das Ergebnis der Statistik bestimmt die Größe der einzelnen Ausschnitte.

Es gibt zwei Typen von statistischen Auswertungsfunktionen. Der Unterschied ist wichtig, da er festlegt, ob Sie eine Wertvariable angeben müssen.

- **Funktionen, die keine Wertvariable benötigen.** Hierbei handelt es sich um Funktionen, die keine Wertvariable benötigen. Alle Anzahl- und Prozentsatzstatistiken fallen in diese Kategorie. Diese Statistiken sind verfügbar, wenn keine Wertvariable definiert ist.
- **Funktionen, die eine Wertvariable benötigen.** Hierbei handelt es sich um Funktionen, die eine Wertvariable benötigen. Für die Funktion *Mittelwert* beispielsweise ist eine Variable erforderlich, anhand deren der Mittelwert berechnet wird. Diese Statistiken sind verfügbar, wenn eine Wertvariable definiert ist.

Wert Dieses Feld wird angezeigt, wenn eine Auswertungsfunktion, die eine metrische Variable erfordert, ausgewählt ist. Wählen Sie eine Variable als metrische Variable aus.

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Balkendiagramme

Balkendiagramme dienen zur Zusammenfassung von kategorialen Variablen. Sie können beispielsweise mithilfe eines Balkendiagramms die Anzahl der Männer bzw. Frauen anzeigen, die an einer Umfrage teilnahmen, oder mithilfe eines Balkendiagramms das mittlere Gehalt von Männern und das mittlere Gehalt von Frauen anzeigen.

Erstellen eines einfachen Balkendiagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Balken**.

Bar •



Der Erstellungsbereich wird aktualisiert, um eine Balkendiagrammvorlage anzuzeigen.

2. Wählen Sie eine kategoriale (nominale oder ordinale) Variable als Variable für die **Kategorie** aus. Sie können eine metrische Variable verwenden, die Ergebnisse sind jedoch nur in einigen Sonderfällen brauchbar. Balkendiagramme sind am klarsten, wenn sie aus einer überschaubaren Anzahl unterschiedlicher Werte bestehen. Wenn Sie ein Balkendiagramm mit einer metrischen Kategorienachse erstellen, ergeben sich sehr dünne Balken, da jeder Balken an einem exakten Wert gezeichnet wird und sich nicht mit anderen stetigen Werten überschneiden kann.
3. Wählen Sie eine Statistik in der Liste **Zusammenfassung** aus. Das Ergebnis der Statistik bestimmt die Höhe der Balken. Wenn die gewünschte Statistik nicht in der Liste **Zusammenfassung** aufgeführt wird, ist dafür möglicherweise eine Variable erforderlich. Wählen Sie eine Variable in der Werteliste aus und prüfen Sie, ob die Statistik jetzt verfügbar ist. Es können noch andere Einschränkungen bezüglich des Diagrammtyps gelten. So können Fehlerbalken nur für bestimmte Statistiken berechnet werden.

Weitere Funktionen

Kategorie

Listet Datasetvariablen auf, die für die x -Achse des Diagramms verfügbar sind.

Sortierbasis

Wählen Sie eine Sortieroption für die Kategorien in der Variablen aus.

Kategorienname

Die Kategorien der Variablen werden nach den Kategoriebeschriftungen sortiert. Diese Beschriftungen werden im Diagramm eingeblendet, in der Regel als Beschriftungen der Teilstriche oder der Legende.

Kategorienwert

Die Kategorien der Variablen werden nach dem im Dataset gespeicherten Wert sortiert. Der Wert der Kategorie identifiziert die Kategorie im Dataset. Dieser Wert weicht häufig von der Beschriftung ab und ist nicht unbedingt aussagekräftig. Der Wert kann beispielsweise aus einer Zahl bestehen (z. B. 1), die Beschriftung dagegen aus einer Textbeschreibung der Kategorie (z. B. *Weiblich*).

Kategorienreihenfolge

Wählen Sie die Reihenfolge aus, in der die Variablenkategorien sortiert werden.

Wie gelesen

Variablenkategorien werden so dargestellt, wie sie im Dataset vorkommen.

Aufsteigend

Sortiert Variablenkategorien in aufsteigender Reihenfolge.

Absteigend

Sortiert Variablenkategorien in absteigender Reihenfolge.

Zusammenfassung

Wählen Sie eine statistische Zusammenfassungsfunktion für das Grafikelement aus. Das Ergebnis der Statistik bestimmt die Position der Grafikelemente auf der y -Achse. In 2-D-Diagrammen wird die Statistik für die einzelnen Werte auf der x -Achse berechnet. In 3-D-Diagrammen wird die Statistik für die Schnittpunkte der Werte auf der x -Achse und der z -Achse berechnet.

Es gibt zwei Typen von statistischen Auswertungsfunktionen. Der Unterschied ist wichtig, da er festlegt, ob Sie eine Wertvariable als y -Achse angeben müssen.

- **Funktionen, die keine Wertvariable benötigen.** Hierbei handelt es sich um Funktionen, die keine Variable für die y -Achsen benötigen. Alle Anzahl- und Prozentsatzstatistiken fallen in diese Kategorie. Diese Statistiken sind verfügbar, wenn keine Wertvariable definiert ist.
- **Funktionen, die eine Wertvariable benötigen.** Hierbei handelt es sich um Funktionen, die eine Wertvariable für die y -Achse benötigen. Für die Funktion *Mittelwert* beispielsweise ist eine Variable erforderlich, anhand deren der Mittelwert berechnet wird. Diese Statistiken sind verfügbar, wenn eine Wertvariable definiert ist.

Wert Dieses Feld wird angezeigt, wenn eine Auswertungsfunktion, die eine y -Achsenvariable erfordert, ausgewählt ist. Wählen Sie eine Variable als y -Achsenwert aus.

Aufteilen nach

Wählen Sie eine kategoriale Variable aus, die eine Tabelle mit Diagrammen erstellt, die für jede Kategorie in der Variablen "Aufteilen nach" eine Zelle enthält. Wie bei der Gruppierung werden dem Diagramm bei der Aufteilung nach Variablen vor allem weitere Dimensionen hinzugefügt, indem Informationen für jede Variablenkategorie angezeigt werden. Weitere Informationen finden Sie in Variablen "Aufteilen nach".

Aufteilungstyp

Wenn eine Variable **Aufteilen nach** ausgewählt ist, können Sie die resultierenden Kategorienbalken als gestapelt oder gruppiert anzeigen. Durch Clustering und Stapelung wird die Dimensionalität im Diagramm erhöht. Beim Clustering wird ein Balken in mehrere Balken aufgestapelt und beim Stapeln werden in jedem Balken Segmente erstellt. Achten Sie darauf, die richtige Statistik für die Stapelung zu verwenden. Beim Addieren (Stapeln) der Werte muss sich ein sinnvolles Resultat ergeben. So ist beispielsweise die Addition und Stapelung von Mittelwerten (Durchschnittswerten) normalerweise nicht sinnvoll.

Transponieren

Bei Aktivierung werden die x - und y -Achsen des Diagramms transponiert.

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Paralleldiagramme

Paralleldiagramme sind nützlich, wenn Sie eine Geometrie mit hohen Dimensionen visualisieren und multivariate Daten analysieren. Paralleldiagramme ähneln Liniendiagrammen für Zeitreihendaten, die Achsen entsprechen jedoch nicht den Zeitpunkten (es gibt keine natürliche Reihenfolge).

Erstellen eines einfachen Paralleldiagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Parallel**.

Parallel



Der Erstellungsbereich wird aktualisiert, um eine Paralleldiagrammvorlage anzuzeigen.

2. Wählen Sie mindestens zwei Variablen als Spaltenvariablen aus. Jede Spalte stellt eine vertikale, parallele Achse im Diagramm dar.

Anmerkung: Die Feldreihenfolge ist bei der Suche nach Merkmalen wichtig. Bei einer Standarddatenanalyse müssen Sie die Reihenfolge der Felder möglicherweise mehrmals ändern.

Weitere Funktionen

Spalten

Listet Datasetvariablen auf, die für die y -Achsen des Diagramms verfügbar sind.

Klicken Sie auf **Weitere Spalte hinzufügen**, um weitere Spalten hinzuzufügen.

Farbzuordnungstabelle

Listet die verfügbaren Variablen der Farbzuordnungstabelle auf. Diese Variablen verwenden den Farbverlauf basierend auf dem Wertebereich in der angegebenen Spalte zur Darstellung in den Diagrammpunkten. Farbzuordnungstabellen sind auch als Choroplethenkarten bekannt.

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Beziehungsdiagramme

Ein Beziehungsdiagramm eignet sich gut zum Ermitteln der Beziehung von Variablen zueinander.

Erstellen eines einfachen Beziehungsdiagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Beziehung**.

Relationship



Der Erstellungsbereich wird aktualisiert, um eine Beziehungsdiagrammvorlage anzuzeigen.

2. Wählen Sie mindestens zwei Variablen als Spaltenvariablen aus.

Weitere Funktionen

Spalten

Listet die verfügbaren Datasetvariablen auf.

Klicken Sie auf **Weitere Spalte hinzufügen**, um weitere Spalten hinzuzufügen.

Liniensstil

Steuert den Liniensstil zwischen verwandten Datenpunkten.

Kurve Bei Auswahl dieser Option werden zwischen verwandten Datenpunkten gekrümmte Linien gezeichnet.

Gerade

Bei Auswahl dieser Option werden zwischen verwandten Datenpunkten gerade Linien gezeichnet.

Beschriftungsschwellenwert

Zeigt Beschriftungen für Datenpunkte an, deren Werte den definierten Wert überschreiten.

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Boxplots

In einem Boxplot-Diagramm werden alle Werte der Fünf-Punkte-Zusammenfassung (Minimum, erstes Quartil, Median, drittes Quartil und Maximum) angezeigt. Es wird zur Anzeige der Verteilung einer metrischen Variablen und zur leichten Identifizierung von Ausreißern verwendet.

Erstellen eines einfachen Boxplots

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Boxplot**.



Der Erstellungsbereich wird aktualisiert, um eine Boxplot-Diagrammvorlage anzuzeigen.

2. Wählen Sie mindestens eine metrische Variable als Spaltenvariable aus.

Anmerkung: Die Statistik für ein Punktdiagramm ist Boxplot. Dies kann nicht geändert werden.

Weitere Funktionen

Spalten

Listet Datasetvariablen auf, die für die x -Achse des Diagramms verfügbar sind.

Klicken Sie auf **Weitere Spalte hinzufügen**, um weitere Spalten hinzuzufügen.

Aufteilen nach

Wählen Sie eine kategoriale Variable aus, die eine Tabelle mit Diagrammen erstellt, die für jede Kategorie in der Variablen "Aufteilen nach" eine Zelle enthält. Wie bei der Gruppierung werden dem Diagramm bei der Aufteilung nach Variablen vor allem weitere Dimensionen hinzugefügt, indem Informationen für jede Variablenkategorie angezeigt werden. Weitere Informationen finden Sie in Variablen "Aufteilen nach".

Kategorienreihenfolge

Wählen Sie die Reihenfolge aus, in der die Variablenkategorien sortiert werden.

Wie gelesen

Variablenkategorien werden so dargestellt, wie sie im Dataset vorkommen.

Aufsteigend

Sortiert Variablenkategorien in aufsteigender Reihenfolge.

Absteigend

Sortiert Variablenkategorien in absteigender Reihenfolge.

Daten normalisieren

Bei Aktivierung transformiert diese Einstellung Daten zu einer Normalverteilung, mit der Sie Daten aus mehreren Datasets oder mehreren Spalten vergleichen können. Diese Einstellung erstellt 100 % Stapelung für Anzahlwerte und wandelt Statistiken in Prozente um.

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Kreispackungsdiagramme

Kreispackungsdiagramme zeigen hierarchische Daten als eine Gruppe von verschachtelten Bereichen an. So können Sie eine große Menge an hierarchisch strukturierten Daten visuell darstellen. Dies ähnelt der Treemap-Darstellung, allerdings werden anstelle von Rechtecken Kreise verwendet. Kreispackungsdiagramme zeigen Hierarchiedaten mithilfe von Einschlüssen an (Verschachtelung).

Erstellen eines einfachen Kreispackungsdiagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Zeitdiagramm**.

Circle packing



Der Erstellungsbereich wird aktualisiert, um eine Vorlage für Kreispackungsdiagramme anzuzeigen.

2. Wählen Sie eine Variable für **Spalten** in der Dropdown-Liste aus.

Anmerkung: Klicken Sie auf **Weitere Spalte hinzufügen**, um weitere Spaltenvariablen einzuschließen.

Weitere Funktionen

Farbe gruppieren

Aktivieren oder inaktivieren Sie Farbgruppierungen.

Zusammenfassung

Wählen Sie eine statistische Zusammenfassungsfunktion aus. Mit dieser Methode werden die einzelnen Kategorien zusammengefasst.

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Blasendiagramme

Blasendiagramme zeigen Kategorien in Ihren Gruppen als gepackte Kreise ohne Hierarchie an. Die Größe jedes Kreises (Blase) ist proportional zu seinem Wert. Blasendiagramme sind für den Vergleich von Beziehungen in Ihren Daten hilfreich.

Erstellen eines einfachen Blasendiagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Blase**.

Bubble



Der Erstellungsbereich wird aktualisiert, um eine Blasendiagrammvorlage anzuzeigen.

2. Wählen Sie eine Variable für **Spalten** in der Dropdown-Liste aus.

Anmerkung: Klicken Sie auf **Weitere Spalte hinzufügen**, um weitere Spaltenvariablen einzuschließen.

Weitere Funktionen

Farbe gruppieren

Aktivieren oder inaktivieren Sie Farbgruppierungen.

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Landkartendiagramme

Landkartendiagramme werden gewöhnlich verwendet, um Werte zu vergleichen und Kategorien in geografischen Regionen anzuzeigen. Sie eignen sich am besten, wenn die Daten geografische Informationen (wie Länder, Regionen, Staaten, Bundesländer, Postleitzahlen usw.) enthalten.

Erstellen eines einfachen Landkartendiagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Karte**.



Der Erstellungsbereich wird aktualisiert, um eine Landkartendiagrammvorlage anzuzeigen.

2. Wählen Sie in der Dropdown-Liste **Längengrad** eine Variable aus, die als Längengradvariable fungieren soll.
3. Wählen Sie in der Dropdown-Liste **Breitengrad** eine Variable aus, die als Breitengradvariable fungieren soll.

Weitere Funktionen

Kartenservice

Listet die Services auf, die für das Bereitstellen von Kartengrafiken verfügbar sind. Siehe „Kartenserviceoptionen“ auf Seite 87.

Typ Listet die Diagrammtypen auf, die für die Darstellung der Daten verfügbar sind.

Längengrad

Listet die Variablen auf, die als Längengradwert zur Verfügung stehen.

Breitengrad

Listet die Variablen auf, die als Breitengradwert zur Verfügung stehen.

QuickInfo-Info

Listet die Variablen auf, die zum Generieren von QuickInfos verwendet werden können, wenn Sie den Mauszeiger auf einen Datenpunkt bewegen.

Größenordnungstabelle

Listet die verfügbaren Variablen der Größenordnungstabelle auf. Diese Variablen verwenden unterschiedliche Größen für die Darstellung in den Zeichenpunkten.

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Kartenserviceoptionen: Sie können im Bereich "Details" unter "Kartendiagramm" den Kartenservice auswählen, der für die Bereitstellung von Kartengrafiken verwendet werden soll.

Konfigurieren von Kartenserviceoptionen

1. Halten Sie den IBM SPSS Modeler Server an.
2. Wenn Sie lokale Kartendateien mit GeoJSON-Formatierung anwenden möchten, müssen Sie diese in das Verzeichnis <Modeler-Installationsverzeichnis>/dataview/conf/public/mapfiles einfügen. Sollte das Verzeichnis nicht vorhanden sein, müssen Sie dieses erstellen.
3. Öffnen Sie die Datei <Modeler-Installationsverzeichnis>/dataview/conf/application.conf in einem Texteditor und fügen Sie folgende Optionen zum Abschnitt map.resources hinzu. In diesem Beispiel werden die beiden neuen Karten Welt - Lokale Karte und Welt - Online-Karte hinzugefügt.

```
map.resources {
  "mapservices" :
  [
    {
      "type"      : "geojson",
      "name"      : "Welt (integriert)",
      "location"  : "built_in_world",
      "id"        : "built_in_world"
    },
    {
      "type"      : "geojson",
      "name"      : "Leer (integriert)",
      "location"  : "built_in_blank",
      "id"        : "built_in_blank"
    },
    {
      "type"      : "geojson",
      "name"      : "Welt - Lokale Karte",
      "location"  : "/mapfiles/world.json",
      "useProxy"  : "false",
      "id"        : "world_local"
    },
    {
      "type"      : "geojson",
      "name"      : "Welt - Online-Karte",
      "location"  : "http://map.example.com/map/world.json",
      "useProxy"  : "true",
      "id"        : "world_online"
    }
  ]
}
```

Folgende Parameter stehen für die Kartenooptionen zur Verfügung:

Tabelle 4. Parameter für den Kartenservice

Option	Wert	Typ	Erforderlich?	Beschreibung
Typ	geojson	Zeichenfolge	Ja	Für diese Option wird derzeit nur der Wert geojson unterstützt.
name	Kartenname	Zeichenfolge	Ja	Geben Sie einen Namen für die Karte ein. Dieser Name wird im Bereich Details für Kartendiagramme im Dropdown-Menü Kartenservice angezeigt.
location	Kartenposition	Zeichenfolge	Ja	Geben Sie die URL oder den lokalen Dateispeicherort der Karte ein. Die URL muss mit http:// oder https:// beginnen. Die lokale Datei muss mit dem Pfad /mapfiles/ beginnen.
useProxy	true oder false	Zeichenfolge	Ja	Verwenden Sie true, wenn die Kartenposition eine URL ist. Verwenden Sie false, wenn die Kartenposition eine lokale Datei ist.
id	Karten-ID	Zeichenfolge	Ja	Verwenden Sie für jede Karte eine eindeutige ID.

Anmerkung:

- Entfernen Sie nicht die Standardoptionen **Welt (integriert)** und **Leer (integriert)**.
- Denken Sie daran, zwischen den Inhaltsblöcken ein Komma zu verwenden (siehe Beispiel), und stellen Sie sicher, dass Ihr neuer Abschnitt map.resources gültige Inhalte in JSON-Formatierung enthält.
- Der IBM SPSS Modeler Server (oder der eigenständige SPSS Modeler Client ohne Server) muss Zugriff auf die Kartenposition haben.

4. Starten Sie den IBM SPSS Modeler Server.

Wenn Sie nun mit Kartendiagrammen arbeiten, wird Ihnen im Bereich **Details** der neue, von Ihnen hinzugefügte Kartenservice angezeigt.

Heat-Map-Diagramme

Heat-Map-Diagramme stellen Daten dar, in denen die einzelnen in einer Matrix enthaltenen Werte als Farben dargestellt werden.

Erstellen eines einfachen Heat-Map-Diagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Heat-Map**.



Der Erstellungsbereich wird aktualisiert, um eine Heat-Map-Diagrammvorlage anzuzeigen.

2. Wählen Sie eine Variable als Spaltenvariable aus. Jede Variablenkategorie wird als einzelne Diagrammspalte dargestellt.
3. Wählen Sie eine Variable als Zeilenvariable aus. Jede Variablenkategorie wird als einzelne Diagrammzeile dargestellt.

Weitere Funktionen

Spalte Listet Datasetvariablen auf, die für die Spalten des Diagramms verfügbar sind. Jede Variablenkategorie wird als einzelne Diagrammspalte dargestellt.

Zeile Listet Datasetvariablen auf, die für die Zeilen des Diagramms verfügbar sind. Jede Variablenkategorie wird als einzelne Diagrammzeile dargestellt.

Kategorienreihenfolge

Wählen Sie die Reihenfolge aus, in der die Variablenkategorien sortiert werden.

Wie gelesen

Variablenkategorien werden so dargestellt, wie sie im Dataset vorkommen.

Aufsteigend

Sortiert Variablenkategorien in aufsteigender Reihenfolge.

Absteigend

Sortiert Variablenkategorien in absteigender Reihenfolge.

Zusammenfassung

Wählen Sie eine statistische Zusammenfassungsfunktion für das Grafikelement aus. Das Ergebnis der Statistik bestimmt die Position der Grafikelemente auf der y -Achse. In 2-D-Diagrammen wird die Statistik für die einzelnen Werte auf der x -Achse berechnet. In 3-D-Diagrammen wird die Statistik für die Schnittpunkte der Werte auf der x -Achse und der z -Achse berechnet.

Es gibt zwei Typen von statistischen Auswertungsfunktionen. Der Unterschied ist wichtig, da er festlegt, ob Sie eine Wertvariable als y -Achse angeben müssen.

- **Funktionen, die keine Wertvariable benötigen.** Hierbei handelt es sich um Funktionen, die keine Variable für die y -Achsen benötigen. Alle Anzahl- und Prozentsatzstatistiken fallen in diese Kategorie. Diese Statistiken sind verfügbar, wenn keine Wertvariable definiert ist.
- **Funktionen, die eine Wertvariable benötigen.** Hierbei handelt es sich um Funktionen, die eine Wertvariable für die y -Achse benötigen. Für die Funktion *Mittelwert* beispielsweise ist eine Variable erforderlich, anhand deren der Mittelwert berechnet wird. Diese Statistiken sind verfügbar, wenn eine Wertvariable definiert ist.

Wert Dieses Feld wird angezeigt, wenn eine Auswertungsfunktion, die eine y -Achsenvariable erfordert, ausgewählt ist. Wählen Sie eine Variable als y -Achsenwert aus.

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

t-SNE-Diagramme

T-distributed Stochastic Neighbor Embedding (t-SNE) ist ein Algorithmus für maschinelles Lernen zur Visualisierung. t-SNE-Diagramme modellieren jedes hochdimensionale Objekt mithilfe eines zwei- oder dreidimensionalen Punkts, sodass ähnliche Objekte mithilfe von nahen Punkten und unähnliche Objekte mithilfe von fernen Punkten mit hoher Wahrscheinlichkeit modelliert werden.

Erstellen eines einfachen t-SNE-Diagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **t-SNE**.



Der Erstellungsbereich wird aktualisiert, um eine t-SNE-Diagrammvorlage anzuzeigen.

2. Legen Sie die Werte für **Perplexität**, **Lernrate** und **Maximalzahl der Iterationen** fest.
3. Wählen Sie optional eine Variable **Farbzuordnungstabelle** aus.

Weitere Funktionen

Perplexität

Legt eine Zahl fest, die eine begründete Vermutung für die Anzahl naher Nachbarn für die einzelnen Datenpunkte ist. Der Zweck besteht darin, die lokalen und globalen Aspekte für Ihre Daten auszugleichen.

Lernrate

Dieser Wert wirkt sich auf die Geschwindigkeit des Lernens aus, indem die Änderung der Gewichtunggröße bei jeder Iteration angegeben wird.

Maximalzahl der Iterationen

Die maximale Anzahl der auszuführenden Iterationen.

Farbzuordnungstabelle

Listet die verfügbaren Variablen der Farbzuordnungstabelle auf. Diese Variablen verwenden den Farbverlauf basierend auf dem Wertebereich in der angegebenen Spalte zur Darstellung in den Diagrammpunkten. Farbzuordnungstabellen sind auch als Choroplethenkarten bekannt.

Primärtitel

Der Diagrammtitel.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Wortwolkendiagramme

Wortwolkendiagramme stellen Daten als Wörter dar, wobei die Größe und Position jedes einzelnen Worts durch seine Gewichtung bestimmt wird.

Erstellen eines einfachen Wortwolkendiagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Wortwolke**.



Der Erstellungsbereich wird aktualisiert, um eine Wortwolkendiagrammvorlage anzuzeigen.

2. Wählen Sie eine Variable als Quellvariable aus. Jede Variablenkategorie wird im Diagramm basierend auf ihrem Gewichtungswert dargestellt.
3. Wählen Sie einen Formwert für das Diagramm aus. Die resultierenden Diagrammdaten werden in der ausgewählten Form dargestellt.

Weitere Funktionen

Quelle

Listet Datasetvariablen auf, die für die Diagrammquelle verfügbar sind. Jede Variablenkategorie wird im Diagramm basierend auf ihrem Gewichtungswert dargestellt.

Form Listet die verfügbaren Diagrammformen auf. Die resultierenden Diagrammdaten werden in der ausgewählten Form dargestellt.

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Fehlerbalkendiagramme

Fehlerbalkendiagramme stellen die Variabilität von Daten dar und geben den Fehler (oder die Unsicherheit) in einer gemeldete Messung an. Mit Fehlerbalken kann ermittelt werden, ob Unterschiede statistisch signifikant sind. Fehlerbalken können außerdem die Anpassungsgüte für eine bestimmte Funktion angeben.

Erstellen eines einfachen Fehlerbalkendiagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Fehlerbalken**.



Der Erstellungsbereich wird aktualisiert, um eine Fehlerbalkendiagrammvorlage anzuzeigen.

2. Wählen Sie eine Variable als Variable für die **Y-Achse** aus. Die Daten dieser Variable werden auf der y -Achse dargestellt.
3. Wählen Sie eine metrische Variable als Variable für die **Kategorie** aus. Die Daten dieser Variable werden auf der x -Achse dargestellt.

Weitere Funktionen

Y-Achse

Listet Datasetvariablen auf, die für die y -Achse des Diagramms verfügbar sind.

Kategorie

Listet Datasetvariablen auf, die für die x -Achse des Diagramms verfügbar sind.

Kategorienreihenfolge

Wählen Sie die Reihenfolge aus, in der die Variablenkategorien sortiert werden.

Wie gelesen

Variablenkategorien werden so dargestellt, wie sie im Dataset vorkommen.

Aufsteigend

Sortiert Variablenkategorien in aufsteigender Reihenfolge.

Absteigend

Sortiert Variablenkategorien in absteigender Reihenfolge.

Aufteilen nach

Wählen Sie eine kategoriale Variable aus, die eine Tabelle mit Diagrammen erstellt, die für jede Kategorie in der Variablen "Aufteilen nach" eine Zelle enthält. Wie bei der Gruppierung werden dem Diagramm bei der Aufteilung nach Variablen vor allem weitere Dimensionen hinzugefügt, indem Informationen für jede Variablenkategorie angezeigt werden. Weitere Informationen finden Sie in Variablen "Aufteilen nach".

Bezugslinie

Zeigt bei Aktivierung eine Bezugslinie im Diagramm an. Die Referenzlinie korreliert mit der ausgewählten statistischen Methode.

Fehlerbalken

Bei Aktivierung werden die Linien, die den Fehlerbereich darstellen, im Diagramm angezeigt.

Maß Wählen Sie den Messtyp aus, der durch die Fehlerbalken dargestellt wird:

Konfidenzintervalle

Legt die Konfidenzintervalle für die ausgewählten Variablen fest. Der Standardwert ist 0,95 (95 %), wie im Feld **Darstellungswert** widergespiegelt.

Standardfehler

Misst den Standardfehler der ausgewählten Variablen.

Standardabweichung

Misst die Standardabweichungen der ausgewählten Variablen.

Konfidenzniveau

Dieser Wert stellt die Konfidenzintervalle für das ausgewählte **Maß** dar. Der Standardwert ist 0,95 (95 %).

Statistische Methode

Wählen Sie die Methode zur Beschreibung der zentralen Tendenz aus:

Mittelwert

Das Ergebnis aus der Summierung aller Quotienten und der anschließenden Division des Ergebnisses durch die Gesamtanzahl der Quotienten.

Median

Der Wert, der dadurch bestimmt ist, dass mindestens die Hälfte der Quotienten kleiner gleich diesem Wert und mindestens die Hälfte größer gleich diesem Wert sind.

Anzeigemodus

Wählen Sie aus, wie die Auswahl **Statistische Methode** angezeigt wird (Balken, Linie oder Kreis).

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

3-D-Diagramme

Mit 3-D-Diagrammen werden gewöhnlich Funktionen mit mehreren Variablen dargestellt. Sie enthalten eine Variable für die z-Achse, die eine Funktion der Variablen für die x- und y-Achsen ist.

Erstellen eines einfachen 3-D-Diagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **3D**.



Der Erstellungsbereich wird aktualisiert, um eine 3-D-Diagrammvorlage anzuzeigen.

2. Wählen Sie den **Typ** in der Dropdown-Liste aus.
3. Wählen Sie eine Variable für die **x-Achse** in der Dropdown-Liste aus.
4. Wählen Sie eine Variable für die **y-Achse** in der Dropdown-Liste aus.
5. Wählen Sie eine Variable für die **z-Achse** in der Dropdown-Liste aus.

Weitere Funktionen

Typ Listet die Diagrammtypen auf, die für die Darstellung der Daten verfügbar sind.

X-Achse

Listet Datasetvariablen auf, die für die x -Achse des Diagramms verfügbar sind.

Y-Achse

Listet Datasetvariablen auf, die für die y -Achse des Diagramms verfügbar sind.

Z-Achse

Listet Datasetvariablen auf, die für die z -Achse des Diagramms verfügbar sind.

QuickInfo-Info

Listet die Variablen auf, die zum Generieren von QuickInfos verwendet werden können, wenn Sie den Mauszeiger auf einen Datenpunkt bewegen.

Farbzuordnungstabelle

Listet die verfügbaren Variablen der Farbzuordnungstabelle auf. Diese Variablen verwenden den Farbverlauf basierend auf dem Wertebereich in der angegebenen Spalte zur Darstellung in den Diagrammpunkten. Farbzuordnungstabellen sind auch als Choroplethenkarten bekannt.

Größenzuordnungstabelle

Listet die verfügbaren Variablen der Größenzuordnungstabelle auf. Diese Variablen verwenden unterschiedliche Größen für die Darstellung in den Zeichenpunkten.

Z-Verhältnis

Legt den Maßstab der Datenwerte der z -Achse relativ zu den x - und y -Achsen fest.

Drehen

Aktiviert und inaktiviert Diagrammrotation.

QuickInfo für Datenpunkte

Steuert, wo die QuickInfo für Datenpunkte angezeigt werden (rechts neben Datenpunkten, in der rechten oberen Ecke des Diagramms oder ausgeblendet).

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Streudiagrammmatrixdiagramme

Mit Streudiagrammmatrizen können Sie mühelos ermitteln, ob zwischen mehreren Variablen lineare Korrelationen vorhanden sind.

Erstellen eines Streudiagrammmatrixdiagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Streudiagrammmatrix**.

Scatterplot matrix



Der Erstellungsbereich wird aktualisiert, um eine Vorlage für ein Streudiagrammmatrixdiagramm anzuzeigen.

2. Wählen Sie mehrere metrische Spaltenvariablen aus.

Jede ausgewählte Variable wird anhand jeder anderen Variablen geplottet, um eine Matrix aus einzelnen Streudiagrammen zu bilden.

Weitere Funktionen

Korrelation.

Bei Aktivierung werden für die ausgewählten Variablen lineare Korrelationsinformationen ("Stark", "Mittel", "Schwach") angezeigt.

Spalten

Wählen Sie mindestens zwei Matrixvariablen aus. Die Variablen müssen numerisch sein (aber kein Datumsformat).

Klicken Sie auf **Weitere Spalte hinzufügen**, um weitere Spalten hinzuzufügen.

Farbzuordnungstabelle

Listet die verfügbaren Variablen der Farbzuordnungstabelle auf. Diese Variablen verwenden den Farbverlauf basierend auf dem Wertebereich in der angegebenen Spalte zur Darstellung in den Diagrammpunkten. Farbzuordnungstabellen sind auch als Choroplethenkarten bekannt.

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Candlestick-Diagramme

Candlestick-Diagramme sind ein Stil von Finanzdiagrammen, mit denen Preisverschiebungen eines Wertpapiers, eines Derivativs oder einer Währung beschrieben werden. Jedes Candlestick-Element zeigt in der Regel einen Tag an. In einem einmonatigen Diagramm können die 20 Handelstage als 20 Candlestick-Elemente angezeigt werden. Candlestick-Diagramme werden am häufigsten bei der Analyse von Aktienkurs- und Währungspreismustern verwendet und ähneln Boxplots.

Die Datasets, mit denen ein Candlestick-Diagramm erstellt wird, müssen Eröffnungswerte, hohe sowie niedrige Werte und Schlusswerte für jeden anzuzeigenden Zeitraum enthalten.

Erstellen eines einfachen Candlestick-Diagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Candlestick**.



Der Erstellungsbereich wird aktualisiert, um eine Candlestick-Diagrammvorlage anzuzeigen.

2. Wählen Sie eine Variable als Variable für die **X-Achse** aus.
3. Wählen Sie eine Variable als Variable für **Hoch** aus.
4. Wählen Sie eine Variable als Variable für **Niedrig** aus.

Weitere Funktionen

X-Achse

Listet Datasetvariablen auf, die für die x -Achse des Diagramms verfügbar sind.

Hoch Listet Datasetvariablen auf, die für die hohen Preisgestaltungswerte des Diagramms verfügbar sind.

Zusammenfassung für hohes Feld

Wählen Sie eine statistische Zusammenfassungsfunktion für die ausgewählte hohe Variable aus.

Niedrig

Listet Datasetvariablen auf, die für die niedrigen Preisgestaltungswerte des Diagramms verfügbar sind.

Zusammenfassung für niedriges Feld

Wählen Sie eine statistische Zusammenfassungsfunktion für die ausgewählte niedrige Variable aus.

Eröffnung

Listet Datasetvariablen auf, die für den Preisgestaltungseröffnungswert des Diagramms verfügbar sind.

Schluss

Listet Datasetvariablen auf, die für den Preisgestaltungsschlusswert des Diagramms verfügbar sind.

Volumen

Listet Datasetvariablen auf, die für die Volumenbalken des Diagramms verfügbar sind.

Candlestick

Schaltet zwischen dem Anzeigen der Diagramm Daten als Candlestick oder Linie um.

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

ThemeRiver-Diagramme

Ein ThemeRiver ist ein spezialisiertes Flussdiagramm, das Änderungen im Zeitverlauf anzeigt.

Erstellen eines einfachen ThemeRiver-Diagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **ThemeRiver**.

Theme River



- Der Erstellungsbereich wird aktualisiert, um eine ThemeRiver-Diagrammvorlage anzuzeigen.
2. Wählen Sie eine Variable für die **X-Achse** aus.
3. Wählen Sie eine **Kategorie** aus.

Einschränkung: Wenn ein Kategoriefeld mehr als 50 unterschiedliche Kategorien aufweist, werden nur die Top 50 der Maximalwerte für Kategorien als Ereignisse im Diagramm verwendet.

Weitere Funktionen

Sortierbasis

Abhängig vom ausgewählten Wert für die X-Achse können Sie möglicherweise angeben, ob die Kategoriereihenfolge auf dem Kategorienamen oder auf dem Kategorienwert basieren soll.

Kategoriereihenfolge der X-Achse

Abhängig vom ausgewählten Wert für die X-Achse können Sie möglicherweise angeben, ob die Kategoriereihenfolge aufsteigend oder absteigend sein soll. Sie kann auch auf den ausgelesenen Werten aus dem Dataset basieren.

Zusammenfassung

Wählen Sie eine statistische Zusammenfassungsfunktion aus. Mit dieser Methode werden die einzelnen Kategorien zusammengefasst.

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Zeitdiagramme

Zeitdiagramme stellen Datenpunkte in aufeinanderfolgenden zeitlichen Intervallen dar. Die von Ihnen dargestellte Zeitreihe muss numerische Werte enthalten. Es wird angenommen, dass diese über einen Zeitraum hinweg auftreten, in dem die Berichtszeiträume einheitlich sind. Zeitdiagramme bieten eine vorläufige Analyse der Merkmale von Zeitreihendaten basierend auf einer Basisstatistik und Tests. Dadurch gewinnen Sie vor der Modellierung wertvolle Erkenntnisse über Ihre Daten. Zeitdiagramme beinhalten Analysemethoden wie Zerlegung, Augmented Dickey-Fuller-Tests (ADF), Korrelationen (ACF/PACF) und Spektralanalyse.

Erstellen eines einfachen Zeitdiagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Zeitdiagramm**.

Time plot



Der Erstellungsbereich wird aktualisiert, um eine Zeitdiagrammvorlage anzuzeigen.

2. Wählen Sie aus der Dropdown-Liste **Werte** einen Wert für die Y-Achse aus.

Weitere Funktionen

Datum

Falls zutreffend, wählen Sie ein Datum aus der Dropdown-Liste aus. Jede Beobachtung wird im Grunde genommen durch dasselbe Zeitintervall getrennt. Wenn Sie eine Datumsvariable auswählen, wird eine Option für das Vornehmen einer erneuten Stichprobe angezeigt. Mit dieser Option können Sie Wertfelder in Übereinstimmung mit dem angegebenen Intervall aggregieren.

Algorithmus des Zeitdiagramms

Der Algorithmus des Zeitdiagramms, der für die Analyse der Zeitreihendaten verwendet werden soll:

- **Zerlegung.** Sie können eine Zeitreihe in drei Komponenten zerlegen ("Trendzyklus", "Saisonal" und "Unregelmäßig"). Die Zerlegung erfolgt additiv.
- **ADF-Test.** Bei Augmented Dickey-Fuller-Tests (ADF) wird die Nullhypothese getestet, dass eine Einheitswurzel in einer Zeitreihe vorhanden und die Zeitreihe nicht stationär ist. Wenn das Testergebnis die Nullhypothese nicht erfüllt, bedeutet dies, dass die Zeitreihe stationär ist oder als stationär mit einem Differenzmodell dargestellt werden kann.
- **ACF/PACF.** Korrelation der Zeitreihe.
- **Spektralanalyse.** Ein Analysetool in der Häufigkeitsdomäne. Die höchste Häufigkeit wird als Raute markiert.

Diagrammposition tauschen

Keht die Positionen der Diagramme im planimetrischen rechteckigen Koordinatensystem und im Polarkoordinatensystem um.

Extrempunkt anzeigen

Blendet die Extrempunkte in den Diagrammen ein oder aus. Basierend auf den Trendzykluskomponenten der Zerlegung der Zeitreihe wird untersucht, ob die Zeitreihe einen allgemeinen Trend aufweist oder bestimmte Extrempunkte zeigt, die die Richtung des Trendmusters ändern.

Ausreißer anzeigen

Blendet Ausreißer ein oder aus. Die Ausreißer der Zeitreihe werden über die Komponente "Unregelmäßig" der Zerlegung einer Zeitreihe analysiert.

Primärtitel

Der Diagrammtitel.

Untertitel

Der Diagrammuntertitel, der direkt unterhalb des Diagrammtitels angezeigt wird.

Fußnote

Die Diagrammfußnote, die unterhalb des Diagramms angezeigt wird.

Benutzerdefinierte Diagramme

Benutzerdefinierte Diagramme ermöglichen das Einfügen/Bearbeiten von JSON-Code zum Erstellen des gewünschten Diagramms.

Erstellen eines benutzerdefinierten Diagramms

1. Klicken Sie in der Diagrammerstellung im Abschnitt **Diagrammtypen** auf das Symbol **Benutzerdefiniert**.

Customized



2. Fügen Sie den JSON-Code mit den Diagrammspezifikationen in das angegebene Feld im Bereich "Details".
3. Klicken Sie auf auf die Option zum Generieren des Diagramms.

Dashboard

Sie können ein Dashboard-Layout erstellen, um mehrere Diagramme gleichzeitig anzuzeigen. Speichern Sie Layouts als Vorlagen und verschieben Sie Ihre gespeicherten Diagramme per Drag-and-Drop zu den Positionen in Ihrem Layout.

1. , nachdem Sie mit der rechten Maustaste auf einen Datenknoten geklickt und **Daten anzeigen** ausgewählt haben, im Abschnitt **Aktionen** auf das Steuerelement "Dashboard".



Abbildung 14. Steuerelement "Dashboard"

Das Dashboard wird angezeigt und stellt folgende Einstellungen bereit.

Vorlage

In diesem Abschnitt können Sie neue Layoutvorlagen erstellen oder aus Ihren vordefinierten bzw. aus Ihren gespeicherten Vorlagen eine Vorlage auswählen.

Layoutvorlage auswählen

Wählen Sie eine der verfügbaren Layoutvorlagen oder beginnen Sie mit einer neuen Vorlage.

Aktionen

Klicken Sie auf das Symbol **Layout bearbeiten**, um die ausgewählte Layoutvorlage zu bearbeiten. Sie können auch eine Dashboarddatei importieren. Wenn Sie fertig sind, speichern Sie Ihre Vorlage.

Klicken Sie auf **Bearbeitungsmodus für Layout verlassen**, wenn Sie fertig sind.

Inhalt In diesem Abschnitt können Sie Elemente per Drag-and-Drop in Ihr Dashboard-Layout ziehen und dort ablegen.

Gespeichertes Diagramm auswählen

Dieser Abschnitt enthält eine Liste Ihrer gespeicherten Diagramme. Verschieben Sie Diagramme per Drag-and-Drop zur gewünschten Position in Ihrem Dashboard-Layout.

Objekt auswählen

Sie können auch HTML-Text oder Grafiken in Ihr Dashboard-Layout ziehen.

Globale Darstellungsvorgaben

Sie können die Standardeinstellungen für Titel, Bereichsschieberegler, Rasterlinien und Mastracking überschreiben. Sie können außerdem eine andere Farbschemavorlage angeben.

1. , nachdem Sie mit der rechten Maustaste auf einen Datenknoten geklickt und **Daten anzeigen** ausgewählt haben, im Abschnitt **Aktionen** auf das Steuerelement "Globale Darstellungsvorgaben".



Abbildung 15. Steuerelement für globale Darstellungsvorgaben

Das Dialogfeld "Globale Darstellungsvorgaben" wird angezeigt und stellt die folgenden Einstellungen bereit.

Titel In diesem Abschnitt werden globale Diagrammtiteleinstellungen bereitgestellt.

Globale Titel

Aktiviert oder inaktiviert die globalen Titel für alle Diagramme.

Globaler Primärtitel

Aktiviert oder inaktiviert die Anzeige von globalen, primären Diagrammtiteln. Bei Aktivierung wird der hier eingegebene Diagrammtitel der höchsten Ebene auf alle Diagramme angewendet und überschreibt dadurch die jeweilige Einstellung **Primärtitel** jeden Diagramms.

Globaler Untertitel

Aktiviert oder inaktiviert die Anzeige von globalen Diagrammuntertiteln. Bei Aktivierung wird der hier eingegebene Diagrammuntertitel auf alle Diagramme angewendet und überschreibt dadurch die jeweilige Einstellung **Untertitel** jeden Diagramms.

Standardtitel

Aktiviert oder inaktiviert die Standardtitel für alle Diagramme.

Tools In diesem Abschnitt werden Optionen bereitgestellt, die das Diagrammverhalten steuern.

Bereichsschieberegler

Aktiviert oder inaktiviert den Bereichsschieberegler für jedes Diagramm. Bei Aktivierung können Sie das Volumen angezeigter Diagrammdaten über einen Bereichsschieberegler steuern, der unter jedem Diagramm bereitgestellt wird.

Rasterlinien

Steuert die Anzeige der Rasterlinien für die X-Achse (vertikal) und die Y-Achse (horizontal).

Mastracker

Bei Aktivierung wird die Mauscursorposition in Beziehung zu den Diagrammdaten überwacht und angezeigt, wenn die Maus über das Diagramm gezogen wird.

Toolbox

Aktiviert oder inaktiviert die Toolbox für jedes Diagramm. Je nach Diagrammtyp stellt die Toolbox am rechten Rand der Anzeige Tools wie "Zoom", "Als Bild speichern", "Wiederherstellen", "Daten auswählen" und "Auswahl löschen" bereit.

Motiv Wählen Sie eine Vorlage aus, um die Farben zu ändern, die in Diagrammen mit einer Gruppierungs- oder Stapelvariablen verwendet werden. Die in der ausgewählten Vorlagendatei definierten Elementattribute überschreiben die Standardvorlageneinstellungen für diese Elementattribute.

Benachrichtigung

Wählen Sie aus, ob Sie mit einer Warnung benachrichtigt werden wollen, wenn Sie den Diagrammtyp wechseln oder zur Startseite navigieren.

2. Klicken Sie auf **Anwenden**, um die Einstellungen zu speichern, oder klicken Sie auf **Abbrechen**, um die Änderungen zu ignorieren.

Kapitel 7. Arbeiten mit Ausgaben

Beim Ausführen einiger Streams sind die Ergebnisse im Viewer über die Registerkarte **Modell** oder die Registerkarte **Erweitert** der Modellnuggetknoten verfügbar. Im Viewer können Sie problemlos zwischen den verschiedenen Teilen der Ausgabe wechseln. Außerdem können Sie die Ausgaben bearbeiten und so Dokumente erstellen, die genau die gewünschten Ausgaben enthalten. Einige Diagrammausgaben verwenden den Viewer.

Der Viewer wird für die folgende Ausgabe in IBM SPSS Modeler verwendet:

- TCM-Modellnuggets
- STP-Modellnuggets
- TwoStep-AS-Clustermodellnuggets
- GSAR-Modellnuggets
- Kartenvisualisierungsdiagrammknoten

Viewer

Ergebnisse werden im Viewer angezeigt. Sie können den Viewer für folgende Vorgänge verwenden:

- Durchsuchen der Ergebnisse
- Ein- und Ausblenden von ausgewählten Tabellen und Diagrammen
- Ändern der Anzeigereihenfolge der Ergebnisse durch Verschieben ausgewählter Elemente
- Verschieben von Elementen zwischen dem Viewer und anderen Anwendungen

Der Viewer ist in zwei Fensterbereiche aufgeteilt:

- Der linke Fensterbereich enthält eine Gliederungsansicht des Inhalts.
- Der rechte Fensterbereich enthält Statistiktabelle, Diagramme und Textausgabe.

Sie können auf ein Element in der Gliederung klicken, um direkt zur zugehörigen Tabelle bzw. dem zugehörigen Diagramm zu wechseln. Wenn Sie die Breite des Gliederungsfensters ändern möchten, können Sie auf dessen rechten Rahmen klicken und ihn mit gedrückter Maustaste auf die gewünschte Breite ziehen.

Ein- und Ausblenden von Ergebnissen

Im Viewer können Sie ausgewählte Tabellen oder alle Ergebnisse einer Prozedur ein- und ausblenden. Dies ist nützlich, wenn Sie möchten, dass im Inhaltsfenster weniger angezeigt wird.

Ausblenden von Tabellen und Diagrammen

1. Doppelklicken Sie im Gliederungsfenster des Viewers auf das Buchsymbol des Elements.
ODER
2. Klicken Sie einmal auf das Element, um es auszuwählen.
3. Wählen Sie in den Menüs Folgendes aus:
Ansicht > Ausblenden
ODER
4. Klicken Sie in der Gliederungssymbolleiste auf die Schaltfläche "Ausblenden", die durch ein geschlossenes Buch dargestellt ist.

Die Schaltfläche "Einblenden", die durch ein geöffnetes Buch dargestellt ist, wird zu einer aktiven Schaltfläche. Dadurch wird angezeigt, dass das Element jetzt ausgeblendet ist.

Ausblenden der Ergebnisse einer Prozedur

1. Klicken Sie im Gliederungsfenster auf das Kästchen neben dem Namen der Prozedur.

Dadurch werden alle Ergebnisse der Prozedur ausgeblendet und die Gliederungsansicht reduziert.

Verschieben, Löschen und Kopieren von Ausgaben

Sie können die Ergebnisse neu anordnen, indem Sie einzelne Elemente oder ganze Elementgruppen kopieren, verschieben oder löschen.

Verschieben von Ausgaben im Viewer

1. Wählen Sie die Elemente im Gliederungs- oder Inhaltsfenster aus.
2. Ziehen Sie die Elemente auf einen anderen Speicherort und legen Sie sie dort ab.

So löschen Sie Ausgabe im Viewer

1. Wählen Sie die Elemente im Gliederungs- oder Inhaltsfenster aus.
2. Drücken Sie die Löschtaste.

ODER

3. Wählen Sie in den Menüs Folgendes aus:

Bearbeiten > Löschen

Ändern der anfänglichen Ausrichtung

In der Standardeinstellung sind alle Ergebnisse linksbündig ausgerichtet. So ändern Sie die ursprüngliche Ausrichtung neuer Ausgabeelemente:

1. Wählen Sie in den Menüs Folgendes aus:
Bearbeiten > Optionen
2. Klicken Sie auf die Registerkarte **Viewer**.
3. Wählen Sie in der Gruppe "Anfänglicher Ausgabestatus" den gewünschten Elementtyp aus (z. B. Pivot-Tabelle, Diagramm, Textausgabe).
4. Wählen Sie die gewünschte Ausrichtungsoption aus.

Ändern der Ausrichtung von Ausgabeelementen

1. Wählen Sie im Gliederungs- bzw. Inhaltsbereich die Elemente aus, die sie ausrichten möchten.
2. Wählen Sie in den Menüs Folgendes aus:

Format > Linksbündig

ODER

Format > Zentriert

ODER

Format > Rechtsbündig

Gliederung für Viewer

Im Gliederungsfenster wird ein Inhaltsverzeichnis des Viewer-Dokuments angezeigt. Hier können Sie die Ergebnisse durchblättern und festlegen, welche Objekte angezeigt werden. Die meisten Aktionen im Gliederungsfenster wirken sich auch auf das Inhaltsfenster aus.

- Beim Auswählen eines Elements im Gliederungsfenster wird das entsprechende Element im Inhaltsfenster angezeigt.
- Beim Verschieben eines Elements im Gliederungsfenster wird das entsprechende Element im Inhaltsfenster ebenfalls verschoben.
- Beim Reduzieren der Gliederungsansicht werden die Ergebnisse aller Elemente in den reduzierten Ebenen ausgeblendet.

Einstellen der Anzeige für die Gliederung. Zum Einstellen der Anzeige für die Gliederung stehen Ihnen die folgenden Möglichkeiten zur Verfügung:

- Erweitern und Reduzieren der Gliederungsansicht
- Ändern der Gliederungsebene von ausgewählten Elementen
- Ändern der Größe von Elementen im Gliederungsfenster
- Ändern der in der Gliederung verwendeten Schriftart

Reduzieren und Erweitern der Gliederungsansicht

1. Klicken Sie auf das Kästchen links neben dem Gliederungselement, das Sie reduzieren oder erweitern möchten.

ODER

2. Klicken Sie in der Gliederung auf das Element.
3. Wählen Sie in den Menüs Folgendes aus:

Ansicht > Reduzieren

ODER

Ansicht > Erweitern

Ändern Gliederungsebene

1. Klicken Sie im Gliederungsfenster auf das Element.
2. Wählen Sie in den Menüs Folgendes aus:

Bearbeiten > Gliederung > Heraufstufen

ODER

Bearbeiten > Gliederung > Herabstufen

Ändern der Größe von Elementen in der Gliederung

1. Wählen Sie in den Menüs Folgendes aus:
Ansicht > Größe der Gliederung
2. Wählen Sie die Größe der Gliederung (**Klein**, **Mittel** oder **Groß**).

Ändern der Schriftart in der Gliederung

1. Wählen Sie in den Menüs Folgendes aus:
Ansicht > Schriftart für Gliederung...
2. Wählen Sie eine Schriftart aus.

Hinzufügen von Elementen zum Viewer

Sie können dem Viewer Elemente hinzufügen, beispielsweise Titel, neue Texte, Diagramme oder Objekte aus anderen Anwendungen.

Einfügen eines neuen Titels oder eines neuen Textobjekts

Sie können dem Viewer Textelemente hinzufügen, die nicht mit einer Tabelle oder einem Diagramm verbunden sind.

1. Klicken Sie auf die Tabelle, das Diagramm oder das Objekt, nach der bzw. dem der Titel oder Text eingefügt werden soll.
2. Wählen Sie in den Menüs Folgendes aus:
Einfügen > Neuer Titel
ODER
Einfügen > Neuer Text
3. Doppelklicken Sie auf das neue Objekt.
4. Geben Sie den gewünschten Text ein.

Hinzufügen einer Textdatei

1. Klicken Sie im Gliederungs- oder Inhaltsfenster des Viewers auf die Tabelle, das Diagramm oder das Objekt, nach der bzw. dem der Text eingefügt werden soll.
2. Wählen Sie in den Menüs Folgendes aus:
Einfügen > Textdatei...
3. Wählen Sie eine Textdatei aus.

Doppelklicken Sie auf die Textdatei, um diese zu bearbeiten.

Einfügen von Objekten in den Viewer

Objekte aus anderen Anwendungen können in den Viewer eingefügt werden. Verwenden Sie hierfür den Befehl **Einfügen nach** oder **Inhalte einfügen**. Dabei wird das neue Objekt nach dem aktuell ausgewählten Objekt im Viewer eingefügt. Verwenden Sie den Befehl **Inhalte einfügen**, wenn Sie das Format für das einzufügende Objekt auswählen möchten.

Suchen und Ersetzen von Informationen im Viewer

1. Um Informationen im Viewer zu suchen bzw. zu ersetzen, wählen Sie in den Menüs Folgendes aus:
Bearbeiten > Suchen
ODER
Bearbeiten > Ersetzen

Sie können "Suchen und Ersetzen" für folgende Vorgänge verwenden:

- Durchsuchen des gesamten Dokuments bzw. der ausgewählten Elemente.
- Suchen nach unten bzw. oben von der aktuellen Position.
- Durchsuchen beider Fensterbereiche oder Einschränkung der Suche auf den Inhalt des Gliederungsfensters.
- Suchen nach ausgeblendeten Elementen. Dazu gehören alle Elemente, die im Inhaltsfenster ausgeblendet sind (z. B. Anmerkungstabellen, die standardmäßig ausgeblendet sind) sowie ausgeblendete Zeilen und Spalten in Pivot-Tabellen.
- Einschränken der Suchkriterien auf Treffer, die auch hinsichtlich der Groß- und Kleinschreibung übereinstimmen.
- Einschränken der Suchkriterien in Pivot-Tabellen auf Treffer, bei denen der gesamte Zelleninhalt übereinstimmt.
- Einschränken der Suchkriterien in Pivot-Tabellen auf Fußnotenmarkierungen. Diese Option steht nicht zur Verfügung, wenn die Auswahl im Viewer andere Elemente als Pivot-Tabellen beinhaltet.

Ausgeblendete Elemente und Schichten von Pivot-Tabellen

- Schichten unterhalb der sichtbaren Schicht einer mehrdimensionalen Pivot-Tabelle gelten nicht als ausgeblendet und sind auch dann in den Suchbereich eingeschlossen, wenn ausgeblendete Elemente nicht in die Suche mit aufgenommen wurden.
- Zu den ausgeblendeten Elementen gehören ausgeblendete Elemente im Inhaltsfenster (Elemente, die im Gliederungsfenster mit einem geschlossenen Buch dargestellt werden oder in reduzierten Blöcken des Gliederungsfensters enthalten sind) und Zeilen und Spalten in Pivot-Tabellen, die entweder standardmäßig ausgeblendet sind (z. B. sind leere Zeilen und Spalten standardmäßig ausgeblendet) oder manuell (durch Bearbeiten der Tabelle und selektives Ausblenden bestimmter Zeilen bzw. Spalten) ausgeblendet sind. Ausgeblendete Elemente werden nur dann in die Suche mit aufgenommen, wenn Sie explizit die Option **Ausgeblendete Elemente einschließen** auswählen.
- In beiden Fällen wird das ausgeblendete bzw. unsichtbare Element, das den Suchtext bzw. -wert enthält beim Auffinden angezeigt, das Element wird jedoch anschließend wieder in den ursprünglichen Zustand zurückversetzt.

Suchen nach einem Wertebereich in Pivot-Tabellen

So finden Sie in Pivot-Tabellen Werte, die in einen angegebenen Wertebereich fallen:

1. Aktivieren Sie eine Pivot-Tabelle oder wählen Sie eine oder mehrere Pivot-Tabellen im Viewer aus. Stellen Sie sicher, dass ausschließlich Pivot-Tabellen ausgewählt sind. Wenn auch andere Optionen ausgewählt sind, steht die Bereichsoption nicht zur Verfügung.
2. Wählen Sie in den Menüs Folgendes aus:
Bearbeiten > Suchen
3. Klicken Sie auf die Registerkarte **Bereich**.
4. Wählen Sie den Typ des Bereichs aus: "Zwischen", "Größer-gleich", "Kleiner-gleich".
5. Wählen Sie den Wert bzw. die Werte, die den Bereich definieren.
 - Wenn ein Wert nicht numerische Zeichen enthält, werden beide Werte als Zeichenfolgen behandelt.
 - Wenn beide Werte Zahlen sind, werden nur numerische Werte durchsucht.
 - Mit der Registerkarte "Bereich" können keine Werte ersetzt werden.

Diese Funktion steht für aus früheren Versionen übernommene Tabellen nicht zur Verfügung. Weitere Informationen finden Sie in „Aus früheren Versionen übernommene Tabellen“ auf Seite 130.

Kopieren von Ausgaben in andere Anwendungen

Ausgabeobjekte können kopiert und in andere Anwendungen, beispielsweise in Textverarbeitungs- oder Tabellenkalkulationsprogramme eingefügt werden. Die Ausgaben können in verschiedenen Formaten eingefügt werden. Je nach Zielanwendung und ausgewählten Ausgabeobjekten stehen einige oder alle der folgenden Formate zur Verfügung:

Metadatei. Metadateiformate WMF und EMF. Diese Formate sind nur unter Windows-Betriebssystemen verfügbar.

RTF (Rich Text Format). Mehrere ausgewählte Objekte, Textausgaben und Pivot-Tabellen können im RTF-Format kopiert und eingefügt werden. Für Pivot-Tabellen bedeutet dies bei den meisten Anwendungen, dass die Tabellen als Tabellen eingefügt werden, die dann in der anderen Anwendung bearbeitet werden können. Pivot-Tabellen, die zu breit für das Dokument sind, werden entweder umgebrochen, auf die Dokumentbreite verkleinert oder unverändert beibehalten, je nach den Optionseinstellungen für Pivot-Tabellen. Weitere Informationen finden Sie in „Optionen: Pivot-Tabellen“ auf Seite 131.

Anmerkung: Besonders breite Tabellen werden von Microsoft Word möglicherweise nicht ordnungsgemäß angezeigt.

Bild. Bildformate JPG und PNG.

BIFF. Pivot-Tabellen und Textausgaben können im BIFF-Format in eine Tabellenkalkulation eingefügt werden. Zahlen in Pivot-Tabellen behalten ihre numerische Genauigkeit bei. Dieses Format ist nur unter Windows-Betriebssystemen verfügbar.

Text. Pivot-Tabellen und Textausgaben können als Text kopiert und eingefügt werden. Dies kann bei Anwendungen wie E-Mail-Programmen nützlich sein, bei denen mit der Anwendung nur Text verarbeitet oder übertragen werden kann.

Microsoft Office-Grafikobjekt. Diagramme, die dieses Format unterstützen, können in Microsoft Office-Anwendungen kopiert und dort wie native Microsoft Office-Diagramme bearbeitet werden. Aufgrund von Unterschieden zwischen SPSS Statistics- bzw. SPSS Modeler-Diagrammen und Microsoft Office-Diagrammen werden einige Funktionen der SPSS Statistics- bzw. SPSS Modeler-Diagramme in der kopierten Version nicht beibehalten. Das Kopieren mehrerer ausgewählter Diagramme im Format "Microsoft Office-Grafikobjekt" wird nicht unterstützt.

Wenn die Zielanwendung mehrere verfügbare Formate unterstützt, verfügt sie möglicherweise über eine Menüoption vom Typ "Inhalte einfügen", mit der Sie das Format auswählen können, oder es wird automatisch eine Liste der verfügbaren Formate angezeigt.

Anmerkung: Microsoft Office Version 16 (oder höher) ist beim Kopieren und Einfügen von Boxplots und Histogrammen erforderlich.

Kopieren und Einfügen mehrerer Ausgabeobjekte

Beim Einfügen mehrerer Ausgabeobjekte in andere Anwendungen gelten folgende Einschränkungen:

- **RTF-Format.** Bei den meisten Anwendungen werden Pivot-Tabellen als Tabellen eingefügt, die in der betreffenden Anwendung bearbeitet werden können. Diagramme, Baumdiagramme und Modellansichten werden als Bilder eingefügt.
- **Metadatei- und Bildformate.** Alle ausgewählten Ausgabeobjekte werden als Einzelobjekt in die andere Anwendung eingefügt.
- **BIFF-Format.** Diagramme, Baumdiagramme und Modellansichten sind ausgeschlossen.

Inhalte kopieren

Beim Kopieren und Einfügen großer Mengen an Ausgaben, insbesondere sehr großer Pivot-Tabellen, können Sie die Geschwindigkeit des Vorgangs erhöhen, indem Sie die Anzahl der in die Zwischenablage kopierten Formate mit **Bearbeiten > Inhalte kopieren** begrenzen.

Sie können die ausgewählten Formate auch als Menge an Formaten speichern, die standardmäßig in die Zwischenablage kopiert werden sollen. Diese Einstellung wird über das Sitzungsende hinweg beibehalten.

Kopieren als

Sie können im Ausgabebrowser mit der rechten Maustaste auf ein ausgewähltes Objekt klicken und **Bearbeiten > Kopieren als** auswählen, um eine Kopie in einem der gängigen Formate zu erstellen (z. B. **Alle**, **Bild** oder **Microsoft Office-Grafikobjekt**). Bei Auswahl von **Bearbeiten > Kopieren** wird Alles kopiert. Wenn **Kopieren als** abgeblendet oder für ein ausgewähltes Objekt nicht verfügbar ist, ist dieses Kopierformat für das jeweilige Objekt nicht verfügbar.

Interaktive Ausgabe

Interaktive Ausgabeobjekte enthalten mehrere zusammengehörige Ausgabeobjekte. Durch eine Auswahl in einem Objekt kann sich die Anzeige oder Hervorhebung in einem anderen Objekt ändern. Wenn Sie z. B. eine Zeile in einer Tabelle auswählen, wird ein Bereich in einer Karte hervorgehoben oder ein Diagramm für eine andere Kategorie angezeigt.

Interaktive Ausgabeobjekte unterstützen keine Bearbeitungsfunktionen wie das Ändern von Text, Farben, Schriftarten oder Tabellenumrandungen. Die einzelnen Objekte können aus dem interaktiven Objekt in den Viewer kopiert werden. Aus interaktiver Ausgabe kopierte Tabellen können im Pivot-Tabellen-Editor bearbeitet werden.

Kopieren von Objekten aus interaktiver Ausgabe

Datei>In Viewer kopieren kopiert einzelne Ausgabeobjekte in das Fenster **Viewer**.

- Die verfügbaren Optionen hängen vom Inhalt der interaktiven Ausgabe ab.
- **Diagramm** und **Karte** erstellen Diagrammobjekte.
- **Tabelle** erstellt eine Pivot-Tabelle, die im Pivot-Tabellen-Editor bearbeitet werden kann.
- **Momentaufnahme** erstellt ein Bild der aktuellen Ansicht.

- **Modell** erstellt eine Kopie des aktuellen interaktiven Ausgabeobjekts.

Bearbeiten > Objekt kopieren kopiert einzelne Ausgabeobjekte in die Zwischenablage.

- Das Einfügen eines kopierten Objekts in den Viewer entspricht der Auswahl von **Datei > In Viewer kopieren**.
- Beim Einfügen des Objekts in eine andere Anwendung wird das Objekt als Bild eingefügt.

Zoomen und Schwenken

Bei Abbildungen können Sie mit **Ansicht > Zoom** die Ansicht zoomen. In einer gezoomten Ansicht können Sie diese mit **Ansicht > Schwenken** verschieben.

Druckeinstellungen

Datei > Druckeinstellungen steuert, wie interaktive Objekte gedruckt werden.

- **Nur sichtbare Ansicht drucken.** Druckt nur die derzeit dargestellte Ansicht. Diese Option ist die Standardeinstellung.
- **Alle Ansichten drucken.** Druckt alle in der interaktiven Ausgabe enthaltenen Ansichten.
- Die ausgewählte Option bestimmt auch die Standardaktion für den Export des Ausgabeobjekts.

Ausgabe exportieren

Mit "Ausgabe exportieren" wird die Viewer-Ausgabe im HTML-, Text-, Word/RTF-, Excel-, PowerPoint- (PowerPoint 97 oder höher erforderlich) oder PDF-Format gespeichert. Außerdem ist der Export von Diagrammen in mehreren verschiedenen Grafikformaten möglich.

Anmerkung: Der Export nach PowerPoint ist nur unter Windows-Betriebssystemen verfügbar.

Exportieren von Ausgaben

1. Machen Sie den Viewer zum aktiven Dialogfeld (klicken Sie auf das Dialogfeld).
2. Klicken Sie auf die Schaltfläche **Exportieren** in der Symbolleiste oder klicken Sie mit der rechten Maustaste in das Ausgabefenster und wählen Sie **Exportieren** aus.
3. Geben Sie einen Dateinamen (bzw. ein Präfix bei Diagrammen) ein und wählen Sie ein Exportformat aus.

Zu exportierende Objekte. Sie können alle Objekte im Viewer, alle sichtbaren Objekte oder nur ausgewählte Objekte exportieren.

Dokumenttyp. Die folgenden Optionen sind verfügbar:

- **Word/RTF (*.doc).** Pivot-Tabellen werden mit sämtlichen Formatierungsattributen (wie Zellenrahmen, Schriftarten und Hintergrundfarben) als Word-Tabellen exportiert. Textausgaben werden als formatierter RTF-Text exportiert. Diagramme, Baumdiagramme und Modellansichten werden im PNG-Format aufgenommen. Besonders breite Tabellen werden von Microsoft Word möglicherweise nicht ordnungsgemäß angezeigt.
- **Excel 97-2004 (*.xls)/Excel 2007 und höher (*.xlsx).** Die Zeilen, Spalten und Zellen von Pivot-Tabellen werden mit sämtlichen Formatierungsattributen wie Zellenrahmen, Schriftarten und Hintergrundfarben als Excel-Zeilen, -Spalten und -Zellen exportiert. Textausgaben werden mit allen Schriftartattributen exportiert. Jede Zeile in der Textausgabe entspricht einer Zeile in der Excel-Datei, wobei der gesamte Inhalt der Zeile in einer einzelnen Zelle enthalten ist. Diagramme, Baumdiagramme und Modellansichten werden im PNG-Format aufgenommen. Die Ausgaben können als *Excel 97-2004* oder als *Excel 2007 und höher* exportiert werden.
- **HTML (*.htm).** Pivot-Tabellen werden als HTML-Tabellen exportiert. Textausgaben werden als vorformatierter HTML-Text exportiert. Diagramme, Baumdiagramme und Modellansichten werden im ausge-

wählten Grafikformat in das Dokument eingebettet. Zum Anzeigen von im HTML-Format exportierten Ausgabedaten ist ein Browser erforderlich, der mit HTML 5 kompatibel ist.

- **Portable Document Format (*.pdf).** Alle Ausgaben werden so exportiert, wie sie in der Druckvorschau/Seitenansicht angezeigt werden. Alle Formatierungsattribute bleiben erhalten.
- **Text - Einfach/UTF8/UTF16 (*.txt).** Zu den Textausgabeformaten gehören einfacher Text, UTF-8 und UTF-16. Pivot-Tabellen können als durch Tabulatoren getrennter Text oder als durch Leerzeichen getrennter Text exportiert werden. Alle Textausgaben werden in durch Leerzeichen getrenntem Format exportiert. Bei Diagrammen, Baumdiagrammen und Modellansichten wird für jede Grafik eine Zeile in die Textdatei eingefügt, in der der Name der Bilddatei angegeben ist.
- **Ohne (nur Grafiken).** Zu den verfügbaren Exportformaten gehören EPS, JPEG, TIFF, PNG und BMP. Unter Windows-Betriebssystemen ist außerdem das Format EMF (Enhanced Metafile, erweiterte Metafile) verfügbar.

Übergeordneten Ordner öffnen. Öffnet den Ordner, der die durch den Export erstellten Dateien enthält.

HTML-Optionen

Für den HTML-Export ist ein Browser erforderlich, der mit HTML 5 kompatibel ist.

Folgende Optionen stehen für den Export von Ausgaben im HTML-Format zur Verfügung:

Schichten in Pivot-Tabellen. Standardmäßig richtet sich die Aufnahme bzw. der Ausschluss von Pivot-Tabellenschichten nach den Tabelleneigenschaften der einzelnen Pivot-Tabellen. Sie können diese Einstellung außer Kraft setzen und alle Schichten aufnehmen oder alle Schichten mit Ausnahme der aktuell sichtbaren Schicht ausschließen. Weitere Informationen finden Sie in „Tabelleneigenschaften: Drucken“ auf Seite 125.

Tabellen mit Schichten als interaktive Tabellen exportieren. Tabellen mit Schichten werden so wie sie im Viewer dargestellt werden angezeigt und Sie können die angezeigte Schicht im Browser interaktiv ändern. Wenn diese Option nicht ausgewählt ist, wird jede Tabellenschicht als separate Tabelle angezeigt.

Tabellen als HTML. Damit werden die bei exportierten Pivot-Tabellen enthaltenen Stilinformationen festgelegt.

- **Export mit Stilen und fester Spaltenbreite.** Alle Stilinformationen (Schriftschnitte, Hintergrundfarben usw.) und Spaltenbreiten der Pivot-Tabellen werden beibehalten.
- **Ohne Stile exportieren.** Pivot-Tabellen werden in HTML-Standardtabellen konvertiert. Es werden keinerlei Stilattribute beibehalten. Die Spaltenbreite wird automatisch festgelegt.

Fußnoten und Titelzeilen aufnehmen. Dient zur Festlegung der Aufnahme bzw. des Ausschlusses aller Fußnoten und Titelzeilen von Pivot-Tabellen.

Modellansichten. Die Aufnahme und der Ausschluss von Modellansichten werden standardmäßig von den Modelleigenschaften der einzelnen Modelle gesteuert. Sie können diese Einstellung außer Kraft setzen und alle Ansichten aufnehmen oder alle Ansichten mit Ausnahme der aktuell angezeigten Ansicht ausschließen. Weitere Informationen finden Sie in Modelleigenschaften. (Hinweis: Alle Modelle einschließlich Tabellen werden als Grafik exportiert.)

Hinweis: Bei HTML können Sie außerdem das Bilddateiformat für exportierte Diagramme festlegen. Weitere Informationen finden Sie in „Grafikformatoptionen“ auf Seite 113.

So legen Sie HTML-Exportoptionen fest:

1. Wählen Sie **HTML** als Exportformat aus.
2. Klicken Sie auf **Optionen ändern**.

Webberichtsoptionen

Ein Webbericht ist ein interaktives Dokument, das mit den meisten Browsern kompatibel ist. Viele der interaktiven Funktionen der im Viewer verfügbaren Pivot-Tabellen stehen in Webberichten ebenfalls zur Verfügung.

Berichtstitel. Der Titel, der in der Kopfzeile des Berichts angezeigt wird. Standardmäßig wird der Dateiname verwendet. Sie können einen benutzerdefinierten Titel angeben, der anstelle des Dateinamens verwendet werden soll.

Format. Für das Berichtsformat stehen zwei Optionen zur Verfügung:

- **SPSS Web Report (HTML 5).** Für dieses Format ist ein Browser erforderlich, der mit HTML 5 kompatibel ist.
- **Cognos Active Report (mht).** Für dieses Format ist ein Browser erforderlich, der MHT-Formatdateien oder die Anwendung "Cognos Active Report" unterstützt.

Objekte ausschließen. Sie können ausgewählte Objekttypen aus dem Bericht ausschließen:

- **Text.** Textobjekte, bei denen es sich nicht um Protokolle handelt. Diese Option umfasst Textobjekte mit Informationen zum aktiven Dataset.
- **Protokolle.** Textobjekte, die eine Liste der ausgeführten Befehlssyntax enthalten. Protokollelemente umfassen auch Warnungen und Fehlernachrichten, die für Befehle auftreten, die keine Viewer-Ausgabe erzeugen.
- **Hinweistabellen.** Die Ausgabe aus statistischen und Diagrammprozeduren umfasst eine Hinweistabelle. Diese Tabelle enthält Informationen zum verwendeten Dataset, zu fehlenden Werten und zur Befehlssyntax, die zum Ausführen der Prozedur verwendet wurde.
- **Warnungen und Fehlernachrichten.** Warnungen und Fehlernachrichten aus Prozeduren für Statistiken und Diagramme.

Tabellen und Diagramm an Stil des Webberichts anpassen. Diese Option wendet den Standardstil von Webberichten auf alle Tabellen und Diagramme an. Diese Einstellung setzt Schriften, Farben oder andere Stile in der Ausgabe des Viewers außer Kraft. Der Standardstil von Webberichten kann nicht geändert werden.

Web-Server-Verbindung. Sie können die URL-Position mindestens eines Anwendungsservers einschließen, der den IBM SPSS Statistics Web Report Application Server ausführt. Der Webanwendungsserver stellt Funktionen für Pivot-Tabellen, zur Bearbeitung von Diagrammen und zum Speichern geänderter Webberichte bereit.

- Wählen Sie **Verwenden** für jeden Anwendungsserver aus, den Sie in den Webbericht einschließen wollen.
- Wenn ein Webbericht eine URL-Angabe enthält, stellt der Webbericht eine Verbindung zum entsprechenden Anwendungsserver her, um die zusätzlichen Bearbeitungsfunktionen bereitzustellen.
- Wenn Sie mehrere URLs angeben, versucht der Webbericht, in der angegebenen Reihenfolge eine Verbindung zu jedem der Server herzustellen.

Der IBM SPSS Statistics Web Report Application Server kann von <http://www.ibm.com/developerworks/spssdevcentral> heruntergeladen werden.

Word/RTF-Optionen

Folgende Optionen stehen für den Export von Ausgaben im Word-Format zur Verfügung:

Schichten in Pivot-Tabellen. Standardmäßig richtet sich die Aufnahme bzw. der Ausschluss von Pivot-Tabellenschichten nach den Tabelleneigenschaften der einzelnen Pivot-Tabellen. Sie können diese Einstel-

lung außer Kraft setzen und alle Schichten aufnehmen oder alle Schichten mit Ausnahme der aktuell sichtbaren Schicht ausschließen. Weitere Informationen finden Sie in „Tabelleneigenschaften: Drucken“ auf Seite 125.

Breite Pivot-Tabellen. Legt fest, wie Tabellen behandelt werden, die für die definierte Dokumentbreite zu breit sind. Standardmäßig wird die Tabelle umgebrochen. Die Tabelle wird in zwei Abschnitte unterteilt und die Zeilenbeschriftungen werden in jedem Abschnitt der Tabelle wiederholt. Alternativ können Sie breite Tabellen auch verkleinern oder keine Änderungen an breiten Tabellen vornehmen und damit zulassen, dass sie über den Rand der definierten Dokumentbreite hinausragen.

Haltepunkte beibehalten. Wenn Sie Haltepunkte definiert haben, werden diese Einstellungen in den Word-Tabellen beibehalten.

Fußnoten und Titelzeilen aufnehmen. Dient zur Festlegung der Aufnahme bzw. des Ausschlusses aller Fußnoten und Titelzeilen von Pivot-Tabellen.

Modellansichten. Die Aufnahme und der Ausschluss von Modellansichten werden standardmäßig von den Modelleigenschaften der einzelnen Modelle gesteuert. Sie können diese Einstellung außer Kraft setzen und alle Ansichten aufnehmen oder alle Ansichten mit Ausnahme der aktuell angezeigten Ansicht ausschließen. Weitere Informationen finden Sie in Modelleigenschaften. (Hinweis: Alle Modelle einschließlich Tabellen werden als Grafik exportiert.)

Druckoptionen für Export. Über diese Option wird ein Dialogfeld geöffnet, in dem Sie die Seitengröße und die Seitenränder für das exportierte Dokument festlegen können. Die Dokumentbreite, die festlegt, ob Tabellen umgebrochen oder verkleinert werden, ergeben sich aus der Seitenbreite minus den linken und rechten Seitenrand.

So legen Sie Exportoptionen für Word-Dokumente fest:

1. Wählen Sie **Word/RTF** als Exportformat aus.
2. Klicken Sie auf **Optionen ändern**.

Excel-Optionen

Folgende Optionen stehen für den Export von Ausgaben im Excel-Format zur Verfügung:

Erstellen Sie ein Arbeitsblatt oder eine Arbeitsmappe oder bearbeiten Sie ein vorhandenes Arbeitsblatt. Standardmäßig wird eine neue Arbeitsmappe erstellt. Wenn bereits eine Datei mit dem angegebenen Namen vorhanden ist, wird sie überschrieben. Wenn Sie die Option für die Erstellung eines Arbeitsblatts wählen und in der angegebenen Datei bereits ein Arbeitsblatt mit dem angegebenen Namen vorhanden ist, wird es überschrieben. Wenn Sie die Option zum Bearbeiten eines bestehenden Arbeitsblatts auswählen, müssen Sie auch den Arbeitsblattnamen angeben. (Beim Erstellen eines Arbeitsblatts optional.) Die Namen von Arbeitsblättern dürfen höchstens 31 Zeichen lang sein und dürfen keine Schrägstriche, umgekehrten Schrägstriche, eckigen Klammern, Fragezeichen oder Sterne enthalten.

Wenn Sie beim Export in Excel 97–2004 ein vorhandenes Arbeitsblatt bearbeiten, werden Diagramme, Modellansichten und Baumdiagramme nicht in die exportierte Ausgabe aufgenommen.

Stelle in Arbeitsblatt. Legt die Position auf dem Arbeitsblatt für die exportierte Ausgabe fest. Standardmäßig wird die exportierte Ausgabe nach der letzten Spalte mit Inhalten beginnend mit der ersten Zeile eingefügt, ohne vorhandene Inhalte zu ändern. Diese Einstellung ist sinnvoll, wenn neue Spalten in ein vorhandenes Arbeitsblatt eingefügt werden. Exportierte Ausgaben nach der letzten Zeile einzufügen, ist sinnvoll, wenn neue Zeilen in ein vorhandenes Arbeitsblatt eingefügt werden sollen. Wenn Sie exportierte Ausgaben ab einer bestimmten Zellenposition einfügen, werden vorhandene Inhalte in dem Bereich überschrieben, in den die exportierten Ausgaben eingefügt werden.

Schichten in Pivot-Tabellen. Standardmäßig richtet sich die Aufnahme bzw. der Ausschluss von Pivot-Tabellenschichten nach den Tabelleneigenschaften der einzelnen Pivot-Tabellen. Sie können diese Einstellung außer Kraft setzen und alle Schichten aufnehmen oder alle Schichten mit Ausnahme der aktuell sichtbaren Schicht ausschließen. Weitere Informationen finden Sie in „Tabelleneigenschaften: Drucken“ auf Seite 125.

Fußnoten und Titelzeilen aufnehmen. Dient zur Festlegung der Aufnahme bzw. des Ausschlusses aller Fußnoten und Titelzeilen von Pivot-Tabellen.

Modellansichten. Die Aufnahme und der Ausschluss von Modellansichten werden standardmäßig von den Modelleigenschaften der einzelnen Modelle gesteuert. Sie können diese Einstellung außer Kraft setzen und alle Ansichten aufnehmen oder alle Ansichten mit Ausnahme der aktuell angezeigten Ansicht ausschließen. Weitere Informationen finden Sie in Modelleigenschaften. (Hinweis: Alle Modelle einschließlich Tabellen werden als Grafik exportiert.)

So legen Sie Exportoptionen für Excel-Dokumente fest:

1. Wählen Sie **Excel** als Exportformat aus.
2. Klicken Sie auf **Optionen ändern**.

PowerPoint-Optionen

Die folgenden Optionen sind für PowerPoint verfügbar:

Schichten in Pivot-Tabellen. Standardmäßig richtet sich die Aufnahme bzw. der Ausschluss von Pivot-Tabellenschichten nach den Tabelleneigenschaften der einzelnen Pivot-Tabellen. Sie können diese Einstellung außer Kraft setzen und alle Schichten aufnehmen oder alle Schichten mit Ausnahme der aktuell sichtbaren Schicht ausschließen. Weitere Informationen finden Sie in „Tabelleneigenschaften: Drucken“ auf Seite 125.

Breite Pivot-Tabellen. Legt fest, wie Tabellen behandelt werden, die für die definierte Dokumentbreite zu breit sind. Standardmäßig wird die Tabelle umgebrochen. Die Tabelle wird in zwei Abschnitte unterteilt und die Zeilenbeschriftungen werden in jedem Abschnitt der Tabelle wiederholt. Alternativ können Sie breite Tabellen auch verkleinern oder keine Änderungen an breiten Tabellen vornehmen und damit zulassen, dass sie über den Rand der definierten Dokumentbreite hinausragen.

Fußnoten und Titelzeilen aufnehmen. Dient zur Festlegung der Aufnahme bzw. des Ausschlusses aller Fußnoten und Titelzeilen von Pivot-Tabellen.

Einträge in der Viewer-Gliederung als Folientitel verwenden. Fügt auf jeder beim Export erstellten Folie einen Titel ein. Jede Folie enthält ein einzelnes Element, das aus dem Viewer exportiert wurde. Der Titel wird aus dem Gliederungseintrag für das Element im Gliederungsfenster des Viewers gebildet.

Modellansichten. Die Aufnahme und der Ausschluss von Modellansichten werden standardmäßig von den Modelleigenschaften der einzelnen Modelle gesteuert. Sie können diese Einstellung außer Kraft setzen und alle Ansichten aufnehmen oder alle Ansichten mit Ausnahme der aktuell angezeigten Ansicht ausschließen. Weitere Informationen finden Sie in Modelleigenschaften. (Hinweis: Alle Modelle einschließlich Tabellen werden als Grafik exportiert.)

Druckoptionen für Export. Über diese Option wird ein Dialogfeld geöffnet, in dem Sie die Seitengröße und die Seitenränder für das exportierte Dokument festlegen können. Die Dokumentbreite, die festlegt, ob Tabellen umgebrochen oder verkleinert werden, ergeben sich aus der Seitenbreite minus den linken und rechten Seitenrand.

So legen Sie Exportoptionen für PowerPoint fest:

1. Wählen Sie **PowerPoint** als Exportformat aus.

2. Klicken Sie auf **Optionen ändern**.

Hinweis: Der Export nach PowerPoint ist nur unter Windows-Betriebssystemen verfügbar.

PDF-Optionen

Die folgenden Optionen sind für PDF verfügbar:

Lesezeichen einbetten. Mit dieser Option werden Lesezeichen in das PDF-Dokument aufgenommen, die den Einträgen in der Viewer-Gliederung entsprechen. Wie das Viewer-Gliederungsfenster können auch Lesezeichen die Navigation in Dokumenten mit einer Vielzahl an Ausgabeobjekten erheblich erleichtern.

Schriftarten einbetten. Durch das Einbetten von Schriftarten wird sichergestellt, dass das PDF-Dokument auf allen Computern gleich dargestellt wird. Andernfalls kann es, wenn im Dokument verwendete Schriftarten auf dem Computer, der zur Anzeige (oder zum Drucken) des PDF-Dokuments verwendet wird, nicht zur Verfügung stehen, durch Schriftartenersetzung zu suboptimalen Ergebnissen kommen.

Schichten in Pivot-Tabellen. Standardmäßig richtet sich die Aufnahme bzw. der Ausschluss von Pivot-Tabellenschichten nach den Tabelleneigenschaften der einzelnen Pivot-Tabellen. Sie können diese Einstellung außer Kraft setzen und alle Schichten aufnehmen oder alle Schichten mit Ausnahme der aktuell sichtbaren Schicht ausschließen. Weitere Informationen finden Sie in „Tabelleneigenschaften: Drucken“ auf Seite 125.

Modellansichten. Die Aufnahme und der Ausschluss von Modellansichten werden standardmäßig von den Modelleigenschaften der einzelnen Modelle gesteuert. Sie können diese Einstellung außer Kraft setzen und alle Ansichten aufnehmen oder alle Ansichten mit Ausnahme der aktuell angezeigten Ansicht ausschließen. Weitere Informationen finden Sie in Modelleigenschaften. (Hinweis: Alle Modelle einschließlich Tabellen werden als Grafik exportiert.)

So legen Sie PDF-Exportoptionen fest:

1. Wählen Sie **Portable Document Format** als Exportformat aus.
2. Klicken Sie auf **Optionen ändern**.

Andere Einstellungen, die die PDF-Ausgabe beeinflussen

Seite einrichten/Seitenattribute. Seitengröße, Ausrichtung, Ränder, Inhalt und Anzeige von Kopf- und Fußzeilen sowie die Größe des gedruckten Diagramms in PDF-Dokumenten werden über die Optionen für die Seiteneinrichtung und Seitenattribute gesteuert.

Tabelleneigenschaften/Tabellenvorlagen. Die Skalierung breiter und/oder langer Tabellen und der Druck von Tabellenschichten werden durch die Tabelleneigenschaften für die einzelnen Tabellen gesteuert. Diese Eigenschaften können auch unter "Tabellenvorlagen" gespeichert werden.

Standarddrucker/aktueller Drucker. Die Auflösung (DPI) des PDF-Dokuments ist die aktuelle Auflösungseinstellung für den Standarddrucker bzw. den aktuell ausgewählten Drucker (kann über "Seite einrichten" geändert werden). Die maximale Auflösung beträgt 1200 DPI. Wenn eine höhere Druckerauflösung eingestellt ist, wird für das PDF-Dokument eine Auflösung von 1200 DPI verwendet.

Hinweis: Dokumente mit höherer Auflösung können beim Drucken auf Druckern mit niedrigerer Auflösung zu schlechten Ergebnissen führen.

Textoptionen

Die folgenden Optionen sind für den Textexport verfügbar:

Pivot-Tabellenformat. Pivot-Tabellen können als durch Tabulatoren getrennter Text oder als durch Leerzeichen getrennter Text exportiert werden. Beim leerzeichengetrennten Format können Sie außerdem folgende Eigenschaften festlegen:

- **Spaltenbreite.** Mit **Automatisch anpassen** werden keine Spalteninhalte umgebrochen und jede Spalte ist so breit, wie die längste Beschriftung bzw. der längste Wert in der Spalte. Mit **Benutzerdefiniert** wird eine maximale Spaltenbreite festgelegt, die für alle Spalten in der Tabelle gilt. Bei Werten, die breiter sind, wird ein Zeilenumbruch durchgeführt, sodass sie sich auch auf die nächste Zeile in der betreffenden Spalte erstrecken.
- **Zeilen-/Spaltenbegrenzungszeichen.** Legt die Zeichen fest, die für Zeilen- und Spaltenrahmen verwendet werden. Um die Anzeige von Zeilen- und Spaltenrahmen zu unterdrücken, geben Sie als Werte Leerzeichen ein.

Schichten in Pivot-Tabellen. Standardmäßig richtet sich die Aufnahme bzw. der Ausschluss von Pivot-Tabellenschichten nach den Tabelleneigenschaften der einzelnen Pivot-Tabellen. Sie können diese Einstellung außer Kraft setzen und alle Schichten aufnehmen oder alle Schichten mit Ausnahme der aktuell sichtbaren Schicht ausschließen. Weitere Informationen finden Sie in „Tabelleneigenschaften: Drucken“ auf Seite 125.

Fußnoten und Titelzeilen aufnehmen. Dient zur Festlegung der Aufnahme bzw. des Ausschlusses aller Fußnoten und Titelzeilen von Pivot-Tabellen.

Modellansichten. Die Aufnahme und der Ausschluss von Modellansichten werden standardmäßig von den Modelleigenschaften der einzelnen Modelle gesteuert. Sie können diese Einstellung außer Kraft setzen und alle Ansichten aufnehmen oder alle Ansichten mit Ausnahme der aktuell angezeigten Ansicht ausschließen. Weitere Informationen finden Sie in Modelleigenschaften. (Hinweis: Alle Modelle einschließlich Tabellen werden als Grafik exportiert.)

So legen Sie Exportoptionen für reine Textdokumente fest:

1. Wählen Sie **Text** als Exportformat aus.
2. Klicken Sie auf **Optionen ändern**.

Nur Grafikoptionen

Die folgenden Optionen sind nur für den Export von Grafiken verfügbar:

Modellansichten. Die Aufnahme und der Ausschluss von Modellansichten werden standardmäßig von den Modelleigenschaften der einzelnen Modelle gesteuert. Sie können diese Einstellung außer Kraft setzen und alle Ansichten aufnehmen oder alle Ansichten mit Ausnahme der aktuell angezeigten Ansicht ausschließen. Weitere Informationen finden Sie in Modelleigenschaften. (Hinweis: Alle Modelle einschließlich Tabellen werden als Grafik exportiert.)

Grafikformatoptionen

Bei HTML- und Textdokumenten und nur zum Exportieren von Diagrammen können Sie das Grafikformat auswählen und für jedes Grafikformat verschiedene optionale Einstellungen festlegen.

So wählen Sie das Grafikformat und die Optionen für exportierte Diagramme aus:

1. Wählen Sie **HTML**, **Text** oder **Ohne (nur Grafiken)** als Dokumenttyp aus.
2. Wählen Sie das Grafikdateiformat aus der Dropdown-Liste aus.
3. Klicken Sie auf **Optionen ändern**, um die Optionen für das ausgewählte Grafikdateiformat zu ändern.

Exportoptionen für JPEG-Diagramme

Bildgröße (%)

Prozentsatz der ursprünglichen Diagrammgröße, bis zu 200 Prozent.

In Graustufen umwandeln

Wandelt die verschiedenen Farben in Grauschattierungen um.

Exportoptionen für BMP-Diagramme

Bildgröße (%)

Prozentsatz der ursprünglichen Diagrammgröße, bis zu 200 Prozent.

Bild zum Reduzieren der Dateigröße komprimieren

Ein verlustfreies Komprimierungsverfahren, das kleinere Dateien ohne Beeinträchtigung der Bildqualität erstellt.

Exportoptionen für PNG-Diagramme

Bildgröße (%)

Prozentsatz der ursprünglichen Diagrammgröße, bis zu 200 Prozent.

Farbtiefe

Bestimmt die Anzahl der Farben im exportierten Diagramm. Ein Diagramm, das mit einer beliebigen Farbtiefe gespeichert wird, verfügt über eine Mindestzahl tatsächlich verwendeter Farben und eine Höchstzahl zulässiger Farben in dieser Farbtiefe. Wenn das Diagramm beispielsweise drei Farben (Rot, Weiß und Schwarz) enthält, jedoch als Bild mit 16 Farben gespeichert wird, verbleibt das Diagramm im dreifarbigem Modus.

Anmerkung: Wenn die Anzahl der Farben im Diagramm die Anzahl der Farben für diese Tiefe übersteigt, werden die Farben zur Replizierung der Diagrammfarben gemischt.

Exportoptionen für EMF- und TIFF-Diagramme

Bildgröße. Prozentsatz der ursprünglichen Diagrammgröße, bis zu 200 Prozent.

Hinweis: Das Format EMF (Enhanced Metafile) ist nur unter Windows-Betriebssystemen verfügbar.

Exportoptionen für EPS-Diagramme

Bildgröße. Sie können die Größe als Prozentsatz der ursprünglichen Bildgröße (bis zu 200 Prozent) angeben oder Sie können eine Bildbreite in Pixel angeben (dabei richtet sich die Höhe nach dem Wert für die Breite und dem Seitenverhältnis). Das exportierte Bild ist immer proportional zum ursprünglichen Bild.

TIFF-Vorschaubild einschließen. Speichert eine Vorschau mit dem EPS-Bild im TIFF-Format zur Anzeige in Anwendungen, bei denen keine EPS-Bilder auf dem Bildschirm dargestellt werden können.

Schriftarten. Steuert die Behandlung von Schriftarten in EPS-Bildern.

- **Schriftartreferenzen verwenden.** Wenn die im Diagramm verwendeten Schriftarten auf dem Ausgabegerät zur Verfügung stehen, werden sie verwendet. Andernfalls verwendet das Ausgabegerät andere Schriftarten.
- **Schriftarten durch Kurven ersetzen.** Wandelt Schriftarten in PostScript-Kurvendaten um. Der Text selbst kann dann nicht mehr bearbeitet werden. Diese Option ist sinnvoll, wenn die im Diagramm verwendeten Schriftarten auf dem Ausgabegerät nicht zur Verfügung stehen.

Drucken von Viewer-Dokumenten

Es stehen zwei Optionen zum Drucken des Inhalts des Viewer-Fensters zur Verfügung:

Alle angezeigten Ausgaben. Hiermit werden nur die gegenwärtig im Inhaltsfenster angezeigten Elemente gedruckt. Ausgeblendete Elemente werden nicht gedruckt. (Ausgeblendete Elemente sind Elemente, die im Gliederungsfenster mit einem geschlossenen Buch dargestellt werden oder in reduzierten Gliederungsschichten verborgen sind.)

Auswahl. Hiermit werden nur die gegenwärtig im Gliederungs- und/oder Inhaltsfenster ausgewählten Elemente gedruckt.

Drucken von Ausgaben und Diagrammen

1. Wechseln Sie in das Viewer-Fenster. (Klicken Sie auf eine beliebige Stelle im Fenster.)
2. Wählen Sie in den Menüs Folgendes aus:
Datei > Drucken...
3. Wählen Sie die gewünschten Druckereinstellungen aus.
4. Klicken Sie zum Drucken auf **OK**.

Seitenansicht

Mit "Seitenansicht" erhalten Sie für jede Seite eine Vorschau des Ausdrucks von Viewer-Dokumenten. Bevor Sie Viewer-Dokumente drucken, sollten Sie diese in der Seitenansicht überprüfen. In der Seitenansicht werden Elemente angezeigt, die möglicherweise nicht sichtbar sind, wenn Sie nur das Inhaltsfenster des Viewers betrachten, beispielsweise:

- Seitenumbrüche
- verborgene Schichten von Pivot-Tabellen
- Umbrüche in breiten Tabellen
- auf jeder Seite zu druckende Kopf- und Fußzeilen

Wenn im Viewer derzeit eine Ausgabe ausgewählt ist, wird in der Seitenansicht nur diese Ausgabe angezeigt. Wenn Sie eine Vorschau für alle Ausgaben sehen möchten, darf im Viewer kein Element ausgewählt sein.

Seitenattribute: Kopf- und Fußzeilen

Kopf- und Fußzeilen sind die Informationen, die am oberen und unteren Rand jeder Seite ausgedruckt werden. Sie können beliebigen Text als Kopf- und Fußzeile eingeben. Außerdem können Sie die Symbolleiste in der Mitte des Dialogfelds verwenden, wenn Sie Folgendes einfügen möchten:

- Datum und Uhrzeit
- Seitennummern
- Dateiname aus dem Viewer
- Beschriftungen der Gliederungsüberschriften
- Titel und Untertitel
- **Als Standard übernehmen** verwendet die hier angegebenen Einstellungen als Standardeinstellungen für neue Viewer-Dokumente. (Hinweis: Dadurch werden die aktuellen Einstellungen auf den Registerkarten "Kopf-/Fußzeile" und "Optionen" zu den Standardeinstellungen.)
- Mit den Beschriftungen der Gliederungsüberschriften werden die Gliederungsüberschriften der ersten, zweiten, dritten und/oder vierten Stufe für das erste Element auf jeder Seite angezeigt.
- Mit den Symbolen für die Seitentitel und die Untertitel werden die aktuellen Titel und Untertitel der Seite gedruckt. Diese Dokumente können mit der Option "Neuer Seitentitel" im Menü "Einfügen" des Viewers erstellt werden oder mit den Befehlen TITLE und SUBTITLE. Wenn Sie keine Seitentitel oder -untertitel angegeben haben, wird diese Einstellung ignoriert.

Hinweis: Die Merkmale der Schriftart für neue Seitentitel und -untertitel werden auf der Registerkarte "Viewer" des Dialogfelds "Optionen" festgelegt (Menü "Bearbeiten", Befehl "Optionen"). Bei bestehenden Seitentiteln und -untertiteln können die Merkmale der Schriftart durch Bearbeiten der Titel direkt im Viewer geändert werden.

Um eine Vorschau des Erscheinungsbilds von Kopf- und Fußzeilen auf der gedruckten Seite zu erhalten, wählen Sie im Menü "Datei" die Option "Seitenansicht" aus.

Einfügen von Seitenkopfzeilen und -fußzeilen

1. Wechseln Sie in das Viewer-Fenster. (Klicken Sie auf eine beliebige Stelle im Fenster.)
2. Wählen Sie in den Menüs Folgendes aus:
Datei > Kopf-/Fußzeile...
3. Geben Sie die gewünschte Kopf- und/oder Fußzeile ein, die auf jeder Seite angezeigt werden soll.

Seitenattribute: Optionen

In diesem Dialogfeld werden die Größe der gedruckten Diagramme, der Abstand zwischen gedruckten Ausgabeelementen und die Seitennummerierung eingestellt.

- **Größe des gedruckten Diagramms.** Hier wird die Größe des gedruckten Diagramms im Verhältnis zur definierten Seitengröße festgelegt. Die Größe des gedruckten Diagramms hat keine Auswirkungen auf das Seitenverhältnis (Verhältnis Breite zu Höhe) des Diagramms. Die Gesamtgröße eines gedruckten Diagramms wird von dessen Höhe und Breite bestimmt. Wenn die äußeren Grenzen eines Diagramms den linken und rechten Rand der Seite berühren, kann das Diagramm in der Höhe nicht weiter vergrößert werden.
- **Abstand zwischen Elementen.** Hier wird der Abstand zwischen gedruckten Elementen festgelegt. Jede Pivot-Tabelle, jedes Diagramm und jedes Textobjekt ist ein separates Element. Diese Einstellung wirkt sich nicht auf die Anzeige von Elementen im Viewer aus.
- **Seitennummerierung beginnen mit.** Die Seiten werden fortlaufend ab der angegebenen Nummer nummeriert.
- **Als Standard übernehmen.** Diese Option verwendet die hier angegebenen Einstellungen als Standardeinstellungen für neue Viewer-Dokumente. (Beachten Sie, dass durch diese Option die aktuellen Kopf-/Fußzeileinstellungen und Optionseinstellungen zu den Standardeinstellungen werden.)

Ändern der Größe des gedruckten Diagramms, der Seitennummerierung und des Abstands zwischen gedruckten Elementen

1. Wechseln Sie in das Viewer-Fenster. (Klicken Sie auf eine beliebige Stelle im Fenster.)
2. Wählen Sie in den Menüs Folgendes aus:
Datei > Seitenattribute...
3. Klicken Sie auf die Registerkarte **Optionen**.
4. Ändern Sie die Einstellungen und klicken Sie auf **OK**.

Speichern der Ausgabe

Der Inhalt des Viewers kann gespeichert werden:

- **Ausgabeobjekt (*.cou).** Dieses Format speichert den gesamten Ausgabecontainer, einschließlich Grafik, Registerkarten, Anmerkungen usw. Dieses Format kann in IBM SPSS Modeler geöffnet und angezeigt werden, zu Projekten hinzugefügt und mithilfe von IBM SPSS Collaboration and Deployment Services Repository veröffentlicht und verfolgt werden. Dieses Format ist nicht kompatibel zu IBM SPSS Statistics.
- **Viewer-Dateien (*.spv).** Das Format, das zum Anzeigen von Dateien im Fenster **Viewer** verwendet wird. Wenn Sie mit diesem Format aus einem Modellnugget in IBM SPSS Modeler speichern, wird nur der Inhalt des Viewers von der Registerkarte **Modell** gespeichert.

Steuern Sie Optionen zum Speichern von Webberichten oder zum Speichern von Ergebnissen in anderen Formaten (beispielsweise Text, Word, Excel) mithilfe der Option **Exportieren** im Menü **Datei**.

Speichern eines Viewer-Dokuments

1. Wählen Sie in den Menüs des Viewer-Fensters Folgendes aus:
Datei > Speichern
2. Geben Sie den Namen des Dokuments ein und klicken Sie dann auf **Speichern**.

Optional können Sie folgende Aktionen ausführen:

Sperren von Dateien, um die Bearbeitung in IBM SPSS Smartreader zu verhindern

Wenn ein Viewer-Dokument gesperrt ist, können Sie Pivot-Tabellen bearbeiten (Zeilen und Spalten vertauschen, die angezeigte Schicht ändern usw.), aber Sie können keine Ausgabe ändern oder Änderungen am Viewer-Dokument in IBM SPSS Smartreader (einem separaten Produkt für die Arbeit mit Viewer-Dokumenten) speichern. Diese Einstellung hat keine Auswirkung auf Viewer-Dokumente, die in IBM SPSS Statistics oder IBM SPSS Modeler geöffnet wurden.

Verschlüsseln von Dateien mit einem Kennwort

Sie können in einem Viewer-Dokument gespeicherte vertrauliche Informationen schützen, indem Sie das Dokument mit einem Kennwort verschlüsseln. Nach der Verschlüsselung kann das Dokument nur unter Angabe des Kennworts geöffnet werden. IBM SPSS Smartreader-Benutzer müssen ebenfalls das Kennwort angeben, um die Datei zu öffnen.

So verschlüsseln Sie ein Viewer-Dokument:

- a. Klicken Sie im Dialogfeld "Ausgabe speichern unter" auf **Datei mit Kennwort verschlüsseln**.
- b. Klicken Sie auf **Speichern**.
- c. Geben Sie im Dialogfeld "Datei verschlüsseln" ein Kennwort an und wiederholen Sie es im Textfeld "Kennwort bestätigen". Kennwörter dürfen maximal 10 Zeichen lang sein. Es wird zwischen Groß- und Kleinschreibung unterschieden.

Warnung: Verloren gegangene Kennwörter können nicht wiederhergestellt werden. Wenn das Kennwort nicht mehr bekannt ist, kann die Datei nicht geöffnet werden.

Erstellen starker Kennwörter

- Verwenden Sie mindestens acht Zeichen.
- Verwenden Sie auch Zahlen, Symbole und Interpunktionszeichen in Ihrem Kennwort.
- Vermeiden Sie Abfolgen von Zahlen und Zeichen, wie beispielsweise "123" und "abc", und vermeiden Sie Wiederholungen, wie "111aaa".
- Erstellen Sie keine Kennwörter mit persönlichen Daten wie Geburtstage oder Spitznamen.
- Ändern Sie das Kennwort in regelmäßigen Abständen.

Hinweis: Das Speichern verschlüsselter Dateien in IBM SPSS Collaboration and Deployment Services Repository wird nicht unterstützt.

Bearbeiten von verschlüsselten Dateien

- Wenn Sie eine verschlüsselte Datei öffnen, Änderungen daran vornehmen und die Optionsfolge "Datei" > "Speichern" verwenden, wird die geänderte Datei mit demselben Kennwort gespeichert.
- Sie können das Kennwort einer verschlüsselten Datei ändern, indem Sie die Datei öffnen, die Schritte für die Verschlüsselung wiederholen und im Dialogfeld "Datei verschlüsseln" ein anderes Kennwort angeben.
- Sie können eine unverschlüsselte Version einer verschlüsselten Datei speichern, indem Sie die Datei öffnen, "Datei" > "Speichern unter" auswählen und im Dialogfeld "Ausgabe speichern unter" die Option **Datei mit Kennwort verschlüsseln** inaktivieren.

Hinweis: Verschlüsselte Datendateien und Ausgabedokumente können nicht in früheren IBM SPSS Statistics-Versionen als Version 21 geöffnet werden. Verschlüsselte Syntaxdateien können nicht in Versionen vor Version 22 geöffnet werden.

Speichern erforderlicher Modellinformationen mit dem Ausgabedokument

Diese Option gilt nur, wenn es im Ausgabedokument Modellviewerelemente gibt, die zusätzliche Informationen erfordern, um einige der interaktiven Funktionen zu aktivieren. Klicken Sie auf **Weitere Informationen**, um eine Liste dieser Modellviewerelemente und die interaktiven Funktionen anzuzeigen, für die zusätzliche Informationen erforderlich sind. Durch das Spei-

chern dieser Informationen steigt die Dokumentgröße möglicherweise erheblich. Wenn Sie sich gegen das Speichern dieser Informationen entscheiden, können Sie diese Ausgabeelemente immer noch öffnen, aber die angegebenen interaktiven Funktionen sind nicht verfügbar.

Pivot-Tabellen

Pivot-Tabellen

Viele Ergebnisse werden in Tabellen dargestellt, die interaktiv pivotiert werden können. Das heißt, Sie können die Zeilen, Spalten und Schichten neu anordnen.

Bearbeiten von Pivot-Tabellen

Für die Bearbeitung von Pivot-Tabellen stehen folgende Optionen zur Verfügung:

- Transponieren von Zeilen und Spalten
- Verschieben von Zeilen und Spalten
- Erstellen von mehrdimensionalen Schichten
- Anlegen und Aufheben von Gruppierungen für Zeilen und Spalten
- Ein- und Ausblenden von Zeilen, Spalten und anderen Informationen
- Drehen von Zeilen- und Spaltenbeschriftungen
- Anzeigen von Definitionen für Terme

Aktivieren von Pivot-Tabellen

Bevor Sie eine Pivot-Tabelle bearbeiten oder ändern können, müssen Sie sie zunächst **aktivieren**. So aktivieren Sie eine Tabelle:

1. Doppelklicken Sie auf die Tabelle.
ODER
2. Klicken Sie mit der rechten Maustaste auf die Tabelle und wählen Sie im Popup-Menü die Option **Inhalt bearbeiten** aus.
3. Wählen Sie im Untermenü entweder **Im Viewer** oder **In separatem Fenster** aus.

Pivotieren einer Tabelle

Eine Tabelle hat drei Dimensionen: Zeilen, Spalten und Schichten. Eine Dimension kann mehrere Elemente (oder überhaupt keine) enthalten. Sie können die Anordnung der Tabelle ändern, indem Sie Elemente zwischen den Dimensionen oder innerhalb der Dimensionen verschieben. Sie können die Elemente einfach mittels Ziehen und Ablegen an die gewünschte Position verschieben.

Ändern der Anzeigereihenfolge der Elemente innerhalb einer Dimension

So ändern Sie die Anzeigereihenfolge von Elementen in einer Tabellendimension (Zeile, Spalte bzw. Schicht):

1. Wenn die Pivot-Leisten noch nicht eingeblendet sind, wählen Sie folgende Befehle im Menü "Pivot-Tabelle" aus:
Pivot > Pivot-Leisten
2. Verschieben Sie mittels Ziehen und Ablegen die Elemente innerhalb der Dimension in der Pivot-Leiste.

Verschieben von Zeilen und Spalten innerhalb eines Dimensionselements

1. Klicken Sie in der Tabelle selbst (nicht in den Pivot-Leisten) auf die Beschriftung der zu verschiebenden Zeile oder Spalte.
2. Ziehen Sie die Beschriftung an die neue Position.

Transponieren von Zeilen und Spalten

Wenn Sie einfach nur die Zeilen und Spalten vertauschen möchten, gibt es eine einfache Alternative zur Verwendung der Pivot-Leisten:

1. Wählen Sie in den Menüs Folgendes aus:
Pivot > Zeilen und Spalten vertauschen

Damit erzielen Sie dasselbe Ergebnis, wie wenn Sie alle Zeilenelemente in die Spaltendimension und alle Spaltenelemente in die Zeilendimension ziehen würden.

Gruppieren von Zeilen oder Spalten

1. Wählen Sie die Beschriftungen für die Zeilen oder Spalten aus, die Sie gruppieren möchten. (Klicken und ziehen Sie mit der Maus oder halten Sie beim Klicken die Umschalttaste gedrückt, um mehrere Beschriftungen auszuwählen).
2. Wählen Sie in den Menüs Folgendes aus:
Bearbeiten > Gruppe

Es wird automatisch eine Gruppenbeschriftung eingefügt. Doppelklicken Sie auf die Gruppenbeschriftung, um den Text zu bearbeiten.

Hinweis: Um Zeilen oder Spalten zu einer bestehenden Gruppe hinzuzufügen, müssen Sie zunächst die Gruppierung der Elemente, die sich derzeit in der Gruppe befinden, aufheben. Anschließend können Sie eine neue Gruppe erstellen, die die zusätzlichen Elemente enthält.

Aufheben der Gruppierung von Zeilen oder Spalten

Beim Aufheben der Gruppierung wird automatisch die Gruppenbeschriftung gelöscht.

Drehen von Zeilen- und Spaltenbeschriftungen

Sie können die Beschriftungen für die innersten Spaltenbeschriftungen und die äußersten Zeilenbeschriftungen in einer Tabelle zwischen horizontaler und vertikaler Anzeige drehen.

1. Wählen Sie in den Menüs Folgendes aus:
Format > Innere Spaltenbeschriftungen drehen
ODER
Format > Äußere Zeilenbeschriftungen drehen

Nur die innersten Spaltenbeschriftungen und die äußersten Zeilenbeschriftungen können gedreht werden.

Sortieren von Zeilen

So sortieren Sie die Zeilen einer Pivot-Tabelle:

1. Aktivieren Sie die Tabelle.
2. Wählen Sie eine beliebige Zelle in der Spalte aus, die Sie zum Sortieren verwenden möchten. Um nur ein bestimmtes Set von Zeilen zu sortieren, wählen Sie zwei oder mehr aneinander angrenzende Zellen in der Spalte aus, die Sie zum Sortieren verwenden möchten.
3. Wählen Sie in den Menüs Folgendes aus:
Bearbeiten > Zeilen sortieren
4. Wählen Sie im Untermenü die Option **Aufsteigend** bzw. **Absteigend** aus.
 - Wenn die Zeilendimension Gruppen enthält, betrifft die Sortierung nur die Gruppe, die die Auswahl enthält.
 - Sortierung ist nicht über Gruppengrenzen hinweg möglich.
 - Es können keine Tabellen mit mehreren Elementen in der Zeilendimension sortiert werden.

Einfügen von Zeilen und Spalten

So fügen Sie eine Zeile bzw. Spalte in eine Pivot-Tabelle ein:

1. Aktivieren Sie die Tabelle.
2. Wählen Sie eine beliebige Zelle in der Tabelle aus.
3. Wählen Sie in den Menüs Folgendes aus:

Einfügen vor

ODER

Einfügen nach

Wählen Sie im Untermenü Folgendes aus:

Zeile

ODER

Spalte

- In jede Zelle der neuen Zeile bzw. Spalte wird ein Pluszeichen (+) eingefügt, um zu verhindern, dass die neue Zeile bzw. Spalte automatisch ausgeblendet wird, da sie leer ist.
- In eine Tabelle mit verschachtelten bzw. geschichteten Dimensionen wird eine Spalte oder Zeile auf jeder zugehörigen Dimensionsebene eingefügt.

Steuern der Anzeige von Variablen- und Wertbeschriftungen

Wenn Variablen deskriptive Variablen- bzw. Wertbeschriftungen enthalten, können Sie die Anzeige der Variablennamen und -beschriftungen sowie der Datenwerte und Wertbeschriftungen in Pivot-Tabellen steuern.

1. Aktivieren Sie die Pivot-Tabelle.
2. Wählen Sie in den Menüs Folgendes aus:

Ansicht > Variablenbeschriftungen

ODER

Ansicht > Wertbeschriftungen

3. Wählen Sie eine der folgenden Optionen aus dem Untermenü aus:
 - **Name** oder **Wert**. Es werden nur Variablennamen (bzw. Werte) angezeigt. Beschreibende Beschriftungen werden nicht angezeigt.
 - **Beschriftung**. Es werden nur beschreibende Beschriftungen angezeigt. Variablennamen (bzw. Werte) werden nicht angezeigt.
 - **Beide**. Sowohl Namen (bzw. Werte) als auch beschreibende Beschriftungen werden angezeigt.

Ändern der Ausgabesprache

So ändern Sie die Ausgabesprache in einer Pivot-Tabelle:

1. Aktivieren Sie die Tabelle.
2. Wählen Sie in den Menüs Folgendes aus:

Ansicht > Sprache
3. Wählen Sie eine der verfügbaren Sprachen aus.

Das Ändern der Sprache wirkt sich nur auf den von der Anwendung generierten Text aus, also beispielsweise auf Zeilen- und Spaltenbeschriftungen sowie auf Fußnotentext. Variablennamen und beschreibende Variablen- und Wertbeschriftungen sind nicht betroffen.

Navigieren in umfangreichen Tabellen

So navigieren Sie mit dem Navigationsfenster in umfangreichen Tabellen:

1. Aktivieren Sie die Tabelle.
2. Wählen Sie in den Menüs Folgendes aus:

Ansicht > Navigation

Rückgängigmachen von Änderungen

Sie haben die Möglichkeit, die aktuellste Änderung oder alle Änderungen an einer aktivierten Pivot-Tabelle rückgängig zu machen. Beide Aktionen gelten nur für Änderungen, die seit der aktuellsten Aktivierung der Tabelle vorgenommen wurden.

So machen Sie die aktuellste Änderung rückgängig:

1. Wählen Sie in den Menüs Folgendes aus:

Bearbeiten > Rückgängig

So machen Sie alle Änderungen rückgängig:

2. Wählen Sie in den Menüs Folgendes aus:

Bearbeiten > Wiederherstellen

Arbeiten mit Schichten

Sie können eine separate zweidimensionale Tabelle für jede Kategorie oder Kombination von Kategorien anzeigen lassen. Die Tabelle ist dabei sozusagen in Schichten gestapelt, und nur die oberste Schicht ist sichtbar.

Erstellen und Anzeigen von Schichten

So erstellen Sie Schichten:

1. Wenn die Pivot-Leisten noch nicht eingeblendet sind, wählen Sie im Menü **Pivot-Tabelle** Folgendes aus:

Pivot > Pivot-Leisten

2. Ziehen Sie ein Element aus der Zeilen- bzw. Spaltendimension in die Schichtdimension.

Durch das Verschieben von Elementen in die Schichtdimension wird eine mehrdimensionale Tabelle erstellt. Es wird jedoch nur ein einziger zweidimensionales Ausschnitt angezeigt. Die sichtbare Tabelle ist die Tabelle für die oberste Schicht. Wenn sich beispielsweise eine kategoriale Variable mit den Kategorien *Ja* und *Nein* in der Schichtdimension befindet, enthält die mehrdimensionale Tabelle zwei Schichten: eine für die Kategorie *Ja* und eine für die Kategorie *Nein*.

Ändern der angezeigten Schicht

1. Wählen Sie eine Kategorie in der Dropdown-Liste der Schichten aus (in der Pivot-Tabelle selbst, nicht in der Pivot-Leiste).

Gehe zu Kategorie in Schicht

Mit "Gehe zu Kategorie in Schicht" können Sie zu einer anderen Schicht in einer Pivot-Tabelle wechseln. Dieses Dialogfeld ist besonders dann nützlich, wenn viele Schichten vorhanden sind oder die ausgewählte Schicht viele Kategorien aufweist.

Ein- und Ausblenden von Elementen

Viele Zellentypen können ausgeblendet werden:

- Dimensionsbeschriftungen
- Kategorien, einschließlich der Zelle mit der Beschriftung und der Datenzellen in einer Zeile oder Spalte
- Kategoriebeschriftungen (wobei die Datenzellen nicht ausgeblendet werden)
- Fußnoten, Titel und Titelzeilen

Ausblenden von Zeilen und Spalten in einer Tabelle

Anzeigen ausgeblendeter Zeilen und Spalten in einer Tabelle

1. Wählen Sie in den Menüs Folgendes aus:

Ansicht > Alle Kategorien einblenden

Dadurch werden alle ausgeblendeten Zeilen und Spalten in der Tabelle angezeigt. Wenn Sie aber im Dialogfeld "Tabelleneigenschaften" für diese Tabelle die Option **Leere Zeilen und Spalten ausblenden** ausgewählt haben, bleiben vollständig leere Zeilen oder Spalten ausgeblendet.

Aus- und Einblenden von Dimensionsbeschriftungen

1. Markieren Sie die Dimensionsbeschriftung oder eine beliebige Kategoriebeschriftung in der Dimension.
2. Wählen Sie im Menü "Ansicht" bzw. im Popup-Menü die Option **Beschriftung für Dimension ausblenden** bzw. **Beschriftung für Dimension einblenden** aus.

Aus- und Einblenden von Tabellentiteln

So blenden Sie einen Titel aus:

1. Aktivieren Sie die Pivot-Tabelle.
2. Wählen Sie den Titel aus.
3. Wählen Sie im Menü "Ansicht" die Option **Ausblenden** aus.

So zeigen Sie ausgeblendete Titel an:

4. Wählen Sie im Menü "Ansicht" die Option **Alles einblenden** aus.

Tabellenvorlagen

Eine Tabellenvorlage umfasst ein Set von Eigenschaften, welche die äußere Form einer Tabelle definieren. Sie können eine vordefinierte Tabellenvorlage auswählen oder eine eigene Tabellenvorlage erstellen.

- Vor und nach dem Zuweisen von Tabellenvorlagen können Sie mithilfe der Zelleigenschaften die Zellenformate einzelner Zellen oder Gruppen von Zellen ändern. Die bearbeiteten Zellenformate bleiben auch nach dem Zuweisen einer neuen Tabellenvorlage erhalten.
- Wahlweise können Sie alle Zellen auf die Zellenformate zurücksetzen, die durch die aktuelle Tabellenvorlage definiert sind. Diese Option setzt die Formate der Zellen zurück, die zuvor bearbeitet wurden. Wenn in der Dateiliste für Tabellenvorlagen **Wie angezeigt** ausgewählt ist, werden alle bearbeiteten Zellen auf die aktuellen Tabelleneigenschaften zurückgesetzt.
- Nur die im Dialogfeld **Tabelleneigenschaften** definierten Tabelleneigenschaften werden in Tabellenvorlagen gespeichert. Tabellenvorlagen enthalten keine Änderungen an einzelnen Zellen.

Zuweisen einer Tabellenvorlage

1. Aktivieren Sie die Pivot-Tabelle.
2. Wählen Sie in den Menüs Folgendes aus:
Format > Tabellenvorlagen...
3. Wählen Sie eine Tabellenvorlage aus der Liste der Dateien für Tabellenvorlagen aus. Wenn Sie eine Datei in einem anderen Verzeichnis auswählen möchten, klicken Sie auf **Durchsuchen**.
4. Klicken Sie auf **OK**, um der ausgewählten Pivot-Tabelle die Tabellenvorlage zuzuweisen.

Bearbeiten oder Erstellen von Tabellenvorlagen

1. Wählen Sie im Dialogfeld "Tabellenvorlagen" aus der Liste der Dateien eine Tabellenvorlage aus.
 2. Klicken Sie auf **Tabellenvorlage bearbeiten**.
 3. Passen Sie die Tabelleneigenschaften an die gewünschten Attribute an und klicken Sie anschließend auf **OK**.
 4. Klicken Sie auf **Vorlage speichern**, um die bearbeitete Tabellenvorlage zu speichern, oder klicken Sie auf **Speichern unter**, um sie als neue Tabellenvorlage zu speichern.
- Änderungen in einer Tabellenvorlage wirken sich nur auf die ausgewählte Pivot-Tabelle aus. Bearbeitete Tabellenvorlagen werden anderen Tabellen, die diese Tabellenvorlage verwenden, nicht automatisch zugewiesen. Hierfür müssen Sie die betreffenden Tabellen auswählen und die Tabellenvorlage erneut zuweisen.

- Nur die im Dialogfeld **Tabelleneigenschaften** definierten Tabelleneigenschaften werden in Tabellenvorlagen gespeichert. Tabellenvorlagen enthalten keine Änderungen an einzelnen Zellen.

Tabelleneigenschaften

Unter "Tabelleneigenschaften" können Sie die allgemeinen Eigenschaften einer Tabelle festlegen und Zelleigenschaften für verschiedene Teile einer Tabelle bestimmen. Sie verfügen über folgende Möglichkeiten:

- Legen Sie allgemeine Eigenschaften, beispielsweise das Ausblenden leerer Zeilen und Spalten und das Anpassen der Druckeigenschaften fest.
- Legen Sie das Format und die Position von Fußnotenzeichen fest.
- Legen Sie spezielle Formate für Zellen im Datenbereich, für Zeilen- und Spaltenbeschriftungen und für andere Bereiche der Tabelle fest.
- Legen Sie die Breite und Farbe für die Rahmenlinien der einzelnen Tabellenbereiche fest.

Ändern der Eigenschaften von Pivot-Tabellen

1. Wählen Sie in den Menüs Folgendes aus:

Format > Tabelleneigenschaften...

2. Wählen Sie eine Registerkarte aus (**Allgemein**, **Fußnoten**, **Zellenformate**, **Rahmen** oder **Drucken**).
3. Wählen Sie die gewünschten Optionen aus.
4. Klicken Sie auf **OK** oder **Anwenden**.

Die neuen Eigenschaften werden der ausgewählten Pivot-Tabelle zugewiesen.

Tabelleneigenschaften: Allgemein

Einige Eigenschaften betreffen die ganze Tabelle. Sie verfügen über folgende Möglichkeiten:

- Blenden Sie leere Zeilen und Spalten ein oder aus. (Bei einer leeren Zeile oder Spalte steht in keiner der Datenzellen ein Wert.)
- Legen Sie die Platzierung der Zeilenbeschriftungen fest, die sich in der oberen linken Ecke befinden oder verschachtelt sein können.
- Legen Sie die maximale und minimale Spaltenbreite (angegeben in Punkt) fest.

So ändern Sie allgemeine Tabelleneigenschaften:

1. Klicken Sie auf die Registerkarte **Allgemein**.
2. Wählen Sie die gewünschten Optionen aus.
3. Klicken Sie auf **OK** oder **Anwenden**.

Anzuzeigende Zeilen festlegen:

Anmerkung: Diese Funktion gilt nur für aus früheren Versionen übernommene Tabellen.

Standardmäßig werden Tabellen mit vielen Zeilen in Abschnitten von 100 Zeilen angezeigt. So legen Sie die Anzahl von in einer Tabelle angezeigten Zeilen fest:

1. Wählen Sie **Tabelle nach Zeilen anzeigen** aus.
2. Klicken Sie auf **Anzuzeigende Zeilen festlegen**.
ODER
3. Wählen Sie im Menü einer aktivierten Pivot-Tabelle **Tabelle nach Zeilen anzeigen** und **Anzuzeigende Zeilen festlegen** aus.

Anzuzeigende Zeilen. Legt die maximale gleichzeitig angezeigte Anzahl Zeilen fest. Mithilfe der Navigationssteuerelemente können Sie zu verschiedenen Abschnitten der Tabelle wechseln. Der Mindestwert ist 10. Der Standardwert ist 100.

Witwen/Waisen-Toleranz. Legt die maximale Zeilenanzahl der innersten Zeilendimension der Tabelle fest, die zwischen den angezeigten Tabellenansichten aufgeteilt werden. Wenn zum Beispiel in jeder Gruppe der innersten Zeilendimension sechs Kategorien vorhanden sind, würde bei Angabe des Werts 6 für jede Gruppe die Aufteilung zwischen angezeigten Ansichten verhindert. Diese Einstellung kann dazu führen, dass die Gesamtzeilenanzahl in einer angezeigten Ansicht die angegebene maximal anzuzeigende Zeilenanzahl übersteigt.

Tabelleneigenschaften: Hinweise

Die Registerkarte "Fußnoten" des Dialogfelds "Tabelleneigenschaften" steuert die Fußnotenformatierung und den Tabellenkommentartext.

Fußnoten. Zu den Eigenschaften von Fußnotenzeichen gehören zum Beispiel Format und Position in Bezug auf den Text.

- Als Format für Fußnotenzeichen können Sie entweder Ziffern (1, 2, 3,...) oder Buchstaben (a, b, c,...) festlegen.
- Die Fußnotenzeichen können hochgestellt oder tiefgestellt eingefügt werden.

Kommentartext. Sie können jeder Tabelle einen Kommentartext hinzufügen.

- Kommentartext wird in einer QuickInfo angezeigt, wenn Sie den Mauszeiger über eine Tabelle im Viewer bewegen.
- Sprachausgabeprogramme lesen den Kommentartext, wenn der Fokus auf die Tabelle verschoben wurde.
- In der QuickInfo im Viewer werden nur die ersten 200 Zeichen des Kommentars angezeigt. Sprachausgabeprogramme lesen jedoch den gesamten Text.
- Wenn Sie eine Ausgabe in eine HTML-Datei oder einen Webbericht exportieren, wird der Kommentartext als alternativer Text verwendet.

Tabelleneigenschaften: Zellenformate

In Bezug auf die Formatierung ist eine Tabelle in verschiedene Bereiche aufgeteilt: Titel, Schichten, Eckenbeschriftungen, Zeilenbeschriftungen, Spaltenbeschriftungen, Daten, Titelzeilen und Fußnoten. Für jeden Bereich der Tabelle können Sie die entsprechenden Zellenformate ändern. Folgende Zellenformate können bearbeitet werden: Texteneigenschaften (z. B. Schriftart, Größe, Farbe und Schnitt), horizontale und vertikale Ausrichtung, Hintergrundfarben sowie die inneren Zellenränder.

Zellenformate werden immer ganzen Bereichen (Informationskategorien) zugewiesen. Sie sind nicht Eigenschaften einzelner Zellen. Diese Unterscheidung ist besonders beim Pivotieren von Tabellen wichtig.

Beispiel:

- Wenn Sie als Zellenformat für Spaltenbeschriftungen eine fette Schriftart festlegen, werden die Spaltenbeschriftungen fett angezeigt, und zwar unabhängig davon, was gerade in der Spaltendimension angezeigt wird. Wenn Sie ein Element aus der Spaltendimension in eine andere Dimension verschieben, bleibt die für Spaltenbeschriftungen geltende Fettformatierung für dieses Element nicht erhalten.
- Wenn Sie hingegen die Spaltenbeschriftungen fett formatieren, indem Sie die Zellen in einer aktivierten Pivot-Tabelle markieren und auf der Symbolleiste auf die Schaltfläche "Fett" klicken, bleibt der Inhalt dieser Zellen auch beim Verschieben in andere Dimensionen immer fett formatiert. Die Spaltenbeschriftungen behalten diese Formatierung dann nicht für andere Elemente, die in die Spaltendimension verschoben werden.

So ändern Sie Zellenformate:

1. Klicken Sie auf die Registerkarte **Zellenformate**.
2. Wählen Sie einen Bereich aus der Dropdown-Liste aus oder klicken Sie auf einen Bereich in der Vorschau.
3. Wählen Sie die Eigenschaften für diesen Bereich aus. Ihre Auswahl wird in der Vorschau angezeigt.

4. Klicken Sie auf **OK** oder **Anwenden**.

Abwechselnde Zeilenfarben

So setzen Sie eine andere Hintergrund- und/oder Textfarbe für abwechselnde Zeilen im Datenbereich der Tabelle:

1. Wählen Sie in der Dropdown-Liste "Bereich" die Option **Daten** aus.
2. Wählen Sie **Zeilenfarbe wechseln** in der Gruppe "Hintergrundfarbe".
3. Wählen Sie die Farben aus, die für den wechselnden Zeilenhintergrund und -text verwendet werden sollen.

Die Einstellung "Zeilenfarbe wechseln" hat nur Auswirkung auf den Datenbereich der Tabelle. Beschriftungsbereiche von Zeilen oder Spalten sind nicht betroffen.

Tabelleneigenschaften: Rahmen

Sie können für jeden Rahmen in einer Tabelle einen Linienstil und eine Farbe auswählen. Wenn Sie **Kein** in der Liste für die Linienstile auswählen, wird an der entsprechenden Position keine Linie gezeichnet.

So ändern Sie die Rahmenlinien für Tabellen:

1. Klicken Sie auf die Registerkarte **Rahmen**.
2. Wählen Sie eine Rahmenposition aus. Klicken Sie hierfür auf den entsprechenden Namen in der Liste oder klicken Sie auf eine Linie in der Vorschau.
3. Wählen Sie einen Linienstil oder die Option **Kein** aus.
4. Wählen Sie eine Farbe aus.
5. Klicken Sie auf **OK** oder **Anwenden**.

Tabelleneigenschaften: Drucken

Sie können folgende Eigenschaften für den Ausdruck von Pivot-Tabellen beeinflussen:

- Drucken aller Schichten oder nur der obersten Schicht der Tabelle und Drucken jeder Schicht auf einer getrennten Seite.
- Horizontales oder vertikales Verkleinern einer Tabelle zur Anpassung an die Seitengröße beim Drucken.
- Legen Sie die Einstellungen für Witwen/Waisen-Zeilen fest, in dem Sie die Mindestanzahl von Zeilen und Spalten festlegen, die in einem beliebigen gedruckten Abschnitt einer Tabelle enthalten sind, wenn die Tabelle zu breit und/oder zu lang für die definierte Seitengröße ist.

Hinweis: Wenn eine Tabelle nicht mehr auf die aktuelle Seite passt, weil sich bereits andere Ausgaben darüber befinden, die Tabelle im Prinzip aber auf eine Seite passen würde, wird sie automatisch auf einer neuen Seite gedruckt. Dies geschieht unabhängig von den Einstellungen für Witwen/Waisen-Zeilen.

- Berücksichtigen von Fortsetzungstexten für Tabellen, die nicht auf eine einzige Seite passen. Fortsetzungstexte können am unteren oder oberen Rand jeder Seite angezeigt werden. Wenn Sie keine dieser Optionen auswählen, werden Fortsetzungstexte nicht angezeigt.

So wählen Sie Druckeigenschaften für Pivot-Tabellen aus:

1. Klicken Sie auf die Registerkarte **Drucken**.
2. Wählen Sie die gewünschten Druckoptionen aus.
3. Klicken Sie auf **OK** oder **Anwenden**.

Zelleneigenschaften

Zelleneigenschaften werden ausgewählten Zellen zugewiesen. Sie können Schriftart, Wertformat, Ausrichtung, Ränder und Farben ändern. Zelleneigenschaften setzen Tabelleneigenschaften außer Kraft. Wenn Sie also die Tabelleneigenschaften ändern, bleiben vorher zugewiesene Zelleneigenschaften erhalten.

So ändern Sie Zelleneigenschaften:

1. Wählen Sie in der Tabelle die Zelle(n) aus.
2. Wählen Sie im Menü "Format" bzw. im Popup-Menü die Option **Zelleneigenschaften** aus.

Schriftart und Hintergrund

Mit der Registerkarte "Schriftart und Hintergrund" werden Schriftschnitt und -farbe sowie die Hintergrundfarbe für die ausgewählten Zellen in der Tabelle festgelegt.

Formatwert

Auf der Registerkarte "Formatwert" werden die Wertformate für die ausgewählten Zellen festgelegt. Sie können die Formate für Zahlen, Datum, Zeit und Währungen bestimmen und die Anzahl der angezeigten Dezimalstellen festlegen.

Ausrichtung und Ränder

Auf der Registerkarte "Ausrichtung und Ränder" können Sie die horizontale und vertikale Ausrichtung von Werten sowie den oberen, unteren, rechten und linken Rand für die ausgewählten Zellen festlegen. Bei einer horizontalen Ausrichtung vom Typ **Gemischt** werden die Inhalte der einzelnen Zellen gemäß ihrem jeweiligen Typ ausgerichtet. So sind beispielsweise Datumsangaben rechtsbündig und Textwerte linksbündig ausgerichtet.

Fußnoten und Titelzeilen

Sie können Fußnoten und Titelzeilen zu einer Tabelle hinzufügen. Außerdem können Sie Fußnoten bzw. Titelzeilen ausblenden, Fußnotenzeichen ändern und Fußnoten neu nummerieren.

Hinzufügen von Fußnoten und Titelzeilen

So fügen Sie Tabellen Titelzeilen hinzu:

1. Wählen Sie im Menü "Einfügen" die Option **Titelzeile** aus.

Sie können jedem Element einer Tabelle eine Fußnote hinzufügen. So fügen Sie eine Fußnote hinzu:

1. Klicken Sie auf einen Titel, eine Zelle oder eine Titelzeile in einer aktivierten Pivot-Tabelle.
2. Wählen Sie im Menü "Einfügen" die Option **Fußnote** aus.
3. Fügen Sie den Fußnotentext in den vorgesehenen Bereich ein.

Ein- und Ausblenden von Titelzeilen

So blenden Sie eine Titelzeile aus:

1. Wählen Sie die Titelzeile aus.
2. Wählen Sie im Menü "Ansicht" die Option **Ausblenden** aus.

So zeigen Sie ausgeblendete Titelzeilen an:

1. Wählen Sie im Menü "Ansicht" die Option **Alles einblenden** aus.

Ein- und Ausblenden von Fußnoten in einer Tabelle

So blenden Sie eine Fußnote aus:

1. Klicken Sie mit der rechten Maustaste auf die Zelle, die den Fußnotenverweis enthält, und wählen Sie im Popup-Menü die Option **Fußnoten ausblenden** aus.

ODER

2. Wählen Sie die Fußnote im Fußnotenbereich der Tabelle aus und wählen Sie im Popup-Menü die Option **Ausblenden** aus.

Hinweis: Wählen Sie bei aus früheren Versionen übernommenen Tabellen den Fußnotenbereich der Tabelle aus, wählen Sie im Popup-Menü die Option **Fußnote bearbeiten** aus und inaktivieren Sie die Eigenschaft **Sichtbar** für alle Fußnoten, die ausgeblendet werden sollen.

Bei Zellen mit mehreren Fußnoten haben Sie mit der zweiten Methode die Möglichkeit, nur einen Teil der Fußnoten auszublenden.

So blenden Sie alle Fußnoten in der Tabelle aus:

1. Wählen Sie alle Fußnoten im Fußnotenbereich der Tabelle aus (verwenden Sie Klicken und Ziehen oder halten Sie beim Klicken die Umschalttaste gedrückt, um die Fußnoten auszuwählen) und wählen Sie im Menü "Ansicht" die Option **Ausblenden** aus.

So zeigen Sie ausgeblendete Fußnoten an:

1. Wählen Sie im Menü "Ansicht" die Option **Alle Fußnoten einblenden** aus.

Fußnotenzeichen

Mit dem Befehl **Fußnotenzeichen** ändern Sie die Zeichen, die zum Kennzeichnen von Fußnoten verwendet werden. Standardmäßig werden abhängig von den Einstellungen in den Tabelleneigenschaften, fortlaufende Buchstaben oder Zahlen als Standardfußnotenzeichen verwendet. Sie können auch eine spezielle Markierung zuweisen. Spezielle Markierungen bleiben unverändert, wenn Sie die Nummerierung von Fußnoten ändern oder zwischen Zahlen und Buchstaben für Standardfußnotenzeichen wechseln. Die Anzeige von Buchstaben bzw. Zahlen für Standardmarkierungen und die hochgestellte bzw. tiefgestellte Position von Fußnotenzeichen werden durch die Registerkarte "Fußnoten" im Dialogfeld "Tabelleneigenschaften" festgelegt.

So ändern Sie Fußnotenzeichen:

1. Wählen Sie eine Fußnote aus.
2. Wählen Sie im Menü **Format** die Option **Fußnotenzeichen** aus.

Spezielle Markierungen sind auf zwei Zeichen beschränkt. Fußnoten mit speziellen Markierungen stehen vor Fußnoten mit fortlaufenden Ziffern oder Buchstaben im Fußnotenbereich der Tabelle. Durch die Änderung zu einer speziellen Markierung kann sich also die Reihenfolge der Fußnotenliste ändern.

Neunummerierung von Fußnoten

Nach dem Pivotieren einer Tabelle durch Verschieben von Zeilen, Spalten und Schichten ist die Reihenfolge der Fußnoten unter Umständen durcheinander. So nummerieren Sie die Fußnoten neu:

1. Wählen Sie im Menü "Format" die Option **Fußnoten neu nummerieren** aus.

Bearbeiten von Fußnoten in aus früheren Versionen übernommenen Tabellen

Bei aus früheren Versionen übernommenen Tabellen können Sie im Dialogfeld "Fußnoten bearbeiten" Fußnotentext eingeben und bearbeiten sowie die Schriftart einstellen, Fußnotenzeichen ändern und einzelne Fußnoten ausblenden oder löschen.

Wenn Sie eine neue Fußnote in eine aus einer früheren Version übernommene Tabelle einfügen, wird das Dialogfeld "Fußnoten bearbeiten" automatisch geöffnet. So bearbeiten Sie im Dialogfeld "Fußnoten bearbeiten" bestehende Fußnoten (ohne eine neue Fußnote zu erstellen):

Markierung. Standardmäßig werden abhängig von den Einstellungen in den Tabelleneigenschaften, fortlaufende Buchstaben oder Zahlen als Standardfußnotenzeichen verwendet. Um ein spezielles Fußnotenzeichen zuzuweisen, geben Sie einfach den neuen Wert in der Spalte "Markierung" ein. Spezielle Markierungen bleiben unverändert, wenn Sie die Nummerierung von Fußnoten ändern oder zwischen Zahlen und Buchstaben für Standardfußnotenzeichen wechseln. Die Anzeige von Buchstaben bzw. Zahlen für Standardmarkierungen und die hochgestellte bzw. tiefgestellte Position von Fußnotenzeichen werden durch die Registerkarte "Fußnoten" im Dialogfeld "Tabelleneigenschaften" festgelegt. Weitere Informationen finden Sie in „Tabelleneigenschaften: Hinweise“ auf Seite 124.

Um eine spezielle Markierung wieder in eine Standardmarkierung zu ändern, klicken Sie mit der rechten Maustaste im Dialogfeld "Fußnoten bearbeiten" auf das entsprechende Zeichen, wählen im Popup-Menü die Option **Fußnotenzeichen** aus und wählen im Dialogfeld "Fußnotenmarkierung" die Option "Standardmarkierung" aus.

Fußnote. Der Inhalt der Fußnote. In der Anzeige sind die aktuellen Einstellungen für Schriftart und Hintergrund zu sehen. Die Schriftarteinstellungen können über das Unterdialogfeld "Format" für einzelne Fußnoten geändert werden. Weitere Informationen finden Sie in „Schriftart- und Farbeinstellungen für Fußnoten“. Für alle Fußnoten wird dieselbe Hintergrundfarbe verwendet. Diese kann auf der Registerkarte "Schriftart und Hintergrund" im Dialogfeld "Zelleneigenschaften" geändert werden. Weitere Informationen finden Sie in „Schriftart und Hintergrund“ auf Seite 126.

Sichtbar. Laut Standardeinstellung sind alle Fußnoten sichtbar. Durch Inaktivieren des Kontrollkästchens "Sichtbar" können Sie Fußnoten ausblenden.

Schriftart- und Farbeinstellungen für Fußnoten: Bei aus früheren Versionen übernommenen Tabellen können Sie im Dialogfeld "Format" Schriftfamilie, Stil, Größe und Farbe für eine oder mehrere ausgewählte Fußnoten ändern:

1. Klicken Sie im Dialogfeld "Fußnoten bearbeiten" auf eine oder mehrere Fußnoten im Fußnotenraster.
2. Klicken Sie auf die Schaltfläche **Format**.

Die getroffene Auswahl in Bezug auf Schriftfamilie, Stil, Größe und Farbe wird auf alle ausgewählten Fußnoten angewendet.

Hintergrundfarbe, Ausrichtung und Ränder können im Dialogfeld "Zelleneigenschaften" festgelegt und auf alle Fußnoten angewendet werden. Es ist nicht möglich, diese Einstellungen für einzelne Fußnoten zu ändern. Weitere Informationen finden Sie in „Schriftart und Hintergrund“ auf Seite 126.

Breite der Datenzellen

Mit "Breite der Datenzellen" können Sie für alle Datenzellen dieselbe Breite festlegen.

So legen Sie die Breite für alle Datenzellen fest:

1. Wählen Sie in den Menüs Folgendes aus:
Format > Breite der Datenzellen...
2. Geben Sie einen Wert für die Zellenbreite an.

Ändern der Spaltenbreite

1. Klicken Sie auf die Spaltenbegrenzung und ziehen Sie daran.

Anzeigen der ausgeblendeten Rahmen in einer Pivot-Tabelle

Bei Tabellen, bei denen nur wenige Rahmen sichtbar sind, können Sie die ausgeblendeten Rahmen anzeigen lassen. Dadurch werden Vorgänge wie das Ändern von Spaltenbreiten vereinfacht.

1. Wählen Sie im Menü "Ansicht" die Option **Rasterlinien** aus.

Auswählen von Zeilen, Spalten und Zellen in Pivot-Tabellen

Sie können eine gesamte Zeile oder Spalte oder ein angegebenes Set von Datenzellen und Beschriftungen auswählen.

Auswahl mehrerer Zellen:

Auswählen > Datenzellen und Beschriftung

Drucken von Pivot-Tabellen

Das Aussehen gedruckter Pivot-Tabellen wird von mehreren Faktoren bestimmt. Diese Faktoren können Sie durch Ändern der Attribute für Pivot-Tabellen beeinflussen.

- Bei multidimensionalen Pivot-Tabellen (Tabellen mit Schichten) können Sie entweder alle Schichten oder nur die oberste (sichtbare) Schicht ausdrucken. Weitere Informationen finden Sie in „Tabelleneigenschaften: Drucken“ auf Seite 125.
- Sie können zu lange oder zu breite Pivot-Tabellen automatisch an die Seitengröße anpassen. Sie können auch die Positionen von Tabellenumbrüchen und Seitenumbrüchen festlegen. Weitere Informationen finden Sie in „Tabelleneigenschaften: Drucken“ auf Seite 125.

Verwenden Sie den Befehl "Seitenansicht" im Menü "Datei", um sich anzeigen zu lassen, wie die Pivot-Tabellen auf der gedruckten Seite aussehen würden.

Festlegen von Tabellenumbrüchen für breite und lange Tabellen

Pivot-Tabellen, die zu breit oder zu lang sind, um innerhalb der definierten Seitengröße gedruckt zu werden, werden automatisch aufgeteilt und in mehreren Abschnitten gedruckt. Sie verfügen über folgende Möglichkeiten:

- Legen Sie die Zeilen und Spalten fest, an denen große Tabellen geteilt werden.
- Geben Sie die Zeilen und Spalten an, bei denen Tabellen nicht getrennt werden sollen.
- Passen Sie umfangreiche Tabellen an die definierte Seitengröße an.

So legen Sie Zeilen- und Spaltenumbrüche für gedruckte Pivot-Tabellen fest:

1. Aktivieren Sie die Pivot-Tabelle.
2. Klicken Sie in eine beliebige Zelle in der Spalte links neben der Stelle oder klicken Sie in eine beliebige Zelle in der Zeile vor der Stelle, an der Sie den Umbruch einfügen möchten.
3. Wählen Sie in den Menüs Folgendes aus:

Format > Haltepunkte > Vertikaler Haltepunkt

ODER

Format > Haltepunkte > Horizontaler Haltepunkt

1. Aktivieren Sie die Pivot-Tabelle.
2. Klicken Sie in eine beliebige Zelle in der Spalte links neben der Stelle oder klicken Sie in eine beliebige Zelle in der Zeile vor der Stelle, an der Sie den Umbruch einfügen möchten.
3. Wählen Sie in den Menüs Folgendes aus:

Format > Haltepunkte > Vertikaler Haltepunkt

ODER

Format > Haltepunkte > Horizontaler Haltepunkt

So legen Sie Zeilen oder Spalten fest, die nicht getrennt werden sollen:

1. Wählen Sie die Beschriftungen der Zeilen oder Spalten aus, die nicht getrennt werden sollen. Sie können mehrere Zeilen- oder Spaltenbeschriftungen auswählen, indem Sie mit gedrückter Maustaste ziehen oder die Umschalttaste gedrückt halten und auf die erste und die letzte auszuwählende Beschriftung klicken.
2. Wählen Sie in den Menüs Folgendes aus:
Format > Haltepunkte > Zusammenhalten

So zeigen Sie Haltepunkte und Gruppen vom Typ "Zusammenhalten" an:

1. Wählen Sie in den Menüs Folgendes aus:

Format > Haltepunkte > Haltepunkte anzeigen

Haltepunkte werden als vertikale oder horizontale Linien angezeigt. Gruppen vom Typ "Zusammenhalten" werden als abgeblendete rechteckige Bereiche angezeigt, die von einem dunkleren Rand umgeben sind.

Hinweis: Die Anzeige von Haltepunkten und Gruppen vom Typ "Zusammenhalten" wird für aus früheren Versionen übernommene Tabellen nicht unterstützt.

So löschen Sie Haltepunkte und Gruppen vom Typ "Zusammenhalten":

So löschen Sie einen Haltepunkt:

1. Klicken Sie in eine beliebige Zelle in der Spalte links neben einem vertikalen Haltepunkt oder klicken Sie in eine beliebige Zelle in der Zeile über einem horizontalen Haltepunkt.

2. Wählen Sie in den Menüs Folgendes aus:

Format > Haltepunkte > Haltepunkt oder Gruppe löschen

So löschen Sie eine Gruppe vom Typ "Zusammenhalten":

3. Wählen Sie die Spalten- bzw. Zeilenbeschriftungen aus, die zur Angabe der Gruppe dienen.

4. Wählen Sie in den Menüs Folgendes aus:

Format > Haltepunkte > Haltepunkt oder Gruppe löschen

Alle Haltepunkte und Gruppen vom Typ "Zusammenhalten" werden automatisch gelöscht, wenn Sie eine Zeile oder Spalte pivotieren bzw. ihre Sortierung ändern. Dieses Verhalten gilt nicht für aus früheren Versionen übernommene Tabellen.

Erstellen eines Diagramms aus einer Pivot-Tabelle

1. Doppelklicken Sie auf die Pivot-Tabelle, um diese zu aktivieren.
2. Wählen Sie die Zeilen, Spalten oder Zellen aus, die im Diagramm angezeigt werden sollen.
3. Klicken Sie mit der rechten Maustaste auf eine beliebige Stelle im ausgewählten Bereich.
4. Wählen Sie im Popup-Menü die Option **Diagramm erstellen** und wählen Sie einen Diagrammtyp aus.

Aus früheren Versionen übernommene Tabellen

Sie können auswählen, dass Tabellen als aus früheren Versionen übernommene Tabellen gerendert werden sollen (in Version 19 als "Tabellen mit allen Funktionen" bezeichnet), die dann vollständig mit IBM SPSS Statistics-Versionen vor 20 kompatibel sind. Aus früheren Versionen übernommene Tabellen werden möglicherweise nur langsam gerendert und werden nur empfohlen, wenn Kompatibilität mit Versionen vor 20 erforderlich ist. Informationen zum Erstellen von aus früheren Versionen übernommenen Tabellen finden Sie in „Optionen: Pivot-Tabellen“ auf Seite 131.

Optionen

Optionen

Optionen steuern verschiedene Einstellungen.

Ändern der Einstellungen für die Optionen

1. Wählen Sie in den Menüs Folgendes aus:

Bearbeiten > Optionen...

2. Klicken Sie zum Ändern der Einstellungen auf die entsprechende Registerkarte.
3. Ändern Sie die Einstellungen.
4. Klicken Sie auf **OK** oder **Anwenden**.

Allgemeine Optionen

Maximale Anzahl Threads

Die Anzahl der Threads, die von Multithread-Prozeduren beim Berechnen von Ergebnissen verwendet werden. Die Einstellung **Automatisch** basiert auf der Anzahl verfügbarer Verarbeitungskerne. Geben Sie einen niedrigeren Wert an, wenn für andere Anwendungen während der Ausführung von Multithread-Prozessen mehr Verarbeitungsressourcen verfügbar sein sollen. Diese Option ist im Modus für verteilte Analysen inaktiviert.

Ausgabe

Führende Null für Dezimalwerte anzeigen. Zeigt führende Nullen für numerische Werte an, die nur aus Dezimalstellen bestehen. Bei Auswahl dieser Option wird beispielsweise der Wert ,123 als 0,123 angezeigt. Diese Einstellung gilt nicht für numerische Werte im Währungs- oder Prozentformat. Mit Ausnahme von ASCII-Dateien mit fester Länge (*.dat) werden führende Nullen beim Speichern der Daten in einer externen Datei nicht beibehalten.

Maßeinheit. Die Maßeinheit (Punkt, Zoll oder Zentimeter), in der Druckparameter, wie zum Beispiel die Zellenränder von Pivot-Tabellen, Zellenbreiten und Abstand zwischen Tabellen, angegeben werden.

Viewer-Optionen

Die Optionen für die Anzeige der Viewer-Ausgabe wirken sich nur auf die neuen Ausgaben aus, die nach dem Ändern der Einstellungen erzeugt werden. Änderungen dieser Einstellungen wirken sich nicht auf die bereits im Viewer angezeigte Ausgabe aus.

Anfänglicher Ausgabestatus. Hiermit legen Sie fest, welche Elemente bei jeder Ausführung einer Prozedur automatisch angezeigt oder ausgeblendet werden und wie die Elemente ausgerichtet sind. Sie können die Anzeige der folgenden Elemente steuern: Protokoll, Warnungen, Hinweise, Titel, Pivot-Tabellen, Diagramme und Textausgabe. Außerdem können Sie die Anzeige von Befehlen im Protokoll aktivieren oder inaktivieren. Sie können Syntaxbefehle aus dem Protokoll kopieren und in einer Syntaxdatei speichern.

Anmerkung: Im Viewer werden alle Ausgabeelemente linksbündig angezeigt. Die Einstellungen für die Ausrichtung wirken sich nur auf gedruckte Ausgaben aus. Zentrierte und rechtsbündige Elemente werden durch ein kleines Symbol gekennzeichnet.

Titel. Hiermit können Sie die Schriftart, Schriftgröße und die Farbe für die Titel neuer Ausgaben festlegen.

Seitentitel. Steuert die Schriftart, Größe und Farbe für neue Seitentitel sowie von Seitentiteln, die durch die Befehlssyntax TITLE und SUBTITLE erstellt wurden, oder von Seitentiteln, die durch die Option **Neuer Seitentitel** im Menü **Einfügen** erstellt wurden.

Textausgabe. Dies ist die für die Textausgabe verwendete Schriftart. Für die Textausgabe sollte ein Zeichensatz mit festem Abstand verwendet werden. Bei Einsatz von Proportionalschriften werden Tabellenausgaben fehlerhaft ausgerichtet.

Standardseiteneinrichtung. Steuert die Standardoptionen für Ausrichtung und Randeinstellungen beim Drucken.

Optionen: Pivot-Tabellen

Mit den Pivot-Tabellenoptionen werden verschiedene Optionen für die Anzeige von Pivot-Tabellen festgelegt.

Tabellenvorlage

Wählen Sie eine Tabellenvorlage aus der Dateiliste aus und klicken Sie auf **OK** oder **Übernehmen**. Sie können eine der mit IBM SPSS Statistics gelieferten Tabellenvorlagen verwenden oder im Pivot-Tabelleneditor Ihre eigene Tabellenvorlage erstellen. Wählen Sie dazu im Menü **Format** die Option **Tabellenvorlagen** aus.

- **Durchsuchen.** Hiermit können Sie eine Tabellenvorlage aus einem anderen Verzeichnis auswählen.
- **Verzeichnis für Tabellenvorlagen.** Hiermit können Sie ein anderes Standardverzeichnis für Tabellenvorlagen angeben. Wechseln Sie mit **Durchsuchen** in das gewünschte Verzeichnis, wählen Sie dort eine Tabellenvorlage aus und wählen Sie anschließend **Verzeichnis für Tabellenvorlagen**.

Hinweis: Tabellenvorlagen, die in früheren Versionen von IBM SPSS Statistics erstellt wurden, können nicht in Version 16.0 oder später verwendet werden.

Spaltenbreiten

Mit diesen Optionen können Sie die automatische Anpassung von Spaltenbreiten in Pivot-Tabellen steuern.

- **Nur an Beschriftungen anpassen.** Hiermit passen Sie die Spaltenbreite an die Breite der Spaltenbeschriftung an. Dies erzeugt kompaktere Tabellen, Datenwerte, die breiter als die Beschriftung sind, werden aber abgeschnitten.
- **Bei allen Tabellen für Beschriftung und Daten anpassen.** Hiermit passen Sie die Spaltenbreite an die Spaltenbeschriftung oder den größten Datenwert an, je nachdem, welcher der beiden länger ist. Dies erzeugt breitere Tabellen, stellt jedoch die Anzeige aller Werte sicher.

Standardbearbeitungsmodus

Mit dieser Option können Sie die Aktivierung von Pivot-Tabellen im Viewer-Fenster oder in einem separaten Fenster steuern. Standardmäßig werden durch Doppelklicken auf eine Pivot-Tabelle alle Tabellen (bis auf extrem große Tabellen) im Viewer-Fenster aktiviert. Sie können Pivot-Tabellen jedoch auch in einem separaten Fenster öffnen oder festlegen, dass kleine Pivot-Tabellen im Viewer-Fenster und Pivot-Tabellen ab einer bestimmten Größe in einem separaten Fenster geöffnet werden.

Kopieren breiter Tabellen im Rich Text Format in die Zwischenablage

Wenn Pivot-Tabellen im Word-/RTF-Format kopiert und eingefügt werden, werden Tabellen, die für die Dokumentbreite zu breit sind, entweder umgebrochen, an die Dokumentgröße angepasst oder nicht verändert.

Ausgabeoptionen

Ausgabeoptionen steuern die Standardeinstellung für verschiedene Ausgabetypen.

Behindertengerechte Bedienung mit Sprachausgabeprogramm. Steuert, wie Zeilen- und Spaltenbeschriftungen von Pivot-Tabellen von Sprachausgabeprogrammen gelesen werden. Sie können für jede Datenzeile die vollständige Zeilen- und Spaltenbeschriftung lesen oder nur die Beschriftungen lesen, die sich ändern, wenn Sie sich zwischen Datenzellen in der Tabelle bewegen.

Kapitel 8. Verarbeiten fehlender Werte

Fehlende Werte - Übersicht

In der Datenvorbereitungsphase des Data-Mining kann es häufig vorkommen, dass Sie fehlende Werte in den Daten ersetzen möchten. **Fehlende Werte** sind Werte im Dataset, die unbekannt, nicht erfasst oder falsch eingegeben sind. In der Regel sind solche Werte für die entsprechenden Felder ungültig. Beispielsweise sollte das Feld *Geschlecht* die Werte *M* und *W* enthalten. Wenn Sie die Werte *Y* oder *Z* in diesem Feld vorfinden, können Sie mit Sicherheit davon ausgehen, dass sie ungültig sind und daher als Leerzeichen interpretiert werden sollten. Ebenso ist ein negativer Wert für das Feld *Alter* sinnlos und sollte als Leerzeichen interpretiert werden. Häufig werden in einem Fragebogen derartige offenkundig falschen Werte absichtlich eingegeben oder es werden Felder leer gelassen, um eine Nichtbeantwortung zu kennzeichnen. Mitunter ist es vielleicht erforderlich, diese Leerzeichen genauer zu untersuchen, um festzustellen, ob eine Nichtbeantwortung, wie eine verweigerter Altersangabe, einen Faktor für die Voraussage eines bestimmten Ergebnisses darstellt.

Manche Modellierungsverfahren gehen besser mit fehlenden Daten um als andere. C5.0 und Apriori kommen beispielsweise gut mit Werten zurecht, die in einem Typknoten ausdrücklich als "fehlend" definiert sind. Bei anderen Modellierungsverfahren treten Probleme bei der Bearbeitung fehlender Werte und eine längere Trainingszeit auf, was zu ungenaueren Modellen führt.

In Clementine gibt es mehrere Arten von fehlenden Werten, die von IBM SPSS Modeler erkannt werden:

- **Nullwerte oder systemdefiniert fehlende Werte.** Bei diesen Werten handelt es sich um Nicht-Zeichenfolgswerte, die in der Datenbank bzw. der Quellendatei leer gelassen und nicht speziell in einem Quellen- oder Typknoten als "fehlend" definiert wurden. Systemdefiniert fehlende Werte werden als **\$null\$** angezeigt. Beachten Sie, dass leere Zeichenfolgen in IBM SPSS Modeler nicht als Nullen betrachtet werden, auch wenn sie von bestimmten Datenbanken (siehe unten) als Nullen behandelt werden können.
- **Leere Zeichenfolgen und leere Bereiche.** Leere Zeichenfolgswerte und leere Bereiche (Zeichenfolgen ohne sichtbare Zeichen) werden anders als Nullwerte behandelt. Leere Zeichenfolgen werden in den meisten Fällen als äquivalent mit leeren Bereichen (Leerzeichen) behandelt. Wenn Sie beispielsweise die Option auswählen, dass leere Bereiche in einem Quellen- oder Typknoten als Leerstellen behandelt werden sollen, gilt diese Einstellung auch für leere Zeichenfolgen.
- **Leere oder benutzerdefiniert fehlende Werte.** Es handelt sich hierbei um Werte wie unbekannt, 99 oder -1, die in einem Quellen- oder Typknoten ausdrücklich als fehlend definiert sind. Optional können Sie auch auswählen dass Nullen und leere Bereiche als Leerzeichen behandelt werden sollen. Dadurch können sie mit Flags für eine spezielle Behandlung versehen und aus den meisten Berechnungen ausgeschlossen werden. Beispielsweise können Sie die Funktion @BLANK verwenden, um diese Werte gemeinsam mit anderen Arten von fehlenden Werten als Leerstellen zu behandeln.

Einlesen gemischter Daten. Beachten Sie, dass beim Einlesen von Feldern mit numerischem Speichertyp (ganze Zahl, reelle Zahl, Zeit, Zeitmarke oder Datum) alle nicht numerischen Werte auf *null* oder auf *systemdefiniert fehlend* gesetzt werden. Dies liegt daran, dass im Gegensatz zu einigen anderen Anwendungen hier keine gemischten Speichertypen innerhalb eines Felds zulässig sind. Um dies zu vermeiden, sollten alle Felder mit gemischten Daten als Zeichenfolgen eingelesen werden, indem der Speichertyp im Quellenknoten bzw. in der externen Anwendung nach Bedarf geändert wird.

Lesen leerer Zeichenfolgen aus Oracle. Beim Lesen aus oder Schreiben in Oracle-Datenbanken sollten Sie darauf achten, dass Oracle im Gegensatz zu IBM SPSS Modeler und den meisten anderen Datenbanken leere Zeichenfolgswerte wie Nullwerte behandelt und speichert. Dies bedeutet, dass dieselben Daten sich unterschiedlich verhalten können und unterschiedliche Ergebnisse ausgeben können, je nachdem ob sie aus einer Oracle-Datenbank oder aus einer anderen Datenbank bzw. einer Datei extrahiert wurden.

Umgang mit fehlenden Werten

Entscheiden Sie vor dem Hintergrund Ihres Geschäfts- oder Domänenwissens, wie fehlende Werte behandelt werden sollen. Um die Trainingszeit zu verkürzen und die Genauigkeit zu erhöhen, können Sie beispielsweise einerseits Leerzeichen aus Ihrem Dataset entfernen. Andererseits kann das Vorhandensein leerer Werte neue Geschäftschancen eröffnen oder für zusätzliche Einblicke sorgen. Bei der Auswahl des besten Verfahrens sollten Sie folgende Aspekte Ihrer Daten berücksichtigen:

- Größe des Datensets
- Anzahl der Felder mit Leerzeichen
- Menge der fehlenden Daten

Im Allgemeinen gibt es zwei Ansätze, die Sie verfolgen können:

- Sie können Felder bzw. Datensätze mit fehlenden Werten ausschließen.
- Sie können fehlende Werte mit einer Reihe von Methoden imputieren, ersetzen oder erzwingen.

Beide Ansätze lassen sich mithilfe des Data Audit-Knotens weitgehend automatisieren. Beispielsweise können Sie einen Filterknoten generieren, der Felder ausschließt, die so viele fehlende Werte aufweisen, dass sie bei der Modellierung nicht mehr sinnvoll einsetzbar sind, und einen Superknoten generieren, der fehlende Werte für bestimmte oder alle verbleibenden Felder imputiert. Hier zeigt sich die Leistungsfähigkeit des Audits: Sie können nicht nur den aktuellen Status Ihrer Daten bewerten, sondern auch aufgrund dieser Bewertung Maßnahmen ergreifen.

Behandeln von Datensätzen mit fehlenden Werten

Wenn sich die Mehrzahl der fehlenden Werte in einer kleinen Anzahl von Datensätzen konzentriert, können Sie diese Datensätze einfach ausschließen. Eine Bank pflegt beispielsweise normalerweise detaillierte und vollständige Datensätze über ihre Darlehenskunden. Wenn die Bank bei der Gewährung von Darlehen an ihre Mitarbeiter weniger restriktiv ist, weisen die für Mitarbeiterdarlehen erhobenen Daten wahrscheinlich einige leere Felder auf. In diesem Fall stehen zwei Optionen zur Behandlung dieser fehlenden Werte zur Verfügung.

- Sie können die Personaldatensätze mithilfe eines Auswahlknotens entfernen.
- Bei einem großen Dataset können Sie alle Datensätze mit Leerstellen verwerfen.

Behandeln von Feldern mit fehlenden Werten

Wenn ein Großteil der fehlenden Werte in einer kleinen Anzahl von Feldern konzentriert ist, können Sie sie anstatt auf der Datensatzebene auf der Feldebene angehen. Mit diesem Ansatz können Sie außerdem mit der relativen Wichtigkeit bestimmter Felder experimentieren, bevor Sie sich für eine Methode zur Handhabung fehlender Werte entscheiden. Wenn ein Feld für die Modellierung unwichtig ist, lohnt es sich wahrscheinlich nicht, es beizubehalten, unabhängig davon, wie viele fehlende Werte es aufweist.

Ein Marktforschungsunternehmen erhebt beispielsweise Daten anhand eines allgemeinen Fragebogens mit 50 Fragen. Zwei Fragen betreffen das Alter und die politische Überzeugung. Diese Informationen werden von vielen Menschen ungern angegeben. In diesem Fall weisen die Felder *Alter* und *Politische_Überzeugung* viele fehlende Werte auf.

Feldmessniveau

Bei der Bestimmung der zu verwendenden Methode sollten Sie auch das Messniveau von Feldern mit fehlenden Werten in die Überlegungen einfließen lassen.

Numerische Felder. Für numerische Feldtypen wie *Stetig* sollten Sie immer nicht numerische Werte eliminieren, bevor Sie ein Modell erstellen, da viele Modelle nicht funktionieren, wenn Leerzeichen in numerischen Feldern enthalten sind.

Kategoriale Felder. Für kategoriale Felder, wie *Nominal* und *Flag*, ist das Ändern fehlender Werte zwar nicht erforderlich, steigert jedoch die Genauigkeit des Modells. Ein Modell mit dem Feld *Geschlecht* funktioniert zwar auch mit sinnlosen Werten, wie *Y* und *Z*, durch Entfernen aller Werte außer *M* und *W* wird jedoch die Genauigkeit des Modells erhöht.

Sichten bzw. Entfernen von Feldern

Zum Ausschluss von Feldern mit zu vielen fehlenden Werten stehen mehrere Optionen zur Verfügung:

- Mit dem Data Audit-Knoten können Sie Felder nach Qualität filtern.
- Mithilfe eines Merkmalauswahlknotens können Sie Felder ausschließen, die mehr als einen angegebenen Prozentsatz fehlender Werte enthalten und eine Rangordnung der Felder auf der Grundlage der Wichtigkeit in Bezug auf ein angegebenes Ziel erstellen.
- Statt die Felder zu entfernen, können Sie die Feldrolle anhand eines Typknotens auf **Keine** festlegen. Dadurch werden die Felder im Dataset belassen, aber aus den Modellierungsprozessen ausgeschlossen.

Behandeln von Datensätzen mit systemdefiniert fehlenden Werten

Was sind systemdefiniert fehlende Werte?

Systemdefiniert fehlende Werte stellen unbekannte oder ungültige Datenwerte dar. In Datenbanken werden diese Werte häufig als *Nullwerte* bezeichnet.

Systemdefiniert fehlende Werte unterscheiden sich von Leerwerten. Leerwerte werden normalerweise im Typknoten als bestimmte Werte oder Wertebereiche definiert, die als benutzerdefiniert fehlend betrachtet werden. Leerwerte werden im Modellierungskontext anders behandelt.

Erstellen von systemdefiniert fehlenden Werten

Systemdefiniert fehlende Werte können in Daten vorhanden sein, die aus einer Datenquelle gelesen werden (z. B. können Datenbanktabellen Nullwerte enthalten).

Systemdefiniert fehlende Werte können durch Verwendung des Werts **undef** in Ausdrücken erstellt werden. Zum Beispiel gibt der folgende CLEM-Ausdruck das Alter zurück, wenn ein Wert kleiner-gleich 30 ist, oder einen fehlenden Wert, wenn ein Wert größer als 30 ist:

```
if Alter > 30 then undef else Alter endif
```

Fehlende Werte können auch beim Ausführen eines Outer Joins, beim Teilen einer Zahl durch null, beim Berechnen der Quadratwurzel einer negativen Zahl und in anderen Situationen erstellt werden.

Anzeigen von systemdefiniert fehlenden Werten

Systemdefiniert fehlende Werte werden in Tabellen und in anderen Ausgaben als \$null\$ angezeigt.

Testen auf systemdefiniert fehlende Werte

Verwenden Sie die Sonderfunktion **@NULL**, damit true zurückgegeben wird, wenn der Argumentwert ein systemdefiniert fehlender Wert ist. Beispiel:

```
if @NULL(MyFieldName) then 'Ist null' else 'Ist nicht null' endif
```

An Funktionen übergebene systemdefiniert fehlende Werte

Systemdefiniert fehlende Werte, die an Funktionen übergeben werden, führen zu fehlenden Werten in der Ausgabe. Wenn z. B. der Wert des Felds *f1* ein systemdefiniert fehlender Wert in einer bestimmten Zeile ist, wird der Ausdruck $\log(f1)$ auch als systemdefiniert fehlender Wert für diese Zeile ausgewertet. Eine Ausnahme ist die Funktion **@NULL**.

Systemdefiniert fehlende Werte in Ausdrücken mit arithmetischen Operatoren

Die Anwendung von arithmetischen Operatoren auf Werte, die einen systemdefiniert fehlenden Wert enthalten, ergibt einen systemdefiniert fehlenden Wert. Wenn z. B. der Wert des Felds f1 ein systemdefiniert fehlender Wert in einer bestimmten Zeile ist, wird der Ausdruck $f1 + 10$ auch als systemdefiniert fehlender Wert für diese Zeile ausgewertet.

Systemdefiniert fehlende Werte in Ausdrücken mit logischen Operatoren

Wenn Sie mit systemdefiniert fehlenden Werten in Ausdrücken mit logischen Operatoren arbeiten, gelten die Regeln einer dreiwertigen Logik (*true*, *false* und *missing*), die in Wahrheitstabellen beschrieben werden können. Die Wahrheitstabellen für die üblichen logischen Operatoren *not*, *and* und *or* werden in den folgenden Tabellen gezeigt.

Table 5. Wahrheitstabelle für NOT

Operand	NOT-Operand
true	false
false	true
missing	missing

Table 6. Wahrheitstabelle für AND

Operand1	Operand2	Operand1 AND Operand2
true	true	true
true	false	false
true	missing	missing
false	true	false
false	false	false
false	missing	false
missing	true	missing
missing	false	false
missing	missing	missing

Table 7. Wahrheitstabelle für OR

Operand1	Operand2	Operand1 OR Operand2
true	true	true
true	false	true
true	missing	true
false	true	true
false	false	false
false	missing	missing
missing	true	true
missing	false	missing
missing	missing	missing

Systemdefiniert fehlende Werte in Ausdrücken mit Vergleichsoperatoren

Wenn Sie einen systemdefiniert fehlenden Wert mit einem nicht systemdefiniert fehlenden Wert vergleichen, wird das Ergebnis als systemdefiniert fehlender Wert und nicht als wahres oder falsches Ergebnis ausgewertet. Systemdefiniert fehlende Werte können miteinander verglichen werden; zwei systemdefiniert fehlende Werte gelten als gleich.

Systemdefiniert fehlende Werte in if/then/else/endif-Ausdrücken

Wenn Sie bedingte Ausdrücke verwenden und der bedingte Ausdruck einen systemdefiniert fehlenden Wert zurückgibt, wird der Wert aus der ELSE-Klausel von dem bedingten Ausdruck zurückgegeben.

Systemdefiniert fehlende Werte im Auswahlknoten

Wird der Auswahlausdruck für einen bestimmten Datensatz als fehlender Wert ausgewertet, wird der Datensatz nicht vom Auswahlknoten zurückgegeben (diese Aktion gilt für die Modi "Einschließen" und "Verwerfen").

Systemdefiniert fehlende Werte im Zusammenführungsknoten

Beim Zusammenführen unter Verwendung eines Schlüssels werden Datensätze, die systemdefiniert fehlende Werte in Schlüsselfeldern enthalten, nicht zusammengeführt.

Systemdefiniert fehlende Werte bei der Aggregation

Beim Aggregieren von Daten in Spalten werden fehlende Werte nicht in die Berechnung einbezogen. Zum Beispiel wird in einer Spalte mit drei Werten {1, 2 und undef} die Summe der Werte in der Spalte als 3 berechnet. Der Mittelwert wird als 1,5 berechnet.

Imputieren oder Auffüllen fehlender Werte

Wenn nur wenige fehlende Werte vorliegen, kann es sinnvoll sein, die Leerstellen durch eingefügte Werte zu ersetzen. Dies kann über den Data Audit-Bericht erfolgen, mit dem Sie nach Bedarf Optionen für bestimmte Felder angeben und dann einen Superknoten erstellen können, der Werte anhand einer Reihe von Methoden imputiert. Dies ist die flexibelste Methode, die außerdem die Bearbeitung einer großen Anzahl von Feldern in einem einzelnen Knoten ermöglicht.

Folgende Methoden stehen zum Imputieren fehlender Werte zur Verfügung:

Fest. Ersetzt einen festen Wert (Feldmittelwert, Mittelpunkt des Bereichs oder eine von Ihnen angegebene Konstante).

Zufällig. Ersetzt einen Zufallswert auf der Grundlage einer Normal- oder Gleichverteilung.

Ausdruck. Ermöglicht die Angabe eines benutzerdefinierten Ausdrucks. Beispielsweise könnten Sie Werte durch eine globale Variable ersetzen, die vom Globalwerteknoten erstellt wurde.

Algorithmus. Ersetzt einen von einem Modell vorhergesagten Wert auf der Grundlage eines C&RT-Algorithmus. Für jedes Feld, das mit dieser Methode imputiert wurde, gibt es ein separates C&RT-Modell sowie einen Füllerknoten, der Leerstellen und Nullen durch den vom Modell vorhergesagten Wert ersetzt. Anschließend werden die vom Modell generierten Vorhersagefelder mithilfe eines Filterknotens entfernt.

Um alternativ Werte für bestimmte Felder zu erzwingen, können Sie mithilfe eines Typknotens sicherstellen, dass die Feldtypen nur zulässige Werte abdecken, und dann die Spalte *Überprüfen* für die Felder, deren leere Werte ersetzt werden müssen, auf **Erzwingen** setzen.

CLEM-Funktionen für fehlende Werte

Zur Behandlung fehlender Werte stehen mehrere Funktionen zur Verfügung. Die folgenden Funktionen werden häufig in Auswahl- und Füllerknoten zum Verwerfen oder Auffüllen fehlender Werte verwendet.

- `count_nulls(LISTE)`
- `@BLANK(FELD)`
- `@NULL(FELD)`
- `undef`

Die @-Funktionen können in Verbindung mit der Funktion `@FIELD` verwendet werden, um vorhandene Leerzeichen oder Nullwerte in einem oder mehreren Feldern zu ermitteln. Die Felder können bei vorhandenen Leerzeichen oder Nullwerten einfach markiert werden oder sie können mit Ersatzwerten ausgefüllt oder in zahlreichen verschiedenen Operationen verwendet werden.

Sie können Nullen folgendermaßen in einer Liste von Feldern zählen:

```
count_nulls(['cardtenure' 'card2tenure' 'card3tenure'])
```

Wenn Sie eine der Funktionen verwenden, die eine Liste von Feldern als Eingabe akzeptieren, können die Sonderfunktionen `@FIELDS_BETWEEN` und `@FIELDS_MATCHING` verwendet werden, wie im folgenden Beispiel gezeigt:

```
count_nulls(@FIELDS_MATCHING('card*'))
```

Mit der Funktion `undef` können Sie Felder mit dem systemdefiniert fehlenden Wert, der als **\$null\$** angezeigt wird, ausfüllen. Um beispielsweise einen beliebigen numerischen Wert zu ersetzen, können Sie eine bedingte Anweisung verwenden. Beispiel:

```
if not(Alter > 17) or not(Alter < 66) then undef else Alter endif
```

Mit dieser Anweisung werden alle Werte ersetzt, die nicht mit einem systemdefiniert fehlenden Wert, der als **\$null\$** angezeigt wird, in einem Bereich liegen. Mit der Funktion `not()` können Sie alle anderen numerischen Werte, einschließlich negativer Werte, erfassen. Weitere Informationen finden Sie in „Funktionen zum Umgang mit Leerstellen und Nullwerten“ auf Seite 188.

Hinweis zum Verwerfen von Datensätzen

Beachten Sie beim Verwenden eines Auswahlknotens zum Verwerfen von Datensätzen, dass die Syntax eine dreiwertige Logik verwendet und automatisch Nullwerte in `SELECT`-Anweisungen integriert. Um Nullwerte (systemdefiniert fehlende Werte) aus einem `SELECT`-Ausdruck auszuschließen, müssen Sie dies ausdrücklich angeben, indem Sie `and` oder `not` im Ausdruck verwenden. Um beispielsweise alle Datensätze, deren Medikamenttyp Medikament C ist, auszuwählen und einzuschließen, wird folgende `SELECT`-Anweisung verwendet:

```
Drug = 'drugC' and not(@NULL(Drug))
```

In früheren Versionen wurden Nullwerte in solchen Fällen ausgeschlossen.

Kapitel 9. Erstellen von CLEM-Ausdrücken

Informationen zu CLEM

CLEM (Control Language for Expression Manipulation) ist eine leistungsfähige Sprache zur Analyse und Bearbeitung von Daten, die durch IBM SPSS Modeler-Streams fließen. Data-Mining-Experten verwenden CLEM in hohem Maße für Streamoperationen, um so einfache Aufgaben durchzuführen, wie das Ableiten von Profit aus Kosten- und Einkommensdaten, oder um so komplexe Aufgaben durchzuführen, wie das Umwandeln von Webprotokolldaten in ein Set von Feldern und Datensätzen mit nützlichen Informationen.

CLEM wird innerhalb von IBM SPSS Modeler für folgende Zwecke eingesetzt:

- Bedingungen in Datensatzfeldern vergleichen und evaluieren
- Werte für neue Felder ableiten
- Neue Werte für vorhandene Felder ableiten
- Reihenfolge der Datensätze erarbeiten
- Daten aus Datensätzen in Berichte einfügen

CLEM-Ausdrücke sind für die Datenvorbereitung in IBM SPSS Modeler unverzichtbar und können in einer Vielzahl von Knoten verwendet werden - von Datensatz- und Feldoperationen (Auswahl, Balancierung, Füller) bis hin zu Plots und Ausgaben (Analyse, Bericht, Tabelle). Sie können beispielsweise einen CLEM-Ausdruck in einem Ableitungsknoten verwenden, um ein neues Feld anhand einer Formel, wie das Verhältnis, zu erstellen.

CLEM-Ausdrücke können auch für globales Suchen und Ersetzen verwendet werden. Der Ausdruck @NULL(@FIELD) beispielsweise kann in einem Füllerknoten verwendet werden, um **systemdefiniert fehlende Werte** durch den ganzzahligen Wert 0 zu ersetzen. (Zum Ersetzen von **benutzerdefiniert fehlenden Werten**, auch als "fehlende Werte" bezeichnet, verwenden Sie die Funktion @BLANK.)

Es können auch komplexere CLEM-Ausdrücke erstellt werden. Beispielsweise können Sie neue Felder auf der Grundlage eines bedingten Regelsets ableiten, etwa eine neue Wertekategorie, die mithilfe folgender Ausdrücke erstellt wird: `If: CardID = @OFFSET(CardID,1), Then: @OFFSET(ValueCategory,1), Else: 'exclude'`.

In diesem Beispiel wird die Funktion @OFFSET für folgende Anweisung verwendet: Wenn der Wert des Felds *CardID* für einen bestimmten Datensatz mit dem für den vorherigen Datensatz identisch ist, soll der Wert des Felds *ValueCategory* für den vorherigen Datensatz zurückgegeben werden. Andernfalls soll die Zeichenfolge "ausschließen" zugewiesen werden. Mit anderen Worten: Wenn die Werte für *CardID* für angrenzende Datensätze identisch sind, sollte ihnen dieselbe Wertekategorie zugewiesen werden. (Datensätze mit der Zeichenfolge "ausschließen" können später mithilfe eines Auswahlknotens entfernt werden.)

CLEM-Beispiele

Zur Veranschaulichung der korrekten Syntax sowie der mit CLEM möglichen Ausdrücke werden im Folgenden Beispielausdrücke genannt.

Einfache Ausdrücke

Formeln können sehr einfach sein. Die folgende Formel leitet beispielsweise ein neues Feld basierend auf den Werten der Felder *Nach* und *Vor* ab:

$$(Nach - Vor) / Vor * 100.0$$

Beachten Sie, dass Feldnamen nicht in Anführungszeichen gesetzt werden, wenn sie sich auf die Feldwerte beziehen.

Ähnlich gibt der folgende Ausdruck einfach den Logarithmus der einzelnen Werte des Felds *Gehalt* zurück.

```
log(Gehalt)
```

Komplexe Ausdrücke

Ausdrücke können auch relativ lang und komplex sein. Der folgende Ausdruck gibt den Wert *true* zurück, wenn der Wert zweier Felder (*\$KX-Kohonen* und *\$KY-Kohonen*) innerhalb der angegebenen Bereiche liegt. Beachten Sie, dass die Feldnamen in diesem Fall in einfache Anführungszeichen gesetzt werden, da die Feldnamen Sonderzeichen enthalten.

```
(' $KX-Kohonen' >= -0.2635771036148072 and ' $KX-Kohonen' <= 0.3146203637123107  
and ' $KY-Kohonen' >= -0.18975617885589602 and  
' $KY-Kohonen' <= 0.17674794197082522) -> T
```

Einige Funktionen, wie Zeichenfolgenfunktionen, erfordern die Eingabe mehrerer Parameter in der korrekten Syntax. Im folgenden Beispiel wird die Funktion *subscr* verwendet, um das erste Zeichen eines Felds *produce_ID* zurückzugeben. Dabei wird angegeben, ob ein Element organisch, genetisch verändert oder herkömmlich ist. Die Ergebnisse eines Ausdrucks werden durch *-> 'Ergebnis'* beschrieben.

```
subscr(1,produce_ID) -> 'c'
```

Ebenso lautet der folgende Ausdruck:

```
stripchar('3','123') -> '12'
```

Beachten Sie unbedingt, dass Zeichen immer in einfache umgekehrte Anführungszeichen gesetzt werden müssen.

Kombinieren von Funktionen in einem Ausdruck

CLEM-Ausdrücke bestehen häufig aus einer Kombination von Funktionen. Die folgende Funktion kombiniert *subscr* und *lowertoupper*, um das erste Zeichen von *produce_ID* zurückzugeben und es in einen Großbuchstaben zu konvertieren.

```
lowertoupper(subscr(1,produce_ID)) -> 'C'
```

Der gleiche Ausdruck kann in Kurzform wie folgt geschrieben werden:

```
lowertoupper(produce_ID(1)) -> 'C'
```

Eine andere häufig verwendete Kombination von Funktionen lautet:

```
locchar_back('n', (length(web_page)), web_page)
```

Mit diesem Ausdruck wird das Zeichen 'n' in den Werten des Felds *web_page* gesucht, wobei vom letzten Zeichen des Feldwerts an rückwärts gelesen wird. Durch Integrieren der Funktion *length* berechnet der Ausdruck dynamisch die Länge des aktuellen Werts, anstatt eine statische Zahl wie 7 zu verwenden. Diese Zahl ist ungültig für Werte, die weniger als sieben Zeichen enthalten.

Sonderfunktionen

Es sind zahlreiche Sonderfunktionen (die mit dem Symbol @ beginnen) verfügbar. Häufig verwendete Funktionen sind u. a.:

```
@BLANK('referrer ID') -> T
```

Häufig werden Sonderfunktionen kombiniert verwendet. Diese Methode wird im Allgemeinen dazu verwendet, Leerstellen in mehr als einem Feld gleichzeitig zu kennzeichnen.

@BLANK(@FIELD)-> T

Weitere Beispiele werden in der CLEM-Dokumentation erörtert. Weitere Informationen finden Sie in „CLEM-Referenz - Überblick“ auf Seite 155.

Werte und Datentypen

CLEM-Ausdrücke ähneln Formeln, die aus Werten, Feldnamen, Operatoren und Funktionen gebildet werden. Der einfachste gültige CLEM-Ausdruck besteht aus einem Wert oder einem Feldnamen. Gültige Werte sind zum Beispiel:

```
3
1.79
'banana'
```

Feldnamen sind zum Beispiel:

```
Product_ID
'$P-NextField'
```

Dabei ist *Product* der Name eines Feldes aus einem Warenkorbdataset, '*\$P-NextField*' der Name eines Parameters und der Wert des Ausdrucks ist der Wert des benannten Feldes. In der Regel beginnen Feldnamen mit einem Buchstaben und können auch Ziffern und Unterstriche (`_`) enthalten. Wenn Sie den Namen in Anführungszeichen setzen, können Sie Namen verwenden, die diese Regeln nicht einhalten.

CLEM-Werte können folgende sein:

- Zeichenfolgen - z. B. "c1", "Typ 2", "freier Text"
- Ganze Zahlen - z. B. 12, 0, -189
- Reelle Zahlen - z. B. 12,34, 0,0, -0,0045
- Datums-/Zeitfelder - z. B. 05/12/2002, 12/05/2002, 12/05/02

Es können auch folgende Elemente verwendet werden:

- Zeichencodes - z. B. `a` oder 3
- Elementlisten - z. B. [1 2 3], ['Typ 1' 'Typ 2']

Zeichencodes und -listen erscheinen normalerweise nicht als Feldwerte. In der Regel werden sie als Argumente für CLEM-Funktionen verwendet.

Regeln für Anführungszeichen

Die Software ist zwar flexibel in Bezug auf die Bestimmung der in einem CLEM-Ausdruck verwendeten Felder, Werte, Parameter und Zeichenfolgen. Die folgenden allgemeinen Regeln dienen jedoch als Liste der "besten Verfahrensweisen" für das Erstellen von Ausdrücken.

- **Zeichenfolgen** - Setzen Sie Zeichenfolgen immer in doppelte Anführungszeichen ("Typ 2" oder "Wert"). Bei einfachen Anführungszeichen besteht die Gefahr, dass sie mit in Anführungszeichen gesetzten Feldern verwechselt werden.
- **Zeichen** - Verwenden Sie immer einfache umgekehrte Anführungszeichen, nämlich ` . Beachten Sie beispielsweise das Zeichen d in der Funktion stripchar(`d`, "drugA"). Die einzige Ausnahme zu dieser Regel ist die Verwendung einer Ganzzahl für den Verweis auf ein bestimmtes Zeichen in einer Zeichenfolge. Beachten Sie beispielsweise das Zeichen 5 in der Funktion lowertoupper("drugA"(5)) -> "A".
Hinweis: Auf einer englischen und US-amerikanischen Standardtastatur befindet sich die Taste für das umgekehrte Anführungszeichen (Accent grave, Unicode 0060) unterhalb der Escapetaste, auf europäischen Standardtastaturen unterhalb der Taste F10.

- **Felder** - Felder werden in der Regel in CLEM-Ausdrücken nicht in Anführungszeichen gesetzt (subscr(2,arrayID)) → CHAR). Wenn Leerzeichen oder andere Sonderzeichen eingeschlossen werden müssen, verwenden Sie einfache Anführungszeichen ('Auftragsnummer'). Felder, die in Anführungszeichen gesetzt sind, im Dataset jedoch nicht definiert sind, werden irrtümlich als Zeichenfolgen interpretiert.
- **Parameter** - Verwenden Sie immer einfache Anführungszeichen ('\$P-threshold').

Ausdrücke und Bedingungen

CLEM-Ausdrücke können ein Ergebnis zurückgeben (beim Ableiten neuer Werte). Beispiel:

```
Weight * 2.2
Age + 1
sqrt(Signal-Echo)
```

Oder sie können *wahr* oder *falsch* überprüfen (beim Auswählen einer Bedingung). Beispiel:

```
Drug = "drugA"
Age < 16
not(PowerFlux) and Power > 2000
```

In CLEM-Ausdrücken können Operatoren und Funktionen willkürlich miteinander kombiniert werden. Beispiel:

```
sqrt(abs(Signal)) * max(T1, T2) + Baseline
```

Die Reihenfolge, in der der Ausdruck berechnet wird, wird durch Klammern und Operatorvorrang bestimmt. Im genannten Beispiel lautet die Evaluierungsreihenfolge wie folgt:

- abs(Signal) wird berechnet und sqrt wird auf das Ergebnis angewendet.
- max(T1, T2) wird berechnet.
- Die beiden Ergebnisse werden multipliziert: x hat Vorrang vor +.
- Schließlich wird Baseline zum Ergebnis addiert.

Die absteigende Vorrangfolge (d. h. Rechenoperationen, die zuerst ausgeführt werden, bis hin zu Operationen, die zuletzt ausgeführt werden) lautet wie folgt:

- Funktionsargumente
- Funktionsaufrufe
- **xx**
- **x / mod div rem**
- + -
- > < >= <= /== == = /=

Wenn Sie den Vorrang umgehen möchten oder in Bezug auf die Evaluierungsreihenfolge unsicher sind, können Sie zur Verdeutlichung Klammern verwenden. Beispiel:

```
sqrt(abs(Signal)) * (max(T1, T2) + Baseline)
```

Stream-, Sitzungs- und Superknotenparameter

Parameter können für die Verwendung in CLEM-Ausdrücken und Scripts definiert werden. Es handelt sich dabei im Grunde um benutzerdefinierte Variablen, die mit dem aktuellen Stream, der aktuellen Sitzung bzw. dem aktuellen Superknoten gespeichert und persistent gemacht werden. Ein Zugriff auf diese Variablen ist über die Benutzerschnittstelle und über Scripts möglich. Wenn Sie also einen Stream speichern, werden alle für diesen Stream festgelegten Parameter ebenfalls gespeichert. (Dies unterscheidet sie von lokalen Scriptvariablen, die nur in dem Script verwendet werden können, in dem sie deklariert sind.)

Parameter werden oft beim Scripting verwendet, um das Verhalten des Scripts zu steuern, indem Informationen zu Feldern und Werten angegeben werden, die nicht im Script fest codiert werden müssen.

Der Bereich eines Parameters hängt davon ab, wo er festgelegt wird:

- Streamparameter können in einem Stream-Script oder im Dialogfeld "Streameigenschaften" festgelegt werden und sie stehen für alle Knoten im Stream zur Verfügung. Sie werden in der Parameterliste im Expression Builder angezeigt.
- Sitzungsparameter können in einem Standalone-Script oder im Dialogfeld "Sitzungsparameter" festgelegt werden. Sie stehen allen Streams zur Verfügung, die in der aktuellen Sitzung verwendet werden (alle Streams auf der Registerkarte "Streams" im Managerbereich).

Parameter können auch für Superknoten festgelegt werden. In diesem Fall sind sie nur für Knoten sichtbar, die in dem betreffenden Superknoten gekapselt sind.

Verwenden von Parametern in CLEM-Ausdrücken

Parameter werden in CLEM-Ausdrücken durch `$P-pname` dargestellt. Dabei ist `pname` der Name des Parameters. Wenn Parameter in CLEM-Ausdrücken verwendet werden, müssen sie in einfache Anführungszeichen eingeschlossen werden. Beispiel: `'$P-scale'`.

Die verfügbaren Parameter können einfach mit dem Expression Builder angezeigt werden. So zeigen Sie die aktuellen Parameter an:

1. Klicken Sie in einem beliebigen Dialogfeld, in das CLEM-Ausdrücke eingegeben werden können, auf die Schaltfläche des Expression Builder.
2. Wählen Sie in der Liste "Felder" die Option **Parameter**.

Sie können in der Liste Parameter auswählen, um sie in den CLEM-Ausdruck einzufügen. Weitere Informationen finden Sie in „Auswählen von Feldern, Parametern und globalen Variablen“ auf Seite 151.

Arbeiten mit Zeichenfolgen

Für Zeichenfolgen stehen eine Reihe von Operationen zur Verfügung, einschließlich:

- Konvertieren einer Zeichenfolge in Groß- oder Kleinbuchstaben - `uppertolower(ZEICHEN)`.
- Entfernen angegebener Zeichen, z. B. `'ID_'` oder `'$'`, aus einer Zeichenfolgevariablen - `stripchar(ZEICHEN,ZEICHENFOLGE)`.
- Bestimmen der Länge (Zeichenanzahl) einer Zeichenfolgevariablen - `length(ZEICHENFOLGE)`.
- Überprüfen der alphabetischen Sortierung von Zeichenfolgewerten - `alphabefore(ZEICHENFOLGE1, ZEICHENFOLGE2)`.
- Entfernen führender oder folgender Leerzeichen aus Werten - `trim(ZEICHENFOLGE)`, `trim_start(ZEICHENFOLGE)` oder `trimend(ZEICHENFOLGE)`.
- Entfernen der ersten bzw. letzten *n* Zeichen aus einer Zeichenfolge - `startstring(LÄNGE, ZEICHENFOLGE)` oder `endstring(LÄNGE, ZEICHENFOLGE)`. Angenommen, es liegt beispielsweise ein Feld mit der Bezeichnung *Element* vor, in dem ein Produktname und ein vierstelliger ID-Code kombiniert sind (ACME CAMERA-D109). Geben Sie die folgende Formel in einem Ableitungsknoten an, um ein neues Feld zu erstellen, das nur den vierstelligen Code enthält:
`endstring(4, Element)`
- Abgleichen eines bestimmten Musters - `ZEICHENFOLGE matches MUSTER`. Um beispielsweise Personen mit dem Wort "market" an einer beliebigen Stelle in ihrer Stellenbezeichnung auszuwählen, können Sie Folgendes in einem Auswahlknoten angeben:
`job_title matches "*market*"`
- Ersetzen aller Instanzen einer Teilzeichenfolge innerhalb einer Zeichenfolge - `replace(TEILZEICHENFOLGE, NEUETEILZEICHENFOLGE, ZEICHENFOLGE)`. Um vor dem Textmining zum Bei-

spiel alle Instanzen eines nicht unterstützten Zeichens (etwa einen senkrechten Strich [|]) durch ein Semikolon zu ersetzen, verwenden Sie in einem Füllerknoten die Funktion `replace`. Wählen Sie unter **Felder ausfüllen** alle Felder aus, in denen das Zeichen möglicherweise vorkommt. Wählen Sie für die Bedingung **Ersetzen** die Option **Immer** aus und geben Sie unter **Ersetzen durch** folgende Bedingung an:

```
replace('|',';',@FIELD)
```

- Ableiten eines Flagfelds basierend auf Vorhandensein einer bestimmten Teilzeichenfolge. Sie können beispielsweise eine Zeichenfolgenfunktion in einem Ableitungsknoten verwenden, um ein gesondertes Flagfeld für jede Antwort mit einem Ausdruck der folgenden Art zu erstellen:

```
hassubstring(museums,"museum_of_design")
```

Weitere Informationen finden Sie in „Zeichenfolgenfunktionen“ auf Seite 172.

Umgang mit Leerzeichen und fehlenden Werten

Zu den gängigen Aufgaben, die Data-Mining-Experten zur Datenvorbereitung ausführen, gehört das Ersetzen von Leerzeichen oder fehlenden Werten. CLEM bietet eine Reihe von Tools zur Automatisierung der Handhabung von Leerzeichen. Mit Leerzeichen wird vor allem im Füllerknoten gearbeitet. Die folgenden Funktionen können jedoch mit jedem Knoten verwendet werden, der CLEM-Ausdrücke akzeptiert:

- Mit `@BLANK(FIELD)` können Datensätze bestimmt werden, deren Werte für ein bestimmtes Feld leer sind, z. B. *Alter*.
- Mit `@NULL(FIELD)` können Datensätze bestimmt werden, deren Werte für das bzw. die angegebenen Felder systemdefiniert fehlen. In IBM SPSS Modeler werden systemdefiniert fehlende Werte als **\$null\$**-Werte angezeigt.

Weitere Informationen finden Sie in „Funktionen zum Umgang mit Leerstellen und Nullwerten“ auf Seite 188.

Arbeiten mit Zahlen

In IBM SPSS Modeler stehen zahlreiche Standardoperationen für numerische Werte zur Verfügung. Dazu gehören:

- Berechnen des Sinus des angegebenen Winkels - `sin(NUM)`
- Berechnen des natürlichen Logarithmus numerischer Felder - `log(NUM)`
- Berechnen der Summe zweier Zahlen - `NUM1 + NUM2`

Weitere Informationen finden Sie in „Numerische Funktionen“ auf Seite 167.

Arbeiten mit Zeit- und Datumsangaben

Die Zeit- und Datumsformate können je nach Datenquelle und Ländereinstellungen variieren. Die Formate für Datum und Uhrzeit sind für jeden Stream spezifisch und werden im Dialogfeld "Streameigenschaften" festgelegt. Im Folgenden sind Beispiele für häufig verwendete Funktionen für die Arbeit mit Datums- und Zeitfeldern aufgeführt.

Berechnen der verstrichenen Zeit

Mit einer Gruppe von Funktionen wie der folgenden können Sie mühelos die Zeit berechnen, die seit einem Basisdatum verstrichen ist. Diese Funktion gibt die Zeit in Monaten und als reelle Zahl zurück, die vom Basisdatum bis zu dem Datum verstrichen ist, das durch die Datumszeichenfolge `DATUM` repräsentiert wird. Dabei handelt es sich um einen Schätzwert, der auf einem Monat mit 30,0 Tagen basiert.

```
date_in_months(Datum)
```

Vergleichen von Datums-/Uhrzeitwerten

Die Werte in Datums- und Uhrzeitfeldern können in den Datensätzen anhand von Funktionen wie der folgenden verglichen werden. Diese Funktion gibt den Wert *Wahr* zurück, wenn die Datumszeichenfolge DATUM1 ein Datum darstellt, das vor dem durch die Datumszeichenfolge DATUM2 repräsentierten Datum liegt. Andernfalls ergibt die Funktion den Wert 0.

```
date_before(Datum1, Datum2)
```

Berechnen von Differenzen

Sie können auch die Differenz zwischen zwei Uhrzeiten und zwei Datumsangaben berechnen. Verwenden Sie dazu Funktionen wie:

```
date_weeks_difference(Datum1, Datum2)
```

Diese Funktion gibt die Zeit in Wochen ab dem durch die Datumszeichenfolge DATUM1 repräsentierten Datum bis zu dem durch die Datumszeichenfolge DATUM2 angegebenen Datum als reelle Zahl zurück. Der Wert basiert auf einer Woche mit 7,0 Tagen. Wenn DATUM2 vor DATUM1 liegt, wird eine negative Zahl zurückgegeben.

Heutiges Datum

Das aktuelle Datum kann mit der Funktion @TODAY zum Dataset hinzugefügt werden. Das heutige Datum wird als Zeichenfolge zum angegebenen oder neuen Feld im Datumsformat hinzugefügt, das im Dialogfeld "Streameigenschaften" ausgewählt wurde. Weitere Informationen finden Sie in „Datums- und Uhrzeitfunktionen“ auf Seite 177.

Zusammenfassen mehrerer Felder

Die CLEM-Sprache beinhaltet eine Reihe von Funktionen, mit denen Übersichtsstatistiken ausgegeben werden, die mehrere Felder umfassen. Diese Funktionen sind besonders nützlich bei der Analyse von Umfragedaten, bei denen mehrere Antworten auf eine Frage in mehreren Feldern gespeichert werden. Weitere Informationen finden Sie in „Arbeiten mit Mehrfachantwortdaten“ auf Seite 146.

Vergleichsfunktionen

Mit den Funktionen min_n und max_n können Sie Werte in mehreren Feldern vergleichen. Beispiel:

```
max_n(['card1fee' 'card2fee' 'card3fee' 'card4fee'])
```

Außerdem können Sie eine Reihe von Zählfunktionen verwenden, um die Anzahl der Werte zu erhalten, die bestimmte Kriterien erfüllen, auch wenn diese Werte in mehreren Feldern gespeichert werden. So können Sie beispielsweise die Anzahl der Karten zählen, die mehr als 5 Jahre im Besitz des Kunden sind:

```
count_greater_than(5, ['cardtenure' 'card2tenure' 'card3tenure'])
```

So können Sie Nullwerte in derselben Gruppe von Feldern zählen:

```
count_nulls(['cardtenure' 'card2tenure' 'card3tenure'])
```

Beachten Sie, dass in diesem Beispiel die Anzahl der Karten im Besitz gezählt werden, nicht die Anzahl der Personen, die sie besitzen. Weitere Informationen finden Sie in „Vergleichsfunktionen“ auf Seite 164.

Mit der Funktion count_equal können Sie zählen, wie oft ein angegebener Wert felderübergreifend vorkommt. Im folgenden Beispiel wird die Anzahl der Felder in der Liste gezählt, die den Wert Y enthalten.

```
count_equal("J",[Antwort1, Antwort2, Antwort3])
```

Wenn die folgenden Werte für die Felder in der Liste gegeben sind, gibt die Funktion die Ergebnisse für den Wert J wie angezeigt aus.

Tabelle 8. Funktionswerte

Antwort1	Antwort2	Antwort3	Anzahl
J	N	J	2
J	N	N	1

Numerische Funktionen

Mithilfe der Funktionen `sum_n`, `mean_n` und `sdev_n` können Sie Statistiken für mehrere Felder erhalten. Beispiele:

```
sum_n(['card1bal' 'card2bal''card3bal'])  
mean_n(['card1bal' 'card2bal''card3bal'])
```

Weitere Informationen finden Sie in „Numerische Funktionen“ auf Seite 167.

Generieren von Feldlisten

Wenn Sie eine der Funktionen verwenden, die eine Liste von Feldern als Eingabe akzeptieren, können die Sonderfunktionen `@FIELDS_BETWEEN(start, ende)` und `@FIELDS_MATCHING(muster)` als Eingabe verwendet werden. Setzt man z. B. voraus, dass die Reihenfolge der Felder derjenigen im vorhergehenden Beispiel für `sum_n` entspricht, wäre Folgendes äquivalent:

```
sum_n(@FIELDS_BETWEEN(card1bal, card3bal))
```

Alternativ können Sie wie folgt die Anzahl der Nullwerte in allen Feldern zählen, die mit "card" beginnen:

```
count_nulls(@FIELDS_MATCHING('card*'))
```

Weitere Informationen finden Sie in „Sonderfelder“ auf Seite 189.

Arbeiten mit Mehrfachantwortdaten

Eine Reihe von Vergleichsfunktionen kann zum Analysieren von Mehrfachantwortdaten herangezogen werden. Zu diesen zählen:

- `value_at`
- `first_index` / `last_index`
- `first_non_null` / `last_non_null`
- `first_non_null_index` / `last_non_null_index`
- `min_index` / `max_index`

Angenommen, in einer Frage mit Mehrfachantworten wird nach den drei wichtigsten Gründen bei der Entscheidung für einen bestimmten Kauf gefragt (z. B. Preis, persönliche Empfehlung, Bewertung, lokaler Lieferant, Sonstiges). In diesem Fall könnten Sie die Wichtigkeit des Preises ermitteln, indem Sie den Index des Felds ableiten, in dem dieser zum ersten Mal verwendet wurde:

```
first_index("Preis", [Grund1 Grund2 Grund3])
```

Nehmen wir analog dazu an, Sie haben Kunden gebeten, drei Autos nach ihrer Kaufwahrscheinlichkeit einzustufen, und die Antworten wie folgt in drei separate Felder codiert:

Tabelle 9. Autobewertungsbeispiele

Kunden-ID	Auto1	Auto2	Auto3
101	1	3	2
102	3	2	1
103	2	3	1

In diesem Fall könnten Sie den Index der Felder für die beliebtesten Autos (an Position 1 bzw. auf dem niedrigsten Rang eingestuft) mithilfe der `min_index`-Funktion ermitteln:

```
min_index(['Auto1' 'Auto2' 'Auto3'])
```

Weitere Informationen finden Sie in „Vergleichsfunktionen“ auf Seite 164.

Referenzieren von Mehrfachantwortsets

Die spezielle `@MULTI_RESPONSE_SET`-Funktion kann dazu verwendet werden, alle Felder in einem Mehrfachantwortset zu referenzieren. Wenn die drei Felder *Auto* aus dem vorherigen Beispiel etwa in einem Mehrfachantwortset mit der Bezeichnung *Autobewertungen* enthalten sind, wird mit folgender Formel das gleiche Ergebnis erzielt:

```
max_index(@MULTI_RESPONSE_SET("Autobewertungen"))
```

Expression Builder

Sie können CLEM-Ausdrücke manuell eingeben oder Expression Builder verwenden. In dieser Funktion wird eine vollständige Liste der CLEM-Funktionen und Operatoren sowie Datenfelder aus dem aktuellen Stream angezeigt, sodass Sie schnell Ausdrücke erstellen können, ohne sich die genauen Namen der Felder bzw. Funktionen merken zu müssen. Außerdem fügt der Builder automatisch die entsprechenden Anführungszeichen für Felder und Werte hinzu, sodass es einfacher wird, syntaktisch korrekte Ausdrücke zu erstellen.

Anmerkung: Expression Builder wird in Script- oder Parametereinstellungen nicht unterstützt.

Anmerkung: Wenn Sie Ihre Datenquelle ändern wollen, sollten Sie vorher prüfen, ob Expression Builder immer noch die von Ihnen ausgewählten Funktionen unterstützt. Da nicht alle Datenbanken alle Funktionen unterstützen, tritt möglicherweise ein Fehler auf, wenn Sie eine neue Datenquelle verwenden.

Zugriff auf Expression Builder

Expression Builder ist in allen Knoten verfügbar, in denen ein CLEM-Ausdruck geschrieben werden muss, einschließlich der Auswahl-, Balancierungs-, Ableitungs-, Füller-, Analyse-, Berichts- und Tabellenknoten. Sie können ihn durch Klicken auf die Schaltfläche für den Taschenrechner direkt rechts neben dem Formelfeld öffnen.

Erstellen von Ausdrücken

Expression Builder bietet nicht nur vollständige Listen mit Feldern, Funktionen und Operatoren, sondern auch Zugriff auf Datenwerte, sofern Ihre Daten instanziiert sind.

So erstellen Sie einen Ausdruck mit Expression Builder:

1. Nehmen Sie Ihre Eingabe im Ausdrucksfeld mithilfe der Funktions- und Feldlisten vor.
oder
2. Wählen Sie die gewünschten Felder und Funktionen aus den Listen aus.
3. Fügen Sie das Feld oder die Funktion durch Doppelklicken oder Klicken auf die gelbe Pfeilschaltfläche zum Ausdrucksfeld hinzu.

4. Fügen Sie die Operationen anhand der Operandenschaltflächen in der Mitte des Dialogfelds in den Ausdruck ein.

Auswählen von Funktionen

In der Funktionsliste werden alle verfügbaren CLEM-Funktionen und -Operatoren angezeigt. Blättern Sie, um eine Funktion aus der Liste auszuwählen, oder zeigen Sie mithilfe der Dropdown-Liste ein Subset von Funktionen oder Operatoren an, um die Suche zu erleichtern. Die verfügbaren Funktionen sind in Kategorien zusammengefasst, um die Suche zu erleichtern.

Die meisten dieser Kategorien sind im Referenzabschnitt der CLEM-Sprachbeschreibung zu finden. Weitere Informationen finden Sie in „Funktionsreferenz“ auf Seite 160.

Die anderen Kategorien lauten wie folgt.

- **Allgemeine Funktionen** enthält eine Auswahl von einigen der am häufigsten verwendeten Funktionen.
- **Zuletzt verwendet** enthält eine Liste der CLEM-Funktionen, die in der aktuellen Sitzung verwendet wurden.
- **@ Functions** enthält eine Liste aller Sonderfunktionen, deren Namen das Zeichen "@" vorangestellt ist.

Anmerkung: Die Funktionen @DIFF1(FELD1,FELD2) und @DIFF2(FELD1,FELD2) erfordern, dass die beiden Feldtypen identisch sind (z. B. beide "Ganzzahl" oder beide "Lang" oder beide "Reelle Zahl").

- **Datenbankfunktionen.** Wenn der Stream eine Datenbankverbindung beinhaltet (mithilfe eines Datenbankquellenknotens) listet diese Auswahl die Funktionen auf, die in dieser Datenbank zur Verfügung stehen, einschließlich benutzerdefinierter Funktionen (UDFs). Weitere Informationen finden Sie in „Datenbankfunktionen“.
- **Datenbankaggregate.** Wenn der Stream eine Datenbankverbindung beinhaltet (mithilfe eines Datenbankquellenknotens) listet diese Auswahl die Aggregationsoptionen auf, die in dieser Datenbank zur Verfügung stehen. Diese Optionen sind in Expression Builder im Aggregatknoten verfügbar.
- **Datenbankfensteraggregate.** Wenn der Stream eine Datenbankverbindung beinhaltet (mithilfe eines Datenbankquellenknotens) listet diese Auswahl die Fensteraggregationsoptionen auf, die Sie in dieser Datenbank verwenden können. Diese Optionen sind nur in Expression Builder in Knoten in der Palette **Feldoperationen** verfügbar.

Anmerkung: Da SPSS Modeler die **Fensteraggregatfunktionen** in der Datenbanksystemansicht abrufen, hängen die verfügbaren Optionen vom Datenbankverhalten ab.

Obwohl diese Optionen "Aggregationsoptionen" genannt werden, sind sie nicht für die Verwendung im Aggregatknoten gedacht, sondern vielmehr für Knoten wie Ableitungs- oder Auswahlknoten anwendbar. Dies liegt daran, dass ihre Ausgabe skalar und kein echtes Aggregat ist. Das bedeutet, sie reduzieren die in der Ausgabe angezeigte Datenmenge nicht so wie der Aggregatknoten. Sie können beispielsweise diese Art der Aggregation verwenden, um einen gleitenden Durchschnitt für Datenzeilen bereitzustellen, z. B. "Durchschnitt der aktuellen Zeile und aller vorherigen Zeilen".

- **Integrierte Aggregate.** Enthält eine Liste der möglichen Aggregationsmodi, die verwendet werden können.
- **Operatoren** listet alle Operatoren auf, die Sie beim Erstellen von Ausdrücken verwenden können. Operatoren stehen auch über die Schaltflächen in der Mitte des Dialogfelds zur Verfügung.
- **Alle Funktionen** enthält eine vollständige Liste der verfügbaren CLEM-Funktionen.

Nachdem Sie eine Gruppe von Funktionen ausgewählt haben, fügen Sie die Funktionen durch Doppelklicken in das Ausdrucksfeld an der Stelle ein, an der sich der Cursor befindet.

Datenbankfunktionen

Datenbankfunktionen können an vielen verschiedenen Speicherorten aufgelistet werden. Die folgende Tabelle enthält die Speicherorte, die SPSS Modeler bei der Suche nach Funktionsdetails durchsucht. Mithilfe

dieser Tabelle können Datenbankadministratoren sicherstellen, dass Benutzer Zugriffsberechtigungen für die erforderlichen Bereiche haben, damit sie die verschiedenen Funktionen verwenden können.

Außerdem listet die Tabelle die Bedingungen auf, die verwendet werden, um auf der Basis des Datenbank- und Funktionstyps zu filtern, wann eine Funktion für die Verwendung verfügbar ist.

Anmerkung: Wenn Sie Datenbankfunktionen aus Amazon Redshift verwenden, muss Ihr Datenbankadministrator Ihnen möglicherweise die Berechtigungen für die folgenden sechs Datenbankobjekte erteilen. Die ersten vier sind Systemkatalogtabellen, die letzten beiden sind Schemas.

- pg_type
- pg_proc
- pg_namespace
- pg_aggregate
- information_schema
- pg_catalog

Tabelle 10. Datenbankfunktionen in Expression Builder

Datenbank	Funktionstyp	Speicherort der Funktionen	Bedingungen für das Filtern von Funktionen
Db2 LUW	UDF	SYSCAT.ROUTINES SYSCAT.ROUTINEPARMS	ROUTINETYPE ist F und FUNCTIONTYPE ist S
Db2 LUW	UDA	SYSCAT.ROUTINES SYSCAT.ROUTINEPARMS	ROUTINETYPE ist F und FUNCTIONTYPE ist C
Db2 iSeries	UDF	QSYS2.SYSROUTINES QSYS2.SYSPARMS	ROUTINE_TYPE ist F und FUNCTION_TYPE ist S
Db2 iSeries	UDA	QSYS2.SYSROUTINES QSYS2.SYSPARMS	ROUTINE_TYPE ist F und FUNCTION_TYPE ist C
Db2 z/OS	UDF	SYSIBM.SYSROUTINES SYSIBM.SYSPARMS	ROUTINETYPE ist F und FUNCTIONTYPE ist S
Db2 z/OS	UDA	SYSIBM.SYSROUTINES SYSIBM.SYSPARMS	ROUTINETYPE ist F und FUNCTIONTYPE ist C
SQL Server	UDF	SYS.ALL_OBJECTS SYS.ALL_PARAMETERS SYS.TYPES	TYPE ist entweder FN oder FS
SQL Server	UDA	SYS.ALL_OBJECTS SYS.ALL_PARAMETERS SYS.TYPES	TYPE ist AF
Oracle	UDF	ALL_ARGUMENTS ALL_PROCEDURES	Alle der folgenden Bedingungen werden erfüllt: <ul style="list-style-type: none"> • OBJECT_TYPE ist FUNCTION • AGGREGATE ist NO • PLS_TYPE ist nicht NULL
Oracle	UDA	ALL_ARGUMENTS ALL_PROCEDURES	Alle der folgenden Bedingungen werden erfüllt: <ul style="list-style-type: none"> • ARGUMENT_NAME ist NULL • AGGREGATE ist YES • PLS_TYPE ist nicht NULL
Teradata	UDF	DBC.FUNCTIONS DBC.ALLRIGHTS	Alle der folgenden Bedingungen werden erfüllt: <ul style="list-style-type: none"> • FUNCTIONTYPE ist F • COLUMNNAME ist RETURN0 • SPPARAMETERTYPE ist 0 • ACCESSRIGHT ist EF

Tabelle 10. Datenbankfunktionen in Expression Builder (Forts.)

Datenbank	Funktionstyp	Speicherort der Funktionen	Bedingungen für das Filtern von Funktionen
Teradata	UDA	DBC.FUNCTIONS DBC.ALLRIGHTS	Alle der folgenden Bedingungen werden erfüllt: <ul style="list-style-type: none"> • FUNCTIONTYPE ist A • COLUMNNAME ist RETURN0 • SPPARAMETERTYPE ist 0 • ACCESSRIGHT ist EF
Netezza	UDF	####.._V_FUNCTION NZA.._V_FUNCTION INZA.._V_FUNCTION	Für ####.._V_FUNCTION gelten die folgenden Bedingungen: <ul style="list-style-type: none"> • RESULT enthält keine Zeichenfolge mit Werten wie TABLE% • FUNCTION enthält keine Zeichenfolge mit Werten wie '/_%' escape '/' • VARARGS ist FALSE <p>Für sowohl NZA.._V_FUNCTION als auch INZA.._V_FUNCTION gelten die folgenden Bedingungen:</p> <ul style="list-style-type: none"> • RESULT enthält keine Zeichenfolge mit Werten wie TABLE% • FUNCTION enthält keine Zeichenfolge mit Werten wie '/_%' escape '/' • BUILTIN ist f • VARARGS ist FALSE
Netezza	UDA	####.._V_AGGREGATE NZA.._V_FUNCTION INZA.._V_FUNCTION	Die beiden folgenden Bedingungen werden erfüllt: <ul style="list-style-type: none"> • AGGTYPE ist ANY oder GROUPED • VARARGS ist FALSE
Netezza	WUDA	####.._V_AGGREGATE NZA.._V_FUNCTION INZA.._V_FUNCTION	Für ####.._V_AGGREGATE gelten die folgenden Bedingungen: <ul style="list-style-type: none"> • AGGTYPE ist ANY oder ANALYTIC • AGGREGATE ist nicht MAX_LABEL • VARARGS ist FALSE <p>Für sowohl NZA.._V_FUNCTION als auch INZA.._V_FUNCTION gelten die folgenden Bedingungen:</p> <ul style="list-style-type: none"> • AGGTYPE ist ANY oder ANALYTIC • BUILTIN ist f • VARARGS ist FALSE

Schlüssel für Begriffe in der Tabelle

- UDF - benutzerdefinierte Funktion (User Defined Function)
- UDA - benutzerdefinierte Aggregierungsfunktion (User Defined Aggregate Function)
- WUDA - benutzerdefinierte Fensteraggregatfunktion (User Defined Window Aggregate Function)
- #### - die Datenbank, mit der Sie zurzeit verbunden sind

Auswählen von Feldern, Parametern und globalen Variablen

In der Feldliste werden alle Felder angezeigt, die an diesem Punkt im Datenstream verfügbar sind. Blättern Sie zu einem Feld, um es aus der Liste auszuwählen. Fügen Sie durch Doppelklicken oder Klicken der gelben Pfeilschaltfläche ein Feld zum Ausdruck hinzu.

Weitere Informationen finden Sie in „Stream-, Sitzungs- und Superknotenparameter“ auf Seite 142.

Neben Feldern können Sie darüber hinaus aus den folgenden Elementen auswählen:

Mehrfachantwortsets. Weitere Informationen finden Sie im Handbuch *IBM SPSS Modeler Quellen-, Prozess- und Ausgabeknoten*.

Zuletzt verwendet enthält eine Liste der Felder, Mehrfachantwortsets, Parameter und globalen Werte, die in der aktuellen Sitzung verwendet wurden.

Parameter. Weitere Informationen finden Sie in „Stream-, Sitzungs- und Superknotenparameter“ auf Seite 142.

Globale Werte. Weitere Informationen finden Sie im Handbuch *IBM SPSS Modeler Quellen-, Prozess- und Ausgabeknoten*.

Anzeigen oder Auswählen von Werten

Feldwerte können an verschiedenen Stellen im System angezeigt werden. Dazu gehören Expression Builder, Data Audit-Berichte und beim Bearbeiten zukünftiger Werte in einem Zeitintervallknoten. Beachten Sie, dass die Daten vollständig in einem Quellen- oder Typknoten instanziiert sein müssen (Speichertypen, Typen und Werte sind bekannt), um diese Funktion verwenden zu können.

Wählen Sie das gewünschte Feld aus und klicken Sie auf die Schaltfläche zur Wertauswahl, um Werte für ein Feld über Expression Builder oder einen Zeitintervallknoten anzuzeigen. Es wird ein Dialogfeld mit Werten für das ausgewählte Feld geöffnet. Sie können dann einen Wert auswählen und auf **Einfügen** klicken, um den Wert in den aktuellen Ausdruck oder die aktuelle Liste einzufügen.



Abbildung 16. Schaltfläche zur Wertauswahl

Bei Flagfeldern und nominalen Feldern werden alle definierten Werte aufgeführt. Bei stetigen Feldern (numerischer Bereich) werden die Minimal- und Maximalwerte angezeigt.

Überprüfen von CLEM-Ausdrücken

Klicken Sie in Expression Builder auf **Überprüfen** (unten rechts), um den Ausdruck zu validieren. Ausdrücke, die nicht geprüft wurden, werden rot angezeigt. Bei Fehlern gibt eine Nachricht die Fehlerursache an.

Folgende Elemente werden überprüft:

- Richtige Anführungszeichen für Werte und Feldnamen
- Richtige Verwendung von Parametern und globalen Variablen
- Zulässige Verwendung von Operatoren
- Vorhandensein von referenzierten Feldern
- Vorhandensein und Definition von referenzierten Globalwerten

Wenn Sie Fehler in der Syntax bemerken, versuchen Sie, den Ausdruck mithilfe der Listen- und Operator-schaltflächen zu erstellen, anstatt den Ausdruck manuell einzugeben. Mit dieser Methode werden automatisch die richtigen Anführungszeichen für Felder und Werte hinzugefügt.

Die folgenden Einschränkungen sind beim Erstellen von Ausdrücken in IBM Analytical Decision Management zu beachten. Ausdrücke können keines der folgenden Elemente enthalten:

- Eine Referenz auf einen IBM SPSS Modeler-Streamparameter
- Eine Referenz auf einen IBM SPSS Modeler-Streamglobalwert
- Eine Referenz auf eine Datenbankfunktion
- Eine Referenz auf eine der folgenden speziellen @-Feldfunktionen oder -Feldwertfunktionen:
 - @TARGET
 - @PREDICTED
 - @FIELD
 - @PARTITION_FIELD
 - @TRAINING_PARTITION
 - @TESTING_PARTITION
 - @VALIDATION_PARTITION

Suchen und Ersetzen

Das Dialogfeld "Suchen/Ersetzen" ist in Situationen verfügbar, in denen Sie Script- oder Ausdruckstext bearbeiten, u. a. im Script-Editor, im CLEM Expression Builder und bei der Definition von Vorlagen im Berichtsknoten. Während der Bearbeitung von Text in einem dieser Bereiche können Sie durch Drücken von Strg+F das Dialogfeld aufrufen. Achten Sie dabei darauf, dass sich der Cursor über einem Textbereich befindet. So können Sie beispielsweise bei der Arbeit in einem Füllerknoten das Dialogfeld aus einem der Textbereiche auf der Registerkarte "Einstellungen" aufrufen oder über das Textfeld im Expression Builder.

1. Achten Sie darauf, dass sich der Cursor in einem Textfeld befindet, und drücken Sie Strg+F, um das Dialogfeld "Suchen/Ersetzen" aufzurufen.
2. Geben Sie den Text ein, nach dem Sie suchen möchten, oder treffen Sie eine Auswahl aus der Dropdown-Liste der kürzlich durchsuchten Elemente.
3. Geben Sie, falls erforderlich, den Ersatztext ein.
4. Klicken Sie auf **Weitersuchen**, um die Suche zu starten.
5. Klicken Sie auf **Ersetzen**, um die aktuelle Auswahl zu ersetzen, oder **Alle ersetzen**, um alle bzw. die ausgewählten Instanzen zu ersetzen.
6. Das Dialogfeld wird nach jedem Vorgang geschlossen. Drücken Sie in einem Textbereich die Taste F3, um den letzten Suchvorgang zu wiederholen, bzw. drücken Sie Strg+F, um erneut auf das Dialogfeld zuzugreifen.

Suchoptionen

Groß-/Kleinschreibung beachten. Gibt an, ob beim Suchvorgang die Groß- und Kleinschreibung berücksichtigt wird; beispielsweise, ob *myvar* als identisch mit *myVar* betrachtet wird. Ersetzungstext wird unabhängig von dieser Einstellung stets genau so eingefügt, wie er eingegeben wurde.

Nur ganze Wörter. Gibt an, ob beim Suchvorgang auch Wortteile gefunden werden. Wenn diese Option ausgewählt ist, werden bei einer Suche nach *Tag* Terme wie *Tagesordnung* oder *Tag-und-Nacht-Gleiche* nicht als Treffer ausgegeben.

Reguläre Ausdrücke. Gibt an, ob die Syntax für reguläre Ausdrücke verwendet werden soll (siehe nächsten Abschnitt). Bei Auswahl dieser Option wird die Option **Nur ganze Wörter** inaktiviert und ihr Wert ignoriert.

Nur **ausgewählter Text**. Legt den Suchumfang bei Verwendung der Option **Alle ersetzen** fest.

Syntax für reguläre Ausdrücke

Mithilfe von regulären Ausdrücken können Sie nach Sonderzeichen (z. B. Tabulatoren oder Zeilenumbrüche), Zeichenklassen bzw. -bereichen, wie *a* bis *d*, beliebige Ziffer oder Nichtziffer, und nach Begrenzungen, wie beispielsweise Zeilenanfang bzw. Zeilenende, suchen. Folgende Arten von Ausdrücken werden unterstützt:

Tabelle 11. Zeichenübereinstimmungen

Zeichen	Übereinstimmung/Bedeutung
x	Das Zeichen x
\\	Umgekehrter Schrägstrich
\0n	Das Zeichen mit dem Oktalwert 0n (0 <= n <= 7)
\0nn	Das Zeichen mit dem Oktalwert 0nn (0 <= n <= 7)
\0mnn	Das Zeichen mit dem Oktalwert 0mnn (0 <= m <= 3; 0 <= n <= 7)
\xhh	Das Zeichen mit dem Hexadezimalwert 0xhh
\uhhhh	Das Zeichen mit dem Hexadezimalwert 0xhhhh
\t	Das Tabulatorzeichen ('\u0009')
\n	Das Zeilenvorschubzeichen ('\u000A')
\r	Das Rücklaufzeichen ('\u000D')
\f	Das Formularvorschubzeichen ('\u000C')
\a	Das Alarmzeichen ('\u0007')
\e	Das Escapezeichen ('\u001B')
\cx	Das Steuerzeichen, das x entspricht

Tabelle 12. Übereinstimmende Zeichenklassen

Zeichenklassen	Übereinstimmung/Bedeutung
[abc]	a, b oder c (einfache Klasse)
[^abc]	Beliebiges Zeichen mit Ausnahme von a, b oder c (Differenzmenge)
[a-zA-Z]	a bis einschließlich z bzw. A bis einschließlich Z (Bereich)
[a-d[m-p]]	a bis d bzw. m bis P (Vereinigungsmenge) Alternative Angabemöglichkeit: [a-dm-p]
[a-z&&[def]]	a bis z und d, e bzw. f (Schnittmenge)
[a-z&&[^bc]]	a bis z mit Ausnahme von b und c (Differenzmenge) Alternative Angabemöglichkeit: [ad-z]
[a-z&&[^m-p]]	a bis z mit Ausnahme von m bis p (Differenzmenge) Alternative Angabemöglichkeit: [a-lq-z]

Tabelle 13. Vordefinierte Zeichenklassen

Vordefinierte Zeichenklassen	Übereinstimmung/Bedeutung
.	Beliebiges Zeichen (evtl. Übereinstimmung mit Zeilenabschlusszeichen)
\d	Beliebige Ziffer: [0-9]
\D	Nichtziffer: [^0-9]
\s	Ein Leerzeichen: [\t\n\x0B\f\r]
\S	Ein Nichtleerzeichen: [^\s]

Tabelle 13. Vordefinierte Zeichenklassen (Forts.)

Vordefinierte Zeichenklassen	Übereinstimmung/Bedeutung
\w	Ein Zeichen: [a-zA-Z_0-9]
\W	Ein Nichtzeichen: [^\w]

Tabelle 14. Übereinstimmungen mit Begrenzungszeichen

Zeichen(folgen) für Begrenzungszeichen	Übereinstimmung/Bedeutung
^	Zeilenanfang
\$	Zeilenende
\b	Wortgrenze
\B	Nichtwortgrenze
\A	Beginn der Eingabe
\Z	Ende der Eingabe mit Ausnahme des letzten Abschlusszeichens, sofern vorhanden
\z	Ende der Eingabe

Kapitel 10. CLEM-Sprachreferenz

CLEM-Referenz - Überblick

In diesem Abschnitt wird CLEM (Control Language for Expression Manipulation) beschrieben, ein leistungsstarkes Tool zur Analyse und Bearbeitung der in IBM SPSS Modeler-Streams verwendeten Daten. CLEM kann in Knoten zur Ausführung verschiedener Aufgaben eingesetzt werden, vom Auswerten von Bedingungen über das Ableiten von Werten bis hin zum Einfügen von Daten in Berichte.

CLEM-Ausdrücke bestehen aus Werten, Feldnamen, Operatoren und Funktionen. Mit der richtigen Syntax können Sie eine Vielzahl leistungsstarker Datenoperationen erstellen.

CLEM-Datentypen

CLEM-Datentypen können aus folgenden Elementen bestehen:

- Ganze Zahlen
- Reelle Zahlen
- Zeichen
- Zeichenfolgen
- Listen
- Felder
- Datum/Uhrzeit

Regeln für die Verwendung von Anführungszeichen

IBM SPSS Modeler ist zwar flexibel bei der Bestimmung der in einem CLEM-Ausdruck verwendeten Felder, Werte, Parameter und Zeichenfolgen. Dennoch ist es ratsam, sich beim Erstellen von Ausdrücken an die unten angegebenen allgemeinen Regeln zu halten.

- Zeichenfolgen - Setzen Sie Zeichenfolgen immer in doppelte Anführungszeichen, beispielsweise "Typ 2". Bei einfachen Anführungszeichen besteht die Gefahr, dass sie mit in Anführungszeichen gesetzten Feldern verwechselt werden.
- Felder - Verwenden Sie einfache Anführungszeichen nur, wenn es erforderlich ist, um Leerzeichen oder andere Sonderzeichen einzuschließen. Beispiel: 'Bestellung Nummer'. Felder, die in Anführungszeichen gesetzt sind, im Dataset jedoch nicht definiert sind, werden irrtümlich als Zeichenfolgen interpretiert.
- Parameter - Setzen Sie Parameter immer in einfache Anführungszeichen, wie beispielsweise '\$P-threshold'.
- Zeichen - Zeichen müssen stets in einfache umgekehrte Anführungszeichen (``) gesetzt werden. Beispiel: stripchar(`d`, "drugA").

Diese Regeln werden in den folgenden Themen genauer behandelt.

Ganze Zahlen

Ganze Zahlen werden als Folge von Dezimalziffern dargestellt. Optional können Sie der ganzen Zahl ein Minuszeichen (-) voranstellen, um eine negative Zahl anzugeben. Beispiele: 1.234, 999, -77.

Die CLEM-Sprache kann mit ganzen Zahlen beliebiger Genauigkeit umgehen. Die maximale Größe der ganzen Zahlen hängt von Ihrer Plattform ab. Wenn die Werte zu groß für die Anzeige in einem Feld ganzer Zahlen sind, kann der Wert normalerweise durch Ändern des Feldtyps in Reelle Zahl wiederhergestellt werden.

Reelle Zahlen

Reelle Zahlen sind Gleitkommazahlen. Reelle Zahlen bestehen aus einer oder mehreren Ziffern, gefolgt von einem Punkt als Dezimaltrennzeichen, gefolgt von mindestens einer weiteren Ziffer. Bei CLEM gilt für die reellen Zahlen doppelte Genauigkeit.

Sie können der ganzen Zahl optional ein Minuszeichen (-) voranstellen, um eine negative Zahl anzugeben — Beispiel: 1,234, 0,999, -77,001. Mit dem Format `<Zahl> e <Exponent>` können Sie eine reelle Zahl in Exponentialnotation angeben. Beispiel `1234,0e5`, `1,7e-2`. Wenn die IBM SPSS Modeler-Anwendung numerische Zeichenfolgen aus Dateien einliest und sie automatisch in Zahlen umwandelt, werden Zahlen ohne führende Ziffer vor dem Dezimalpunkt oder ohne Ziffer nach dem Dezimaltrennzeichen akzeptiert. Beispiel: `999`, oder `,11`. Bei CLEM-Ausdrücken sind diese Formate jedoch nicht zulässig.

Hinweis: Bei der Verwendung reeller Zahlen in CLEM-Ausdrücken muss ein Punkt als Dezimaltrennzeichen verwendet werden, unabhängig von den verwendeten Ländereinstellungen und den Einstellungen für den aktuellen Stream. So müssen Sie beispielsweise angeben:

Na > 0.6

Und nicht:

Na > 0,6

Dies gilt auch dann, wenn im Dialogfeld für die Streameigenschaften das Komma als Dezimaltrennzeichen ausgewählt wurde. Diese Vorgehensweise entspricht der allgemeinen Richtlinie, dass Codesyntax unabhängig von bestimmten Ländereinstellungen bzw. Konventionen sein sollte.

Zeichen

Zeichen (in der Regel als ZEICHEN angezeigt) werden in CLEM-Ausdrücken normalerweise zur Durchführung von Tests an Zeichenfolgen verwendet. Mit der Funktion `isuppercode` können Sie beispielsweise ermitteln, ob das erste Zeichen einer Zeichenfolge ein Großbuchstabe ist. Im folgenden CLEM-Ausdruck wird ein Zeichen verwendet, um anzugeben, dass der Test für das erste Zeichen der Zeichenfolge durchgeführt werden soll.

```
isuppercode(subscrs(1, "MyString"))
```

Um den Code (im Gegensatz zur Position) eines bestimmten Zeichens in einem CLEM-Ausdruck anzugeben, verwenden Sie einzelne umgekehrte Anführungszeichen in der Form ``<Zeichen>``. Beispiel: ``A``, ``Z``.

Hinweis: Bei Feldern gibt es keinen Speichertyp CHAR. Wenn also ein Feld abgeleitet oder mit einem Ausdruck ausgefüllt wird, der zu CHAR führt, wird das Ergebnis in eine Zeichenfolge umgewandelt.

Zeichenfolgen

Im Allgemeinen sollten Zeichenfolgen in doppelte Anführungszeichen gesetzt werden. Beispiele für Zeichenfolgen sind `"c35Produkt2"` und `"ReferenzID"`. Zur Kennzeichnung von Sonderzeichen in Zeichenfolgen werden umgekehrte Schrägstriche verwendet, z. B.: `"\$65443"`. (Zur Kennzeichnung eines umgekehrten Schrägstrichs wird ein doppelter umgekehrter Schrägstrich verwendet: `\\`.) Zeichenfolgen können auch in einfache Anführungszeichen gesetzt werden. Das Ergebnis lässt sich jedoch nicht von einem in Anführungszeichen gesetzten Feld unterscheiden (`'ReferenzID'`). Weitere Informationen finden Sie in „Zeichenfolgefunktionen“ auf Seite 172.

Listen

Eine Liste ist eine geordnete Abfolge von Elementen, die verschiedenen Typen angehören können. Listen werden in eckige Klammern (`[]`) eingeschlossen. Beispiele für Listen sind `[1 2 4 16]` und `["abc" "def"]`. Listen werden nicht als Wert von IBM SPSS Modeler-Feldern verwendet. Sie werden verwendet, um Argumente für Funktionen, beispielsweise `member` und `oneof`, bereitzustellen

Anmerkung: Listen können nur aus statischen Objekten (z. B. einer Zeichenfolge, einer Zahl oder einem Feldnamen) und nicht aus Funktionsaufrufen bestehen.

Felder

Namen in CLEM-Ausdrücken, bei denen es sich nicht um Namen von Funktionen handelt, werden als Feldnamen betrachtet. Namen können einfach als `Energie`, `val27`, `state_flag` usw. geschrieben werden, aber wenn der Name mit einer Ziffer beginnt oder nicht alphabetische Zeichen, beispielsweise Leerzeichen, (Ausnahme: Unterstrich) enthält, muss der Name in einfache Anführungszeichen gesetzt werden. Beispiel: `'Anstieg der Energie'`, `'2. Antwort'`, `'#101'`, `'$P-NextField'`.

Hinweis: Felder, die in Anführungszeichen gesetzt sind, im Dataset jedoch nicht definiert sind, werden irrtümlich als Zeichenfolgen interpretiert.

Datumsangaben

Die Datumsberechnungen beruhen auf einem "Basisdatum", das im Dialogfeld "Streameigenschaften" angegeben wird. Standardmäßig wird als Basisdatum der 1. Januar 1900 verwendet.

Die CLEM-Sprache unterstützt folgende Datumsformate.

Tabelle 15. Datumsformate der CLEM-Sprache

Format	Beispiele
TTMMJJ	150163
MMTTJJ	011563
JJMMTT	630115
JJJMMTT	19630115
JJJJTTT	Vierstellige Jahreszahl, gefolgt von einer dreistelligen Zahl für den Tag im Jahr. So steht 2000032 für den 32. Tag des Jahres 2000, also für den 1. Februar 2000.
DAY	Tag der Woche in der aktuellen Ländereinstellung - z. B. Montag, Dienstag, ..., für Deutsch.
MONTH	Monat in der aktuellen Ländereinstellung - z. B. Januar, Februar,
TT/MM/JJ	15/01/63
TT/MM/JJJJ	15/01/1963
MM/TT/JJ	01/15/63
MM/TT/JJJJ	01/15/1963
TT-MM-JJ	15-01-63
TT-MM-JJJJ	15-01-1963
MM-TT-JJ	01-15-63
MM-TT-JJJJ	01-15-1963
TT.MM.JJ	15.01.63
TT.MM.JJJJ	15.01.1963
MM.TT.JJ	01.15.63
MM.TT.JJJJ	01.15.1963
TT-MON-JJ	15-JAN-63, 15-jan-63, 15-Jan-63
TT/MON/JJ	15/JAN/63, 15/jan/63, 15/Jan/63
TT.MON.JJ	15.JAN.63, 15.jan.63, 15.Jan.63

Tabelle 15. Datumsformate der CLEM-Sprache (Forts.)

Format	Beispiele
TT-MON-JJJJ	15-JAN-1963, 15-jan-1963, 15-Jan-1963
TT/MON/JJJJ	15/JAN/1963, 15/jan/1963, 15/Jan/1963
TT.MON.JJJJ	15.JAN.1963, 15.jan.1963, 15.Jan.1963
MON JJJJ	Jan 2004
q Q JJJJ	Datum als Ziffer (1-4) für das Quartal, gefolgt vom Buchstaben Q und einer vierstelligen Jahresangabe. So wird beispielsweise das Datum 25. Dez. 2004 als 4 Q 2004 dargestellt.
ww WK JJJJ	Zweistellige Zahl für die Kalenderwoche, gefolgt von den Buchstaben WK und dann von einer vierstelligen Jahreszahl. Bei der Berechnung der Kalenderwoche wird davon ausgegangen, dass Montag der erste Wochentag ist, und dass mindestens ein Tag in der ersten Kalenderwoche liegt.

Zeit

Die CLEM-Sprache unterstützt folgende Zeitformate.

Tabelle 16. Zeitformate der CLEM-Sprache

Format	Beispiele
HHMMSS	120112, 010101, 221212
HHMM	1223, 0745, 2207
MMSS	5558, 0100
HH:MM:SS	12:01:12, 01:01:01, 22:12:12
HH:MM	12:23, 07:45, 22:07
MM:SS	55:58, 01:00
(H)H:(M)M:(S)S	12:1:12, 1:1:1, 22:12:12
(H)H:(M)M	12:23, 7:45, 22:7
(M)M:(S)S	55:58, 1:0
HH.MM.SS	12.01.12, 01.01.01, 22.12.12
HH.MM	12.23, 07.45, 22.07
MM.SS	55.58, 01.00
(H)H.(M)M.(S)S	12.1.12, 1.1.1, 22.12.12
(H)H.(M)M	12.23, 7.45, 22.7
(M)M.(S)S	55.58, 1.0

CLEM-Operatoren

Folgende Operatoren stehen zur Verfügung.

Tabelle 17. CLEM-Sprachoperatoren

Operation	Kommentare	Rangfolge (siehe nächsten Abschnitt)
or	Wird zwischen zwei CLEM-Ausdrücken verwendet. Ergibt den Wert "Wahr", wenn mindestens einer der beiden Terme wahr ist.	10

Tabelle 17. CLEM-Sprachoperatoren (Forts.)

Operation	Kommentare	Rangfolge (siehe nächsten Abschnitt)
and	Wird zwischen zwei CLEM-Ausdrücken verwendet. Ergibt den Wert "Wahr", wenn beide Terme "wahr" sind.	9
=	Wird zwischen zwei beliebigen miteinander vergleichbaren Elementen verwendet. Ergibt den Wert "Wahr", wenn ELEMENT1 gleich ELEMENT2 ist.	7
==	Identisch mit =.	7
/=	Wird zwischen zwei beliebigen miteinander vergleichbaren Elementen verwendet. Ergibt den Wert "Wahr", wenn ELEMENT1 <i>nicht</i> gleich ELEMENT2 ist.	7
/==	Identisch mit /=.	7
>	Wird zwischen zwei beliebigen miteinander vergleichbaren Elementen verwendet. Ergibt den Wert "Wahr", wenn ELEMENT1 größer als ELEMENT2 ist.	6
>=	Wird zwischen zwei beliebigen miteinander vergleichbaren Elementen verwendet. Ergibt den Wert "Wahr", wenn ELEMENT1 größer-gleich ELEMENT2 ist.	6
<	Wird zwischen zwei beliebigen miteinander vergleichbaren Elementen verwendet. Ergibt den Wert "Wahr", wenn ELEMENT1 kleiner als ELEMENT2 ist.	6
<=	Wird zwischen zwei beliebigen miteinander vergleichbaren Elementen verwendet. Ergibt den Wert "Wahr", wenn ELEMENT1 kleiner-gleich ELEMENT2 ist.	6
&&=_0	Wird zwischen zwei ganzen Zahlen verwendet. Entspricht dem booleschen Ausdruck GANZZ1 && GANZZ2 = 0.	6
&&/=_0	Wird zwischen zwei ganzen Zahlen verwendet. Entspricht dem booleschen Ausdruck GANZZ1 && GANZZ2 /= 0.	6
+	Addiert zwei Zahlen: ZAHL1 + ZAHL2.	5
>>	Verkettet zwei Zeichenfolgen; beispielsweise ZEICHENFOLGE1 >> ZEICHENFOLGE2.	5
-	Subtrahiert eine Zahl von einer anderen Zahl: ZAHL1 - ZAHL2. Kann auch vor einer Zahl verwendet werden: - NUM.	5
*	Multipliziert zwei Zahlen mit einander: ZAHL1 * ZAHL2.	4
&&	Wird zwischen zwei ganzen Zahlen verwendet. Ergibt das bitweise "Und" der ganzen Zahlen GANZZ1 und GANZZ2.	4
&&~	Wird zwischen zwei ganzen Zahlen verwendet. Ergibt das bitweise "Und" von GANZZ1 und das bitweise Komplement von GANZZ2.	4

Tabelle 17. CLEM-Sprachoperatoren (Forts.)

Operation	Kommentare	Rangfolge (siehe nächsten Abschnitt)
	Wird zwischen zwei ganzen Zahlen verwendet. Ergibt das bitweise "Oder" von GANZZ1 und GANZZ2.	4
~~	Wird vor einer ganzen Zahl verwendet. Ergibt das bitweise Komplement von GANZZ.	4
/&	Wird zwischen zwei ganzen Zahlen verwendet. Ergibt das bitweise "exklusive Oder" von GANZZ1 und GANZZ2.	4
GANZ1 << N	Wird zwischen zwei ganzen Zahlen verwendet. Ergibt das Bitmuster von GANZZ, um N Positionen nach links verschoben.	4
GANZ1 >> N	Wird zwischen zwei ganzen Zahlen verwendet. Ergibt das Bitmuster von GANZZ, um N Positionen nach rechts verschoben.	4
/	Dividiert eine Zahl durch eine andere: ZAHL1 / ZAHL2.	4
**	Wird zwischen zwei Zahlen verwendet: BASIS ** POTENZ. Ergibt BASIS hoch POTENZ.	3
rem	Wird zwischen zwei ganzen Zahlen verwendet: GANZZ1 rem GANZZ2. Ergibt den Rest: GANZZ1 - (GANZZ1 div GANZZ2) * GANZZ2.	2
div	Wird zwischen zwei ganzen Zahlen verwendet: GANZZ1 div GANZZ2. Führt eine Division zwischen zwei ganzen Zahlen durch.	2

Operatorenrangfolge

Rangfolgen bestimmen die Analyse komplexer Ausdrücke, insbesondere von Ausdrücken ohne Klammern mit mehreren Infix-Operatoren. Beispiel:

3 + 4 * 5

wird als 3 + (4 * 5) und nicht als (3 + 4) * 5 analysiert, da die relative Rangfolge festlegt, dass * vor + verarbeitet werden muss. Jedem Operator in der CLEM-Sprache ist ein Rangfolgewert zugeordnet; je niedriger dieser Wert ist, desto wichtiger ist er auf der Analyseliste, was bedeutet, dass er vor anderen Operatoren mit höheren Rangfolgewerten verarbeitet wird.

Funktionsreferenz

Folgende CLEM-Funktionen stehen bei der Arbeit mit Daten in IBM SPSS Modeler zur Verfügung. Diese Funktionen können in verschiedenen Dialogfeldern (z. B. "Ableitungsknoten" und "Dichotomknoten") als Code eingegeben werden. Alternativ können Sie mit dem Expression Builder gültige CLEM-Ausdrücke erstellen, ohne sich Funktionslisten oder Feldnamen einprägen zu müssen.

Tabelle 18. CLEM-Funktionen für die Verwendung mit IBM SPSS Modeler-Daten

Funktionstyp	Beschreibung
Informationsfunktionen	Dienen zum Gewinnen von Erkenntnissen über Feldwerte. Beispiel: Die Funktion is_string gibt für alle Datensätze vom Typ string den Wert "Wahr" zurück.

Tabelle 18. CLEM-Funktionen für die Verwendung mit IBM SPSS Modeler-Daten (Forts.)

Funktionstyp	Beschreibung
Konvertierungsfunktionen	Dienen zum Erstellen neuer Felder oder zum Konvertieren des Speichertyps. Die Funktion <code>to_timestamp</code> konvertiert das ausgewählte Feld beispielsweise in eine Zeitmarke.
Vergleich	Dienen zum Vergleichen von Feldwerten untereinander oder mit einer angegebenen Zeichenfolge. Beispiel: Mit <code><=</code> verglichen, ob die Werte zweier Felder kleiner oder gleich sind.
Logische Funktionen	Dienen zur Durchführung logischer Operationen wie <code>if</code> (wenn), <code>then</code> (dann), <code>else</code> (sonst).
Numerische Funktionen	Dienen zur Durchführung numerischer Berechnungen, wie dem natürlichen Logarithmus von Feldwerten.
Trigonometrische Funktionen	Dienen zur Durchführung trigonometrischer Berechnungen, wie Berechnung des Arkuskosinus eines bestimmten Winkels.
Wahrscheinlichkeitsfunktionen	Geben Wahrscheinlichkeiten zurück, die auf verschiedenen Verteilungen basieren, z. B. die Wahrscheinlichkeit, dass ein Wert aus der <i>t</i> -Verteilung in der Studentenversion kleiner ist als ein bestimmter Wert.
Räumliche Funktionen	Werden zur Durchführung räumlicher Berechnungen mit georäumlichen Daten verwendet.
Bitweise Operationen	Dienen zur Bearbeitung von ganzen Zahlen als Bitmuster.
Zufallsfunktionen	Dienen zur zufälligen Auswahl von Elementen oder Generierung von Zahlen.
Zeichenfolgefunktionen	Dienen zur Durchführung verschiedener Operationen an Zeichenfolgen, wie beispielsweise <code>stripchar</code> zum Entfernen eines angegebenen Zeichens.
SoundEx-Funktionen	Dienen zum Auffinden von Zeichenfolgen, deren genaue Schreibweise nicht bekannt ist, und zwar auf der Grundlage fonetischer Annahmen, wie bestimmte Buchstaben ausgesprochen werden.
Datums- und Uhrzeitfunktionen	Dienen zur Durchführung verschiedener Operationen von Datums-, Zeit- und Zeitmarkenfeldern.
Sequenzfunktionen	Dienen zur Gewinnung von Erkenntnissen über die Datensatzsequenz eines Datasets oder zur Durchführung von Operationen auf der Basis dieser Sequenz.
Globale Funktionen	Ermöglichen den Zugriff auf Globalwerte, die von einem Globalwerteknoten erstellt wurden. Beispiel: <code>@MEAN</code> wird als Verweis auf den mittleren Durchschnitt aller Werte für ein Feld im gesamten Dataset verwendet.
Leerzeichen und Nullwerte	Dienen zum Zugreifen, Kennzeichnen und häufig zum Ausfüllen benutzerdefinierter Leerstellen oder systemdefiniert fehlender Werte. Beispiel: <code>@BLANK(FELD)</code> wird verwendet, um Datensätze mit Leerstellen als "wahr" zu kennzeichnen.
Sonderfelder	Dienen zur Angabe der gesonderten Felder, die untersucht werden. <code>@FIELD</code> wird beispielsweise beim Ableiten mehrerer Felder verwendet.

Konventionen bei Funktionsbeschreibungen

Folgende Konventionen werden überall in diesem Dokument für Elemente in einer Funktion verwendet.

Tabelle 19. Konventionen bei Funktionsbeschreibungen

Konvention	Beschreibung
<code>BOOL</code>	Ein boolescher Wert bzw. ein boolesches Flag, wie "wahr" oder "falsch".

Tabelle 19. Konventionen bei Funktionsbeschreibungen (Forts.)

Konvention	Beschreibung
ZAHL, ZAHL1, ZAHL2	Eine beliebige Zahl.
REELL, REELL1, REELL2	Eine reelle Zahl, wie 1,234 oder -77,01.
GANZZ, GANZZ1, GANZZ2	Eine ganze Zahl, wie 1 oder -77.
ZEICHEN	Eine Zeichencode, wie beispielsweise `A`.
ZEICHENFOLGE	Ein Zeichenfolge, wie beispielsweise "ReferenzID".
LISTE	Eine Liste mit Elementen, wie beispielsweise ["abc" "def"].
ELEMENT	Ein Feld, wie beispielsweise Kunde oder Konzept_extrahieren.
DATUM	Ein Datumsfeld, beispielsweise Startdatum, mit Werten in Formaten wie TT-MON-JJJJ.
ZEIT	Ein Zeitfeld, beispielsweise Energiefluss, mit Werten in Formaten wie HHMMSS.

Die Funktionen in diesem Handbuch sind wie folgt aufgeführt: Die Funktion in der ersten Spalte, der Ergebnistyp (ganze Zahl, Zeichenfolge usw.) in der zweiten und gegebenenfalls eine Beschreibung in der dritten Spalte. Folgendes Beispiel ist eine Beschreibung der Funktion *rem*.

Tabelle 20. Beschreibung der Funktion "rem"

Funktion	Ergebnis	Beschreibung
GANZZ1 rem GANZZ2	Zahl	Ergibt den Rest des Quotienten aus GANZZ1 und GANZZ2. Beispiel: GANZZ1 - (GANZZ1 div GANZZ2) * GANZZ2.

Einzelheiten zur den Verwendungskonventionen, beispielsweise zur Auflistung von Elementen oder zur Angabe von Zeichen in einer Funktion, werden an anderer Stelle beschrieben. Weitere Informationen finden Sie in „CLEM-Datentypen“ auf Seite 155.

Informationsfunktionen

Informationsfunktionen werden verwendet, um einen Einblick in die Werte eines bestimmten Felds zu gewinnen. Sie werden üblicherweise verwendet, um Flagfelder abzuleiten. Beispiel: Mit der Funktion @BLANK können Sie ein Flagfeld erstellen, das die Datensätze angibt, deren Werte für das ausgewählte Feld leer sind. In ähnlicher Weise können Sie den Speichertyp für ein Feld mit einer der Speichertypfunktionen überprüfen, beispielsweise mit *is_string*.

Tabelle 21. CLEM-Informationsfunktionen.

Funktion	Ergebnis	Beschreibung
@BLANK(FELD)	boolesch	Ergibt den Wert "Wahr" für alle Datensätze, deren Werte gemäß den Regeln zum Umgang mit Leerstellen, die in einem vorgeordneten Typknoten oder Quellenknoten (Registerkarte "Typen") festgelegt wurden, Leerstellen sind.
@NULL(ELEMENT)	boolesch	Ergibt den Wert "Wahr" für alle Datensätze, deren Werte nicht definiert sind. Nicht definierte Werte sind systemdefinierte Nullwerte, die in IBM SPSS Modeler als \$null\$ angezeigt werden.
is_date(ELEMENT)	boolesch	Ergibt den Wert "Wahr" für alle Datensätze, deren Typ ein Datum ist.
is_datetime(ELEMENT)	boolesch	Ergibt den Wert "Wahr" für alle Datensätze, deren Typ "Datum", "Zeit" oder "Zeitmarke" ist.

Tabelle 21. CLEM-Informationsfunktionen (Forts.).

Funktion	Ergebnis	Beschreibung
is_integer(ELEMENT)	boolesch	Ergibt den Wert "Wahr" für alle Datensätze, deren Typ eine ganze Zahl ist.
is_number(ELEMENT)	boolesch	Ergibt den Wert "Wahr" für alle Datensätze, deren Typ eine Zahl ist.
is_real(ELEMENT)	boolesch	Ergibt den Wert "Wahr" für alle Datensätze, deren Typ eine reelle Zahl ist.
is_string(ELEMENT)	boolesch	Ergibt den Wert "Wahr" für alle Datensätze, deren Typ eine Zeichenfolge ist.
is_time(ELEMENT)	boolesch	Ergibt den Wert "Wahr" für alle Datensätze, deren Typ eine Zeitangabe ist.
is_timestamp(ELEMENT)	boolesch	Ergibt den Wert "Wahr" für alle Datensätze, deren Typ eine Zeitmarke ist.

Konvertierungsfunktionen

Mit Konvertierungsfunktionen können Sie neue Felder erstellen und den Speichertyp bestehender Felder umwandeln. Beispielsweise können Sie neue Zeichenfolgen bilden, indem Sie Zeichenfolgen miteinander verbinden oder zerlegen. Zwei Zeichenfolgen können mit dem Operator >< miteinander verbunden werden. Wenn das Feld Standort beispielsweise den Wert "EISENACH" aufweist, ergibt "xx" >< Standort den Wert "xxEISENACH". Das Ergebnis von >< ist immer eine Zeichenfolge, selbst wenn die einzelnen Argumente keine Zeichenfolgen sind. Wenn also Feld V1 den Wert 3 und Feld V2 den Wert 5 aufweist, ergibt V1 >< V2 den Wert "35" (eine Zeichenfolge, keine Zahl).

Konvertierungsfunktionen (und alle anderen Funktionen, für die ein spezieller Eingabetyp, wie beispielsweise ein Wert für Datum oder Uhrzeit, erforderlich ist) hängen von den aktuell im Dialogfeld für die Streamoptionen angegebenen Formaten ab. Wenn Sie beispielsweise ein Zeichenfolgefild mit den Werten *Jan 2003*, *Feb 2003* usw. konvertieren möchten, müssen Sie das zugehörige Datumsformat, **MON JJJJ**, als Standarddatumsformat für den Stream auswählen.

Tabelle 22. CLEM-Konvertierungsfunktionen

Funktion	Ergebnis	Beschreibung
ELEMENT1 >< ELEMENT2	Zeichenfolge	Zieht die Werte der beiden Felder zusammen und ergibt die resultierende Zeichenfolge als <i>ELEMENT1ELEMENT2</i> .
to_integer(ELEMENT)	ganze Zahl	Konvertiert den Speichertyp des angegebenen Felds in eine ganze Zahl.
to_real(ELEMENT)	reelle Zahl	Konvertiert den Speichertyp des angegebenen Felds in eine reelle Zahl.
to_number(ELEMENT)	Zahl	Konvertiert den Speichertyp des angegebenen Felds in eine Zahl.
to_string(ELEMENT)	Zeichenfolge	Konvertiert den Speichertyp des angegebenen Felds in eine Zeichenfolge. Wenn eine reelle Zahl mit dieser Funktion in eine Zeichenfolge konvertiert wird, wird ein Wert mit sechs Ziffern nach dem Radixpunkt zurückgegeben.
to_time(ELEMENT)	Zeit	Konvertiert den Speichertyp des angegebenen Felds in eine Zeit.
to_date(ELEMENT)	Datum	Konvertiert den Speichertyp des angegebenen Felds in ein Datum.
to_timestamp(ELEMENT)	Zeitmarke	Konvertiert den Speichertyp des angegebenen Felds in eine Zeitmarke.

Tabelle 22. CLEM-Konvertierungsfunktionen (Forts.)

Funktion	Ergebnis	Beschreibung
to_datetime(ELEMENT)	Datum/Uhrzeit	Konvertiert den Speichertyp des angegebenen Felds in einen Wert vom Typ "Datum", "Zeit" oder "Zeitmarke".
datetime_date(ELEMENT)	Datum	Gibt den Datumswert für eine <i>Zahl</i> , eine <i>Zeichenfolge</i> oder <i>Zeitmarke</i> zurück. Beachten Sie, dass dies die einzige Funktion ist, mit der Sie eine Zahl (in wenigen Sekunden) wieder in ein Datum zurückkonvertieren können. Wenn es sich bei ELEMENT um eine Zeichenfolge handelt, wird ein Datum erstellt, indem eine Zeichenfolge im aktuellen Datumsformat analysiert wird. Das im Dialogfeld für die Streameigenschaften angegebene Datumsformat muss korrekt sein, damit diese Funktion erfolgreich ausgeführt werden kann. Wenn es sich bei ELEMENT um eine Zahl handelt, wird sie als Anzahl von Sekunden seit dem Basisdatum (Epoche) interpretiert. Die Bruchteile eines Tages werden gekürzt. Wenn es sich bei ELEMENT um eine Zeitmarke handelt, wird der Datumsteil der Zeitmarke ausgegeben. Wenn es sich bei ELEMENT um ein Datum handelt, wird es unverändert ausgegeben.
stb_centroid_latitude(ELEMENT)	ganze Zahl	Gibt einen ganzzahligen Wert für den Breitengrad zurück, der dem Zentroid des Geohash-Arguments entspricht.
stb_centroid_longitude(ELEMENT)	ganze Zahl	Gibt einen ganzzahligen Wert für den Längengrad zurück, der dem Zentroid des Geohash-Arguments entspricht.
to_geohash(ELEMENT)	Zeichenfolge	Gibt die Geohash-Zeichenfolge zurück, die dem Breitengrad und Längengrad bei der angegebenen für die Dichte Anzahl Bit entspricht. Ein Geohash ist ein Code, der zur Angabe einer Gruppe geographischer Koordinaten auf der Basis der Breiten- und Längengraddetails dient. Für to_geohash gibt es die folgenden drei Parameter: <ul style="list-style-type: none"> • <i>latitude</i>: Bereich (-180, 180); Einheiten sind Gradangaben im WGS84-Koordinatensystem. • <i>longitude</i>: Bereich (-90, 90); Einheiten sind Gradangaben im WGS84-Koordinatensystem. • <i>bits</i>: Die Anzahl der zum Speichern des Hash verwendeten Bit. Bereich [1,75]. Dies wirkt sich sowohl auf die Länge der zurückgegebenen Zeichenfolge (pro jeweils 5 Bit wird ein Zeichen verwendet) als auch auf die Hashgenauigkeit aus. 5 Bit (1 Zeichen) stellen beispielsweise ca 2500 Kilometer oder 45 Bit (9 Zeichen) stellen ca. 2,3 Meter dar.

Vergleichsfunktionen

Mit Vergleichsfunktionen werden Feldwerte miteinander oder mit einer angegebenen Zeichenfolge verglichen. Beispielsweise können Sie Zeichenfolgen mit = auf Gleichheit überprüfen. Ein Beispiel für die Überprüfung der Zeichenfolgengleichheit ist: Klasse = "Klasse 1".

Bei numerischen Vergleichen bedeutet *größer* näher an der positiven Unendlichkeit und *kleiner* näher an der negativen Unendlichkeit. Negativen Zahlen sind also stets kleiner als jede positive Zahl.

Tabelle 23. CLEM-Vergleichsfunktionen

Funktion	Ergebnis	Beschreibung
count_equal(ELEMENT1, LISTE)	ganze Zahl	Ergibt die Anzahl der Werte aus einer Felderliste, die gleich ELEMENT1 sind, bzw. den Wert null, wenn ELEMENT1 gleich null ist.
count_greater_than(ELEMENT1, LISTE)	ganze Zahl	Ergibt die Anzahl der Werte aus einer Felderliste, die größer als ELEMENT1 sind, bzw. den Wert null, wenn ELEMENT1 gleich null ist.
count_less_than(ELEMENT1, LISTE)	ganze Zahl	Ergibt die Anzahl der Werte aus einer Felderliste, die kleiner als ELEMENT1 sind, bzw. den Wert null, wenn ELEMENT1 gleich null ist.
count_not_equal(ELEMENT1, LISTE)	ganze Zahl	Ergibt die Anzahl der Werte aus einer Felderliste, die ungleich ELEMENT1 sind, bzw. den Wert null, wenn ELEMENT1 gleich null ist.
count_nulls(LISTE)	ganze Zahl	Ergibt die Anzahl der Nullwerte aus einer Felderliste.
count_non_nulls(LISTE)	ganze Zahl	Ergibt die Anzahl der Nichtnullwerte aus einer Felderliste.
date_before(DATUM1, DATUM2)	boolesch	Dient zur Überprüfung der alphabetischen Sortierung von Datumswerten. Ergibt den Wert "Wahr", wenn DATUM1 vor DATUM2 liegt.
first_index(ELEMENT, LISTE)	ganze Zahl	Ergibt den Index des ersten Felds aus einer LISTE von Feldern, das ELEMENT enthält, oder "0", wenn der Wert nicht gefunden wird. Nur für die Typen "Zeichenfolge", "Ganze Zahl" und "Reelle Zahl" unterstützt.
first_non_null(LISTE)	beliebig	Ergibt den ersten Nichtnullwert in der angegebenen Felderliste. Alle Speichertypen werden unterstützt.
first_non_null_index(LISTE)	ganze Zahl	Ergibt den Index des ersten Felds in der angegebenen LISTE, das einen Nichtnullwert enthält, oder "0", wenn alle Werte null sind. Alle Speichertypen werden unterstützt.
ELEMENT1 = ELEMENT2	boolesch	Ergibt den Wert "Wahr" für Datensätze, bei denen ELEMENT1 gleich ELEMENT2 ist.
ELEMENT1 /= ELEMENT2	boolesch	Ergibt den Wert "Wahr", wenn die beiden Zeichenfolgen nicht identisch sind, oder "0", wenn sie identisch sind.
ELEMENT1 < ELEMENT2	boolesch	Ergibt den Wert "Wahr" für Datensätze, bei denen ELEMENT1 kleiner als ELEMENT2 ist.
ELEMENT1 <= ELEMENT2	boolesch	Ergibt den Wert "Wahr" für Datensätze, bei denen ELEMENT1 kleiner-gleich ELEMENT2 ist.
ELEMENT1 > ELEMENT2	boolesch	Ergibt den Wert "Wahr" für Datensätze, bei denen ELEMENT1 größer als ELEMENT2 ist.
ELEMENT1 >= ELEMENT2	boolesch	Ergibt den Wert "Wahr" für Datensätze, bei denen ELEMENT1 größer-gleich ELEMENT2 ist.
last_index(ELEMENT, LISTE)	ganze Zahl	Ergibt den Index des letzten Felds aus einer LISTE von Feldern, das ELEMENT enthält, oder "0", wenn der Wert nicht gefunden wird. Nur für die Typen "Zeichenfolge", "Ganze Zahl" und "Reelle Zahl" unterstützt.
last_non_null(LISTE)	beliebig	Ergibt den letzten Nichtnullwert in der angegebenen Felderliste. Alle Speichertypen werden unterstützt.
last_non_null_index(LISTE)	ganze Zahl	Ergibt den Index des letzten Felds in der angegebenen LISTE, das einen Nichtnullwert enthält, oder "0", wenn alle Werte null sind. Alle Speichertypen werden unterstützt.
max(ELEMENT1, ELEMENT2)	beliebig	Ergibt das größere der beiden Elemente: ELEMENT1 oder ELEMENT2.

Table 23. CLEM-Vergleichsfunktionen (Forts.)

Funktion	Ergebnis	Beschreibung
max_index(LISTE)	ganze Zahl	Ergibt den Index des Felds mit dem größten Wert aus einer Liste numerischer Felder bzw. "0", wenn alle Feldwerte null sind. Wenn beispielsweise das dritte Feld in der Liste den größten Wert enthält, wird der Indexwert 3 ausgegeben. Wenn mehrere Felder den größten Wert enthalten, wird das zuerst aufgelistete (von links nach rechts) ausgegeben.
max_n(LISTE)	Zahl	Ergibt den größten Wert aus einer Liste numerischer Felder bzw. den Wert null, wenn alle Feldwerte gleich null sind.
member(ELEMENT, LISTE)	boolesch	Ergibt den Wert "Wahr", wenn ELEMENT ein Mitglied der angegebenen LISTE ist. Andernfalls ergibt sich ein Falsch-Wert. Auch eine Liste mit Feldnamen kann angegeben werden.
min(ELEMENT1, ELEMENT2)	beliebig	Ergibt das kleinere der beiden Elemente: ELEMENT1 oder ELEMENT2.
min_index(LISTE)	ganze Zahl	Ergibt den Index des Felds mit dem kleinsten Wert aus einer Liste numerischer Felder bzw. "0", wenn alle Feldwerte null sind. Wenn beispielsweise das dritte Feld in der Liste den kleinsten Wert enthält, wird der Indexwert 3 ausgegeben. Wenn mehrere Felder den kleinsten Wert enthalten, wird das zuerst aufgelistete (von links nach rechts) ausgegeben.
min_n(LISTE)	Zahl	Ergibt den kleinsten Wert aus einer Liste numerischer Felder bzw. den Wert null, wenn alle Feldwerte gleich null sind.
time_before(ZEIT1, ZEIT2)	boolesch	Dient zur Überprüfung der alphabetischen Sortierung von Zeitwerten. Ergibt den Wert "Wahr", wenn ZEIT1 vor ZEIT2 liegt.
value_at(GANZZ, LISTE)		Gibt den Wert jedes aufgelisteten Felds bei Offset GANZZ zurück oder NULL, wenn das Offset außerhalb des Bereichs der gültigen Werte liegt (d. h. kleiner als 1 oder größer als die Anzahl der aufgelisteten Felder ist). Alle Speichertypen werden unterstützt.

Logische Funktionen

CLEM-Ausdrücke können zur Durchführung logischer Operationen verwendet werden.

Table 24. Logische Funktionen von CLEM

Funktion	Ergebnis	Beschreibung
BEDG1 and BEDG2	boolesch	Bei dieser Operation handelt es sich um eine logische Konjunktion, die den Wert "Wahr" ergibt, wenn sowohl BEDG1 als auch BEDG2 wahr sind. Wenn BEDG1 falsch ist, wird BEDG2 nicht ausgewertet; dadurch sind Konjunktionen möglich, bei denen BEDG1 zuerst getestet, ob eine Operation in BEDG2 zulässig ist. Beispiele length(Label) >=6 und Label(6) = 'x'.
BEDG1 or BEDG2	boolesch	Bei dieser Operation handelt es sich um eine logische (inklusive) Disjunktion, die den Wert "Wahr" ergibt, wenn entweder BEDG1 oder BEDG2 wahr ist oder wenn beide wahr sind. Wenn BEDG1 wahr ist, wird BEDG2 nicht ausgewertet.
not(BEDG)	boolesch	Bei dieser Operation handelt es sich um eine logische Negation, die den Wert "Wahr" ergibt, wenn BEDG falsch ist. Andernfalls ergibt diese Operation den Wert 0.
if BEDG then AUSDR1 else AUSDR2 endif	beliebig	Bei dieser Operation handelt es sich um eine bedingte Evaluierung. Wenn BEDG wahr ist, ergibt die Operation das Ergebnis von AUSDR1. Andernfalls ergibt sich das Ergebnis der Auswertung von AUSDR2.

Tabelle 24. Logische Funktionen von CLEM (Forts.)

Funktion	Ergebnis	Beschreibung
if BEDG1 then AUSDR1 elseif BEDG2 then AUSDR2 else AUSDR_N endif	beliebig	Bei dieser Operation handelt es sich um eine bedingte Evaluierung mit mehreren Zweigen. Wenn <i>BEDG1</i> wahr ist, ergibt die Operation das Ergebnis von <i>AUSDR1</i> . Andernfalls, wenn <i>BEDG2</i> wahr ist, ergibt die Operation das Ergebnis der Auswertung von <i>AUSDR2</i> . Andernfalls ergibt sich das Ergebnis der Auswertung von <i>AUSDR_N</i> .

Numerische Funktionen

CLEM enthält eine Reihe häufig verwendeter numerischer Funktionen.

Tabelle 25. Numerische Funktionen von CLEM

Funktion	Ergebnis	Beschreibung
-ZAHL	Zahl	Wird zur Umkehrung des Vorzeichens von <i>ZAHL</i> verwendet. Ergibt die entsprechende Zahl mit dem entgegengesetzten Vorzeichen.
ZAHL1 + ZAHL2	Zahl	Ergibt die Summe von <i>ZAHL1</i> und <i>ZAHL2</i> .
ZAHL1 - ZAHL2	Zahl	Ergibt die Differenz von <i>ZAHL1</i> und <i>ZAHL2</i> .
ZAHL1 * ZAHL2	Zahl	Ergibt das Produkt aus <i>ZAHL1</i> und <i>ZAHL2</i> .
ZAHL1 / ZAHL2	Zahl	Ergibt den Quotienten aus <i>ZAHL1</i> und <i>ZAHL2</i> .
GANZZ1 div GANZZ2	Zahl	Wird zur Durchführung einer Division zwischen zwei ganzen Zahlen verwendet. Ergibt den Wert von <i>GANZZ1</i> dividiert durch <i>GANZZ2</i> .
GANZZ1 rem GANZZ2	Zahl	Ergibt den Rest des Quotienten aus <i>GANZZ1</i> und <i>GANZZ2</i> . Beispiel: $GANZZ1 - (GANZZ1 \text{ div } GANZZ2) * GANZZ2$.
GANZZ1 mod GANZZ2	Zahl	Diese Funktion wird nicht mehr verwendet. Verwenden Sie stattdessen die Funktion <i>rem</i> .
BASIS ** POTENZ	Zahl	Ergibt den Wert von <i>BASIS</i> hoch <i>POTENZ</i> . Für beide Werte ist eine beliebige Zahl möglich (mit der Ausnahme, dass <i>BASIS</i> nicht null sein darf, wenn <i>POTENZ</i> einen anderen Nullwert als die ganze Zahl 0 aufweist). Wenn <i>POTENZ</i> eine ganze Zahl ist, erfolgt die Berechnung, indem <i>BASIS</i> mehrmals nacheinander mit sich selbst multipliziert wird. Wenn <i>BASIS</i> also eine ganze Zahl ist, ist das Ergebnis ebenfalls eine ganze Zahl. Wenn <i>POTENZ</i> die ganze Zahl 0 ist, ist das Ergebnis immer eine 1 mit demselben Typ wie <i>BASIS</i> . Andernfalls, wenn <i>POTENZ</i> keine ganze Zahl ist, wird das Ergebnis als $\exp(POTENZ * \log(BASIS))$ berechnet.
abs(ZAHL)	Zahl	Ergibt den absoluten Wert von <i>ZAHL</i> , bei dem es sich immer um eine Zahl desselben Typs handelt.
exp(ZAHL)	reelle Zahl	Ergibt den Wert von <i>e</i> hoch <i>ZAHL</i> . Dabei ist <i>e</i> die Basis eines natürlichen Logarithmus.
fracof(ZAHL)	reelle Zahl	Ergibt den Bruchbereich (Nachkommastellen) von <i>ZAHL</i> , der definiert ist als $ZAHL - \text{intof}(ZAHL)$.
intof(ZAHL)	ganze Zahl	Kürzt das Argument zu einer ganzen Zahl. Ergibt eine ganze Zahl mit demselben Vorzeichen wie <i>Zahl</i> und mit dem größten Absolutwert, sodass $\text{abs}(GANZZ) \leq \text{abs}(ZAHL)$.
log(ZAHL)	reelle Zahl	Ergibt den natürlichen Logarithmus (Basis <i>e</i>) von <i>ZAHL</i> (darf keine Art von Nullwert sein).

Tabelle 25. Numerische Funktionen von CLEM (Forts.)

Funktion	Ergebnis	Beschreibung
log10(ZAHL)	reelle Zahl	Ergibt den Logarithmus zur Basis 10 von ZAHL (darf keine Art von Nullwert sein). Diese Funktion ist als $\log(ZAHL) / \log(10)$ definiert.
negate(ZAHL)	Zahl	Wird zur Umkehrung des Vorzeichens von ZAHL verwendet. Ergibt die entsprechende Zahl mit dem entgegengesetzten Vorzeichen.
round(ZAHL)	ganze Zahl	Dient zum Runden von ZAHL auf eine ganze Zahl. Dabei wird $\text{intof}(ZAHL+0,5)$ verwendet, wenn ZAHL positiv ist, oder $\text{intof}(ZAHL-0,5)$, wenn ZAHL negativ ist.
sign(ZAHL)	Zahl	Wird zur Ermittlung des Vorzeichens von ZAHL verwendet. Diese Operation ergibt -1, 0 oder 1, wenn ZAHL eine ganze Zahl ist. Wenn ZAHL eine reelle Zahl ist, ergibt sich -1,0; 0,0 oder 1,0, je nachdem ob ZAHL negativ, null oder positiv ist.
sqrt(ZAHL)	reelle Zahl	Ergibt die Quadratwurzel von ZAHL. ZAHL muss positiv sein.
sum_n(LISTE)	Zahl	Ergibt die Summe der Werte aus einer Liste numerischer Felder bzw. den Wert null, wenn alle Feldwerte gleich null sind.
mean_n(LISTE)	Zahl	Ergibt den Mittelwert aus einer Liste numerischer Felder bzw. den Wert null, wenn alle Feldwerte gleich null sind.
sdev_n(LISTE)	Zahl	Ergibt die Standardabweichung aus einer Liste numerischer Felder bzw. den Wert null, wenn alle Feldwerte gleich null sind.

Trigonometrische Funktionen

Alle Funktionen in diesem Abschnitt verwenden entweder einen Winkel als Argument oder ergeben einen Winkel als Ergebnis. In beiden Fällen richtet sich die für den Winkel verwendete Einheit (Radianten oder Grad) nach die Einstellung in der entsprechenden Streamoption.

Tabelle 26. Trigonometrische Funktionen von CLEM

Funktion	Ergebnis	Beschreibung
arccos(ZAHL)	reelle Zahl	Berechnet den Arcuscosinus des angegebenen Winkels.
arccosh(ZAHL)	reelle Zahl	Berechnet den hyperbolischen Arcuscosinus des angegebenen Winkels.
arcsin(ZAHL)	reelle Zahl	Berechnet den Arcussinus des angegebenen Winkels.
arcsinh(ZAHL)	reelle Zahl	Berechnet den hyperbolischen Arcussinus des angegebenen Winkels.
arctan(ZAHL)	reelle Zahl	Berechnet den Arcustangens des angegebenen Winkels.
arctan2(ZAHL_Y, ZAHL_X)	reelle Zahl	Berechnet den Arcustangens von ZAHL_Y / ZAHL_X und verwendet die Vorzeichen der beiden Zahlen zur Ableitung von Quadranteninformationen. Das Ergebnis ist eine reelle Zahl im Bereich $-\pi < \text{WINKEL} \leq \pi$ (rad) - $180 < \text{ANGLE} \leq 180$ (deg)
arctanh(ZAHL)	reelle Zahl	Berechnet den hyperbolischen Arcustangens des angegebenen Winkels.
cos(ZAHL)	reelle Zahl	Berechnet den Cosinus des angegebenen Winkels.
cosh(ZAHL)	reelle Zahl	Berechnet den hyperbolischen Cosinus des angegebenen Winkels.
pi	reelle Zahl	Diese Konstante ist die beste reelle Annäherung an Pi.
sin(ZAHL)	reelle Zahl	Berechnet den Sinus des angegebenen Winkels.
sinh(ZAHL)	reelle Zahl	Berechnet den hyperbolischen Sinus des angegebenen Winkels.
tan(ZAHL)	reelle Zahl	Berechnet den Tangens des angegebenen Winkels.

Table 26. Trigonometrische Funktionen von CLEM (Forts.)

Funktion	Ergebnis	Beschreibung
tanh(ZAHL)	reelle Zahl	Berechnet den hyperbolischen Tangens des angegebenen Winkels.

Wahrscheinlichkeitsfunktionen

Wahrscheinlichkeitsfunktionen ergeben Wahrscheinlichkeiten auf der Grundlage verschiedener Verteilungen, z. B. die Wahrscheinlichkeit, dass ein Wert aus der *t*-Verteilung in der Studentenversion kleiner ist als ein bestimmter Wert.

Table 27. CLEM-Wahrscheinlichkeitsfunktionen

Funktion	Ergebnis	Beschreibung
cdf_chisq(ZAHL, DF)	reelle Zahl	Ergibt die Wahrscheinlichkeit, dass ein Wert aus der Chi-Quadrat-Verteilung mit den angegebenen Freiheitsgraden kleiner ist als die angegebene Zahl.
cdf_f(ZAHL, DF1, DF2)	reelle Zahl	Ergibt die Wahrscheinlichkeit, dass ein Wert aus der <i>F</i> -Verteilung mit den angegebenen Freiheitsgraden <i>DF1</i> und <i>DF2</i> kleiner ist als die angegebene Zahl.
cdf_normal(ZAHL, MITTELWERT, STDABW)	reelle Zahl	Ergibt die Wahrscheinlichkeit, dass ein Wert aus der Normalverteilung mit dem angegebenen Mittelwert und der angegebenen Standardabweichung kleiner ist als die angegebene Zahl.
cdf_t(ZAHL, DF)	reelle Zahl	Ergibt die Wahrscheinlichkeit, dass ein Wert aus der Chi-Quadrat-Verteilung aus einer <i>t</i> -Verteilung in der Studentenversion mit den angegebenen Freiheitsgraden kleiner ist als die angegebene Zahl.

Räumliche Funktionen

Räumliche Funktionen können mit georäumlichen Daten verwendet werden. Sie ermöglichen z. B. das Berechnen der Abstände zwischen zwei Punkten, der Fläche eines Polygons usw. Es kann auch Situationen geben, in denen eine Zusammenführung mehrerer georäumlicher Datasets erforderlich ist, die auf einem Prädikat für räumliche Daten (innerhalb, nahe usw.) basieren, was über eine Zusammenführungsbedingung bewerkstelligt werden kann.

Diese räumlichen Funktionen arbeiten mit dem über **Tools > Streameigenschaften > Optionen > Räumlich** angegebenen Koordinatensystem.

Anmerkung: Diese räumlichen Funktionen gelten nicht für dreidimensionale Daten. Wenn dreidimensionale Daten in den Stream importiert werden, werden nur die ersten beiden Dimensionen von diesen Funktionen verwendet. Die Werte der z-Achse werden ignoriert.

Table 28. Räumliche Funktionen von CLEM

Funktion	Ergebnis	Beschreibung
close_to(SHAPE, SHAPE, NUM)	boolesch	Testet, ob zwei Formen sich innerhalb einer bestimmten DISTANZ voneinander befinden. Wenn ein projiziertes Koordinatensystem verwendet wird, wird die Distanz in Meter angegeben. Wird kein Koordinatensystem angegeben, gilt eine beliebige Einheit.
crosses(SHAPE, SHAPE)	boolesch	Testet, ob zwei Formen einander überschneiden. Diese Funktion eignet sich für zwei Linienfolgenformen oder eine Linienfolge und ein Polygon.

Tabelle 28. Räumliche Funktionen von CLEM (Forts.)

Funktion	Ergebnis	Beschreibung
overlap(SHAPE,SHAPE)	boolesch	Testet, ob ein Schnittpunkt zwischen zwei Polygonen vorhanden ist und ob der Schnittpunkt innerhalb der beiden Formen liegt.
within(SHAPE,SHAPE)	boolesch	Testet, ob die Gesamtheit von SHAPE1 innerhalb eines Polygons (POLYGON) liegt.
area(SHAPE)	reelle Zahl	Gibt die Fläche des angegebenen Polygons (POLYGON) zurück. Wenn ein projiziertes System verwendet wird, gibt die Funktion Quadratmeter zurück. Wird kein Koordinatensystem angegeben, gilt eine beliebige Einheit. Die Form muss ein POLYGON oder ein MULTIPOLYGON sein.
num_points(SHAPE,LIST)	ganze Zahl	Gibt die Anzahl von Punkten aus einem Punktfeld (MULTIPOINT) zurück, die innerhalb der Grenzen eines Polygons (POLYGON) liegen. SHAPE1 muss ein POLYGON oder ein MULTIPOLYGON sein.
distance(SHAPE,SHAPE)	reelle Zahl	Gibt den Abstand zwischen SHAPE1 und SHAPE2 zurück. Wenn ein projiziertes Koordinatensystem verwendet wird, gibt die Funktion Meter zurück. Wird kein Koordinatensystem angegeben, gilt eine beliebige Einheit. SHAPE1 und SHAPE2 können jeden geografischen Messtyp haben.

Bitweise Operationen mit ganzen Zahlen

Mit dieser Funktion können Ganzzahlen als Bitmuster behandelt werden, die die Zweierkomplementwerte darstellen, bei denen die Bitposition N den Wert 2^{*N} aufweist. Bit werden beginnend bei 0 nummeriert. Diese Operationen funktionieren so, als ob das Vorzeichenbit einer ganzen Zahl endlos nach links erweitert wird. Daher weisen positive ganze Zahlen überall oberhalb des werthöchsten 0-Bits auf, negative ganze Zahlen 1-Bits.

Tabelle 29. Bitweise Ganzzahloperationen von CLEM

Funktion	Ergebnis	Beschreibung
\sim GANZZ1	ganze Zahl	Ergibt das bitweise Komplement der ganzen Zahl GANZZ1. Im Ergebnis steht also eine 1 für jede Bitposition, bei der GANZZ1 den Wert 0 aufweist. Es gilt immer: \sim GANZZ = -(GANZZ + 1).
GANZZ1 GANZZ2	ganze Zahl	Das Ergebnis dieser Operation ist das bitweise "inklusive Oder" von GANZZ1 und GANZZ2. Im Ergebnis steht also eine 1 für jede Bitposition, bei der bei GANZZ1 und/oder bei GANZZ2 der Wert 1 steht.
GANZZ1 /& GANZZ2	ganze Zahl	Das Ergebnis dieser Operation ist das bitweise "exklusive Oder" von GANZZ1 und GANZZ2. Im Ergebnis steht also eine 1 für jede Bitposition, bei der entweder bei GANZZ1 oder bei GANZZ2 der Wert 1 steht, nicht jedoch bei beiden.
GANZZ1 && GANZZ2	ganze Zahl	Ergibt das bitweise "Und" der ganzen Zahlen GANZZ1 und GANZZ2. Im Ergebnis steht also eine 1 für jede Bitposition, bei der sowohl bei GANZZ1 als auch bei GANZZ2 der Wert 1 steht.
GANZZ1 && \sim GANZZ2	ganze Zahl	Ergibt das bitweise "Und" von GANZZ1 und das bitweise Komplement von GANZZ2. Im Ergebnis steht also eine 1 für jede Bitposition, bei der bei GANZZ1 der Wert 1 und bei GANZZ2 der Wert 0 steht. Diese Operation ist identisch mit GANZZ1 && (\sim GANZZ2) und ist sinnvoll zum Bereinigen von in GANZZ2 gesetzten Bits von GANZZ1.

Tabelle 29. Bitweise Ganzzahloperationen von CLEM (Forts.)

Funktion	Ergebnis	Beschreibung
<code>GANZZ << N</code>	<i>ganze Zahl</i>	Ergibt das Bitmuster von <code>GANZZ1</code> , um <code>N</code> Positionen nach links verschoben. Ein negativer Wert von <code>N</code> führt zu einer Verschiebung nach rechts.
<code>GANZZ >> N</code>	<i>ganze Zahl</i>	Ergibt das Bitmuster von <code>GANZZ1</code> , um <code>N</code> Positionen nach rechts verschoben. Ein negativer Wert von <code>N</code> führt zu einer Verschiebung nach links.
<code>GANZZ1 &&=_0 GANZZ2</code>	<i>boolesch</i>	Entspricht dem booleschen Ausdruck <code>GANZZ1 && GANZZ2 != 0</code> , ist jedoch effizienter.
<code>GANZZ1 &&/= _0 GANZZ2</code>	<i>boolesch</i>	Entspricht dem booleschen Ausdruck <code>GANZZ1 && GANZZ2 == 0</code> , ist jedoch effizienter.
<code>integer_bitcount(GANZZ)</code>	<i>ganze Zahl</i>	Zählt die Anzahl der 1- oder 0-Bits in der Zweierkomplementdarstellung von <code>GANZZ</code> . Wenn <code>GANZZ</code> nicht negativ ist, ist <code>N</code> die Anzahl der 1-Bits. Wenn <code>GANZZ</code> negativ ist, gibt dieser Wert die Anzahl der 0-Bits an. Aufgrund der Vorzeichenerweiterung gibt es unendlich viele 0-Bits in nicht negativen ganzen Zahlen bzw. 1-Bits in negativen ganzen Zahlen. Es gilt immer: <code>integer_bitcount(GANZZ) = integer_bitcount(-(GANZZ+1))</code> .
<code>integer_leastbit(GANZZ)</code>	<i>ganze Zahl</i>	Gibt die Bitposition <code>N</code> des niedrigstwertigen Bits an, das in der ganzen Zahl <code>GANZZ</code> gesetzt wurde. <code>N</code> ist die höchste Potenz von 2, durch die sich <code>GANZZ</code> ohne Rest teilen lässt.
<code>integer_length(GANZZ)</code>	<i>ganze Zahl</i>	Ergibt die Länge von <code>GANZZ</code> als ganze Zweierkomplementzahl in Bit. <code>N</code> ist also die kleinste ganze Zahl, sodass gilt: <code>GANZZ < (1 << N)</code> , wenn <code>GANZZ >= 0</code> . <code>GANZZ >= (-1 << N)</code> , wenn <code>GANZZ < 0</code> . Wenn <code>GANZZ</code> nicht negativ ist, ist für die Darstellung von <code>GANZZ</code> als ganze Zahl ohne Vorzeichen ein Feld mit mindestens <code>N</code> Bit erforderlich. Alternativ sind, unabhängig vom Vorzeichen, mindestens <code>N+1</code> Bit erforderlich, um <code>GANZZ</code> als ganze Zahl mit Vorzeichen darzustellen.
<code>testbit(GANZZ, N)</code>	<i>boolesch</i>	Testet das Bit an Position <code>N</code> in der ganzen Zahl <code>GANZZ</code> und ergibt den Status von Bit <code>N</code> als booleschen Wert, der bei 1 "wahr" und bei 0 "falsch" ist.

Zufallsfunktionen

Folgende Funktionen dienen zur zufälligen Auswahl von Elementen oder zur Generierung von Zufallszahlen.

Tabelle 30. CLEM-Zufallsfunktionen

Funktion	Ergebnis	Beschreibung
<code>oneof(LISTE)</code>	<i>beliebig</i>	Ergibt ein zufällig ausgewähltes Element von <code>LISTE</code> . Die Listenelemente sollten wie folgt eingegeben werden <code>[ELEMENT1, ELEMENT2, ..., ELEMENT_N]</code> . Auch eine Liste mit Feldnamen kann angegeben werden.
<code>random(ZAHL)</code>	<i>Zahl</i>	Ergibt eine gleichmäßig verteilte Zufallszahl desselben Typs (<code>GANZZ</code> oder <code>REELL</code>). Es gilt der Bereich von 1 bis <code>ZAHL</code> . Bei Verwendung einer ganzen Zahl ergeben sich nur ganze Zahlen. Bei Verwendung einer reellen Zahl (Dezimalzahl), ergeben sich reelle Zahlen (die Anzahl der Dezimalstellen werden in den Streamoptionen festgelegt). Die größte Zufallszahl, die die Funktion ergibt, kann gleich <code>ZAHL</code> sein.

Tabelle 30. CLEM-Zufallsfunktionen (Forts.)

Funktion	Ergebnis	Beschreibung
random0(ZAHL)	Zahl	Diese Operation hat dieselben Eigenschaften wie random(ZAHL), der Bereich beginnt jedoch bei 0. Die größte Zufallszahl, die die Funktion ergibt, ist nie gleich ZAHL.

Zeichenfolgefunktionen

In CLEM können Sie folgende Operationen an Zeichenfolgen durchführen:

- Vergleichen von Zeichenfolgen
- Erstellen von Zeichenfolgen
- Zugreifen auf Zeichen

In CLEM-Ausdrücken ist eine Zeichenfolge jede Folge von Zeichen zwischen zusammengehörenden (doppelten) Anführungszeichen ("Zeichenfolge in Anführungszeichen"). Zeichen (ZEICHEN) können beliebige alphanumerische Zeichen sein. Sie werden in CLEM-Ausdrücken mithilfe einfacher umgekehrter Anführungsstriche in der Form ``<Zeichen>`` deklariert, wie beispielsweise ``z``, ``A`` oder ``2``. Zeichen, die außerhalb des zulässigen Bereichs liegen, oder negative Indizes bei einer Zeichenfolge führen zu einem undefinierten Verhalten.

Anmerkung: Vergleiche zwischen Zeichenfolgen, die SQL-Pushback verwenden oder nicht, können unterschiedliche Ergebnisse generieren, wenn nachfolgende Leerzeichen vorhanden sind.

Tabelle 31. CLEM-Zeichenfolgefunktionen

Funktion	Ergebnis	Beschreibung
allbutfirst(N, ZEICHENFOLGE)	Zeichenfolge	Ergibt eine Zeichenfolge, bei der es sich um ZEICHENFOLGE nach der Entfernung der ersten N Zeichen handelt.
allbutlast(N, ZEICHENFOLGE)	Zeichenfolge	Ergibt eine Zeichenfolge, bei der es sich um ZEICHENFOLGE nach der Entfernung der letzten N Zeichen handelt.
alphabefore(ZEICHENFOLGE1, ZEICHENFOLGE2)	boolesch	Dient zur Überprüfung der alphabetischen Sortierung von Zeichenfolgen. Ergibt den Wert "Wahr", wenn ZEICHENFOLGE1 der ZEICHENFOLGE2 vorausgeht.
endstring(LÄNGE, ZEICHENFOLGE)	Zeichenfolge	Extrahiert die letzten N Zeichen aus der angegebenen Zeichenfolge. Ist eine Zeichenfolge kleinergleich der angegebenen Länge, dann bleibt die Zeichenfolge unverändert.
hasendstring(ZEICHENFOLGE, TEILZEICHENFOLGE)	ganze Zahl	Diese Funktion entspricht isendstring(TEILZEICHENFOLGE, ZEICHENFOLGE).
hasmidstring(ZEICHENFOLGE, TEILZEICHENFOLGE)	ganze Zahl	Diese Funktion entspricht ismidstring(TEILZEICHENFOLGE, ZEICHENFOLGE) (eingebettete Teilzeichenfolge).
hasstartstring(ZEICHENFOLGE, TEILZEICHENFOLGE)	ganze Zahl	Diese Funktion entspricht isstartstring(TEILZEICHENFOLGE, ZEICHENFOLGE).
hassubstring(ZEICHENFOLGE, N, TEILZEICHENFOLGE)	ganze Zahl	Diese Funktion entspricht issubstring(TEILZEICHENFOLGE, N, ZEICHENFOLGE). Dabei hat N standardmäßig den Wert 1.

Tabelle 31. CLEM-Zeichenfolgefunktionen (Forts.)

Funktion	Ergebnis	Beschreibung
count_substring(ZEICHENFOLGE, TEILZEICHENFOLGE)	ganze Zahl	Ergibt den Wert, wie oft die angegebene Teilzeichenfolge in der Zeichenfolge auftritt. Beispiel: count_substring("foooo.txt", "oo") ergibt den Wert 3.
hassubstring(ZEICHENFOLGE, TEILZEICHENFOLGE)	ganze Zahl	Diese Funktion entspricht issubstring(TEILZEICHENFOLGE, 1, ZEICHENFOLGE). Dabei hat N standardmäßig den Wert 1.
isalphacode(ZEICHEN)	boolesch	Ergibt den Wert "Wahr", wenn ZEICHEN ein Zeichen in der angegebenen Zeichenfolge (häufig ein Feldname) ist, dessen Zeichencode ein Buchstabe ist. Andernfalls ergibt die Funktion den Wert 0. Beispiel: isalphacode(produce_num(1)).
isendstring(TEILZEICHENFOLGE, ZEICHENFOLGE)	ganze Zahl	Wenn die Zeichenfolge ZEICHENFOLGE mit der Teilzeichenfolge TEILZEICHENFOLGE endet, ergibt diese Funktion den ganzzahligen Index von TEILZEICHENFOLGE in ZEICHENFOLGE. Andernfalls ergibt die Funktion den Wert 0.
islowercode(ZEICHEN)	boolesch	Ergibt den Wert "Wahr", wenn ZEICHEN ein Kleinbuchstabe der angegebenen Zeichenfolge (häufig ein Feldname) ist. Andernfalls ergibt die Funktion den Wert 0. Beispielsweise sind sowohl islowercode(``) als auch islowercode(country_name(2)) gültige Ausdrücke.
ismidstring(TEILZEICHENFOLGE, ZEICHENFOLGE)	ganze Zahl	Wenn TEILZEICHENFOLGE eine Teilzeichenfolge von ZEICHENFOLGE ist, jedoch nicht beim ersten Zeichen von ZEICHENFOLGE beginnt oder mit dem letzten Zeichen endet, ergibt diese Funktion den Index, bei dem die Teilzeichenfolge beginnt. Andernfalls ergibt die Funktion den Wert 0.
isnumbercode(ZEICHEN)	boolesch	Ergibt den Wert "Wahr", wenn ZEICHEN bei der angegebenen Zeichenfolge (häufig ein Feldname) ein Zeichen ist, dessen Zeichencode eine Ziffer ist. Andernfalls ergibt die Funktion den Wert 0. Beispiel: isnumbercode(product_id(2)).
isstartstring(TEILZEICHENFOLGE, ZEICHENFOLGE)	ganze Zahl	Wenn die Zeichenfolge ZEICHENFOLGE mit der Teilzeichenfolge TEILZEICHENFOLGE beginnt, ergibt diese Funktion den Index 1. Andernfalls ergibt die Funktion den Wert 0.
issubstring(TEILZEICHENFOLGE, N, ZEICHENFOLGE)	ganze Zahl	Durchsucht die Zeichenfolge ZEICHENFOLGE beginnend beim N. Zeichen nach einer Teilzeichenfolge, die mit TEILZEICHENFOLGE übereinstimmt. Wenn die Teilzeichenfolge gefunden wird, ergibt diese Funktion den Index, bei dem die Teilzeichenfolge beginnt. Andernfalls ergibt die Funktion den Wert 0. Wenn N nicht angegeben ist, wird standardmäßig der Wert 1 verwendet.

Tabelle 31. CLEM-Zeichenfolgefunktionen (Forts.)

Funktion	Ergebnis	Beschreibung
issubstring(TEILZEICHENFOLGE, ZEICHENFOLGE)	ganze Zahl	Durchsucht die Zeichenfolge <i>ZEICHENFOLGE</i> beginnend beim <i>N.</i> Zeichen nach einer Teilzeichenfolge, die mit <i>TEILZEICHENFOLGE</i> übereinstimmt. Wenn die Teilzeichenfolge gefunden wird, ergibt diese Funktion den Index, bei dem die Teilzeichenfolge beginnt. Andernfalls ergibt die Funktion den Wert 0. Wenn <i>N</i> nicht angegeben ist, wird standardmäßig der Wert 1 verwendet.
issubstring_count(TEILZEICHENFOLGE, N, ZEICHENFOLGE):	ganze Zahl	Ergibt den Index für das <i>N</i> -te Vorkommen von <i>TEILZEICHENFOLGE</i> in der angegebenen <i>ZEICHENFOLGE</i> . Falls <i>TEILZEICHENFOLGE</i> weniger als <i>N</i> -mal auftritt, ergibt sich der Wert 0.
issubstring_lim(TEILZEICHENFOLGE, N, STARTLIM, ENDLIM, ZEICHENFOLGE)	ganze Zahl	Diese Funktion entspricht <i>issubstring</i> , der Treffer muss jedoch beim oder vor dem Index <i>STARTLIM</i> beginnen und beim oder vor dem Index <i>ENDLIM</i> enden. Die Beschränkungen <i>STARTLIM</i> bzw. <i>ENDLIM</i> können inaktiviert werden, indem ein Falsch-Wert für eines der Argumente angegeben wird. So ist beispielsweise <i>issubstring_lim(TEILZEICHENFOLGE, N, false, false, ZEICHENFOLGE)</i> mit <i>issubstring</i> identisch.
isuppercode(ZEICHEN)	boolesch	Ergibt den Wert "Wahr", wenn <i>ZEICHEN</i> ein Großbuchstabe ist. Andernfalls ergibt die Funktion den Wert 0. Beispielsweise sind sowohl <i>isuppercode(``)</i> als auch <i>isuppercode(country_name(2))</i> gültige Ausdrücke.
last(ZEICHEN)	Zeichenfolge	Ergibt das letzte Zeichen <i>ZEICHEN</i> von <i>ZEICHENFOLGE</i> (muss mindestens ein Zeichen lang sein).
length(ZEICHENFOLGE)	ganze Zahl	Ergibt die Länge der Zeichenfolge <i>ZEICHENFOLGE</i> , d. h. die Anzahl der Zeichen in der Zeichenfolge.
locchar(ZEICHEN, N, ZEICHENFOLGE)	ganze Zahl	Mit dieser Funktion kann die Position der Zeichen in symbolischen Feldern ermittelt werden. Mit dieser Funktion wird die Zeichenfolge <i>ZEICHENFOLGE</i> nach dem Zeichen <i>ZEICHEN</i> durchsucht, wobei der Suchvorgang beim <i>N.</i> Zeichen von <i>ZEICHENFOLGE</i> gestartet wird. Diese Funktion ergibt einen Wert (mit Startwert <i>N</i>), der die Position angibt, an der das Zeichen gefunden wurde. Wenn das Zeichen nicht gefunden wurde, ergibt die Funktion den Wert 0. Wenn die Funktion ein ungültiges Offset (<i>N</i>) aufweist (z. B. ein Offset, das außerhalb der Länge der Zeichenfolge liegt), ergibt die Funktion \$null\$. Beispiel: <i>locchar(`n`, 2, Webseite)</i> durchsucht das Feld <i>Webseite</i> nach dem Zeichen <i>`n`</i> und beginnt mit der Suche beim zweiten Zeichen im Feldwert. <i>Hinweis:</i> Achten Sie darauf, das angegebene Zeichen in einzelne umgekehrte Anführungszeichen einzuschließen.

Tabelle 31. CLEM-Zeichenfolgefunktionen (Forts.)

Funktion	Ergebnis	Beschreibung
locchar_back(ZEICHEN, N, ZEICHENFOLGE)	<i>ganze Zahl</i>	Ähnlich wie locchar, außer dass die Suche ausgehend vom N-ten Zeichen rückwärts durchgeführt wird. Beispiel: occhar_back(`n`, 9, Webseite) durchsucht das Feld <i>Webseite</i> ausgehend vom 9. Zeichen rückwärts bis zum Anfang der Zeichenfolge. Wenn die Funktion ein ungültiges Offset aufweist, (z. B. ein Offset, das außerhalb der Länge der Zeichenfolge liegt), ergibt die Funktion \$null\$. Idealerweise sollte locchar_back in Verbindung mit der Funktion length(<Field>) verwendet werden, um die Länge des aktuellen Werts des Felds dynamisch zu verwenden. Beispiel: locchar_back(`n`, (length(Webseite)), Webseite).
lowertoupper(ZEICHEN) lowertoupper (ZEICHENFOLGE)	<i>ZEICHEN oder Zeichenfolge</i>	Die Eingabe kann entweder eine Zeichenfolge oder ein Zeichen sein. Diese wird in dieser Funktion verwendet, um ein neues Element desselben Typs auszugeben, wobei alle Kleinbuchstaben in ie entsprechenden Großbuchstaben konvertiert wurden. Beispiele für gültige Ausdrücke sind: lowertoupper(`a`), lowertoupper("Meine Zeichenfolge") und lowertoupper(feldname(2)).
matches	<i>boolesch</i>	Ergibt den Wert "wahr", wenn eine Zeichenfolge mit einem bestimmten Muster übereinstimmt. Das Muster muss aus einer Literalzeichenfolge bestehen; der Name eines Felds, das ein Muster enthält, ist nicht zulässig. Das Fragezeichen (?) als Platzhalterzeichen kann die Stelle von genau einem beliebigen Zeichen einnehmen, der Stern (*) entspricht keinem, einem oder mehreren Zeichen. Soll ein Fragezeichen oder ein Stern nicht als Platzhalter fungieren, sondern bei der Suche berücksichtigt werden, geben Sie den umgekehrten Schrägstrich (\) als Escapezeichen ein.
replace(TEILZEICHENFOLGE, NEUE_TEILZEICHENFOLGE, ZEICHENFOLGE)	<i>Zeichenfolge</i>	Innerhalb der angegebenen ZEICHENFOLGE werden alle Instanzen von TEILZEICHENFOLGE durch NEUE_TEILZEICHENFOLGE ersetzt.
replicate(ANZAHL, ZEICHENFOLGE)	<i>Zeichenfolge</i>	Ergibt eine Zeichenfolge, die aus der ursprünglichen Zeichenfolge besteht, die so oft wie angegeben kopiert wurde.
stripchar(ZEICHEN,ZEICHENFOLGE)	<i>Zeichenfolge</i>	Mit dieser Funktion können Sie die angegebenen Zeichen aus einer Zeichenfolge oder einem Feld entfernen. Mit dieser Funktion können Sie beispielsweise zusätzliche Symbole, wie Währungsangaben, aus den Daten entfernen, um eine einfache Zahl oder einen einfachen Namen zu erhalten. Mit der Syntax stripchar(`€`, 'Kosten') beispielsweise erhalten Sie ein neues Feld, bei dem das Euro-Zeichen aus allen Werten entfernt ist. <i>Hinweis:</i> Achten Sie darauf, das angegebene Zeichen in einzelne umgekehrte Anführungszeichen einzuschließen.

Tabelle 31. CLEM-Zeichenfolgefunktionen (Forts.)

Funktion	Ergebnis	Beschreibung
skipchar(ZEICHEN, N, ZEICHENFOLGE)	ganze Zahl	Durchsucht die Zeichenfolge <i>ZEICHENFOLGE</i> nach allen Zeichen außer <i>ZEICHEN</i> und beginnt dabei beim <i>N</i> . Zeichen. Diese Funktion ergibt eine ganzzahlige Teilzeichenfolge, die den Punkt angibt, an dem ein solches Zeichen gefunden wurde, oder 0, wenn es sich bei jedem Zeichen vom <i>N</i> . Zeichen an um <i>ZEICHEN</i> handelt. Wenn die Funktion ein ungültiges Offset aufweist, (z. B. ein Offset, das außerhalb der Länge der Zeichenfolge liegt), ergibt die Funktion \$null\$. <i>locchar</i> wird häufig in Verbindung mit den Funktionen vom Typ <i>skipchar</i> verwendet, um den Wert von <i>N</i> (die Stelle, an der mit der Suche begonnen werden soll) zu bestimmen. Beispiel: <code>skipchar(`s`, (locchar(`s`, 1, "MeinString")), "MeinString")</code> .
skipchar_back(ZEICHEN, N, ZEICHENFOLGE)	ganze Zahl	Ähnlich wie <i>skipchar</i> , außer dass die Suche ausgehend vom <i>N</i> -ten Zeichen rückwärts durchgeführt wird.
startstring(LÄNGE, ZEICHENFOLGE)	Zeichenfolge	Extrahiert die ersten <i>N</i> Zeichen aus der angegebenen Zeichenfolge. Ist eine Zeichenfolge kleiner-gleich der angegebenen Länge, dann bleibt die Zeichenfolge unverändert.
strmember(ZEICHEN, ZEICHENFOLGE)	ganze Zahl	Entspricht <code>locchar(ZEICHEN, 1, ZEICHENFOLGE)</code> . Ergibt eine ganzzahlige Teilzeichenfolge, die entweder den Punkt angibt, an dem <i>CHAR</i> zum ersten Mal auftritt, oder den Wert 0. Wenn die Funktion ein ungültiges Offset aufweist (z. B. ein Offset, das außerhalb der Länge der Zeichenfolge liegt), ergibt diese Funktion \$null\$.
subscrs(N, ZEICHENFOLGE)	ZEICHEN	Ergibt das <i>N</i> . Zeichen <i>ZEICHEN</i> der Eingabezeichenfolge <i>ZEICHENFOLGE</i> . Diese Funktion kann auch in einer abgekürzten Form geschrieben werden, nämlich als <code>ZEICHENFOLGE(N)</code> . Der Ausdruck <code>lowertoupper("name"(1))</code> ist beispielsweise ein gültiger Ausdruck.
substring(N, LÄNGE, ZEICHENFOLGE)	Zeichenfolge	Ergibt die Zeichenfolge <i>TEILZEICHENFOLGE</i> , die aus den <i>LÄNGE</i> Zeichen der Zeichenfolge <i>ZEICHENFOLGE</i> besteht, wobei mit dem Zeichen bei Index <i>N</i> begonnen wird.
substring_between(N1, N2, ZEICHENFOLGE)	Zeichenfolge	Ergibt die Teilzeichenfolge von <i>ZEICHENFOLGE</i> , die bei Index <i>N1</i> beginnt und bei Index <i>N2</i> endet.
trim(ZEICHENFOLGE)	Zeichenfolge	Entfernt führende und nachgestellte Leerzeichen aus der angegebenen Zeichenfolge.
trim_start(ZEICHENFOLGE)	Zeichenfolge	Entfernt führende Leerzeichen aus der angegebenen Zeichenfolge.
trimend(ZEICHENFOLGE)	Zeichenfolge	Entfernt nachgestellte Leerzeichen aus der angegebenen Zeichenfolge.
unicode_char(ZAHL)	ZEICHEN	Die Eingabe muss aus Dezimalwerten, nicht Hexadezimalwerten, bestehen. Ergibt das Zeichen mit dem Unicode-Wert von <i>ZAHL</i> .

Tabelle 31. CLEM-Zeichenfolgefunktionen (Forts.)

Funktion	Ergebnis	Beschreibung
unicode_value(ZEICHEN)	ZAHL	Ergibt den Unicode-Wert für das Zeichen ZEICHEN.
uppertolower (ZEICHEN) uppertolower (ZEICHENFOLGE)	ZEICHEN oder Zeichenfolge	Die Eingabe kann entweder eine Zeichenfolge oder ein Zeichen sein. Diese wird in dieser Funktion verwendet, um ein neues Element desselben Typs auszugeben, wobei alle Großbuchstaben in die entsprechenden Kleinbuchstaben konvertiert wurden. <i>Hinweis:</i> Achten Sie darauf, Zeichenfolgen in doppelte Anführungszeichen und Zeichen in einzelne, umgekehrte Anführungszeichen zu setzen. Einzelne Feldnamen sollten ohne Anführungszeichen angegeben werden.

SoundEx-Funktionen

SoundEx ist eine Methode zum Auffinden von Zeichenfolgen, deren Klang bekannt ist, nicht jedoch deren genaue Schreibweise. Die Methode wurde im Jahr 1918 entwickelt. Hiermit werden Wörter mit ähnlichen Lauten ermittelt und zwar auf der Grundlage phonetischer Annahmen, wie bestimmte Buchstaben ausgesprochen werden. Suchen Sie beispielsweise Namen in einer Datenbank, bei denen die Schreibweise und die Betonung ähnlicher Namen voneinander abweichen können. Der grundlegende SoundEx-Algorithmus ist in zahlreichen Schriften dokumentiert. Trotz bekannter Einschränkungen (führende Buchstaben und Buchstabenfolgen wie ph und f gelten beispielsweise nicht als Übereinstimmung, obwohl sie gleich klingen) wird diese Methode in den meisten Datenbanken in irgendeiner Form genutzt.

Tabelle 32. Soundex-Funktionen von CLEM

Funktion	Ergebnis	Beschreibung
soundex(ZEICHENFOLGE)	ganze Zahl	Ergibt den vierbuchstabigen SoundEx-Code für die angegebene ZEICHENFOLGE.
soundex_difference(ZEICHENFOLGE1, ZEICHENFOLGE2)	ganze Zahl	Ergibt eine ganze Zahl zwischen 0 und 4, aus der die Anzahl der Zeichen hervorgeht, die im SoundEx-Code für die beiden Zeichenfolgen übereinstimmen. Der Wert 0 bedeutet dabei, dass gar keine Ähnlichkeit zwischen den Zeichenfolgen vorliegt, der Wert 4 weist auf sehr ähnliche oder sogar identische Zeichenfolgen hin.

Datums- und Uhrzeitfunktionen

CLEM beinhaltet eine Gruppe von Funktionen zum Umgang mit Feldern mit dem Speichertyp "Datum/ Uhrzeit" von Zeichenfolgevariablen, die für Datums- und Zeitangaben stehen. Die Formate für Datum und Uhrzeit sind für jeden Stream spezifisch und werden im Dialogfeld "Streameigenschaften" angegeben. Die Datums- und Zeitfunktionen analysieren Datums- und Zeitzeichenfolgen gemäß dem aktuell ausgewählten Format.

Wenn Sie ein Jahr in einem Datumsformat angeben, bei dem nur zwei Ziffern verwendet werden (d. h., das Jahrhundert ist nicht angegeben), verwendet IBM SPSS Modeler das im Dialogfeld "Streameigenschaften" angegebene Standardjahrhundert.

Anmerkung: Wenn die Datenfunktion per Pushback an SQL oder IBM SPSS Analytic Server zurückübertragen wird, müssen in einer Verzweigung, die auf eine Analytic Server-Datenquelle folgt, alle Datumsformatzeichenfolgen (to_date) innerhalb dieser Daten mit dem Datumsformat übereinstimmen, das im SPSS Modeler-Stream angegeben ist.

Tabelle 33. CLEM-Datums- und Uhrzeitfunktionen.

Funktion	Ergebnis	Beschreibung
@TODAY	Zeichenfolge	Bei Auswahl von Tage/Minuten übertragen im Dialogfeld "Streameigenschaften" ergibt diese Funktion das aktuelle Datum als Zeichenfolge im aktuellen Datumsformat. Wenn Sie ein zweistelliges Datumsformat verwenden und nicht die Option Tage/Minuten übertragen verwenden, ergibt diese Funktion \$null\$ auf dem aktuellen Server.
to_time(ELEMENT)	Zeit	Konvertiert den Speichertyp des angegebenen Felds in eine Zeit.
to_date(ELEMENT)	Datum	Konvertiert den Speichertyp des angegebenen Felds in ein Datum.
to_timestamp(ELEMENT)	Zeitmarke	Konvertiert den Speichertyp des angegebenen Felds in eine Zeitmarke.
to_datetime(ELEMENT)	Datum/Uhrzeit	Konvertiert den Speichertyp des angegebenen Felds in einen Wert vom Typ "Datum", "Zeit" oder "Zeitmarke".
datetime_date(ELEMENT)	Datum	Gibt den Datumswert für eine <i>Zahl</i> , eine <i>Zeichenfolge</i> oder <i>Zeitmarke</i> zurück. Beachten Sie, dass dies die einzige Funktion ist, mit der Sie eine Zahl (in wenigen Sekunden) wieder in ein Datum zurückkonvertieren können. Wenn es sich bei ELEMENT um eine Zeichenfolge handelt, wird ein Datum erstellt, indem eine Zeichenfolge im aktuellen Datumsformat analysiert wird. Das im Dialogfeld für die Streameigenschaften angegebene Datumsformat muss korrekt sein, damit diese Funktion erfolgreich ausgeführt werden kann. Wenn es sich bei ELEMENT um eine Zahl handelt, wird sie als Anzahl von Sekunden seit dem Basisdatum (Epoche) interpretiert. Die Bruchteile eines Tages werden gekürzt. Wenn es sich bei ELEMENT um eine Zeitmarke handelt, wird der Datumsteil der Zeitmarke ausgegeben. Wenn es sich bei ELEMENT um ein Datum handelt, wird es unverändert ausgegeben.
date_before(DATUM1, DATUM2)	boolesch	Ergibt den Wert "Wahr", wenn <i>DATUM1</i> ein Datum oder eine Zeitmarke darstellt, das/der vor dem durch <i>DATUM2</i> dargestellten Datum liegt. Andernfalls ergibt die Funktion den Wert 0.
date_days_difference(DATUM1, DATUM2)	ganze Zahl	Ergibt die Zeit in Tagen (als ganze Zahl), die zwischen dem durch <i>DATUM1</i> und dem durch <i>DATUM2</i> angegebenen Datum bzw. der Zeitmarke liegt. Wenn <i>DATUM2</i> vor <i>DATUM1</i> liegt, ergibt diese Funktion eine negative Zahl.
date_in_days(DATUM)	ganze Zahl	Ergibt die Zeit in Tagen (als ganze Zahl), die zwischen dem Basisdatum und dem durch <i>DATUM</i> angegebenen Datum oder der Zeitmarke liegt. Wenn <i>DATUM</i> vor dem Basisdatum liegt, ergibt diese Funktion eine negative Zahl. Es muss ein gültiges Datum angegeben werden, damit die Berechnung ordnungsgemäß erfolgen kann. Beispielsweise sollten Sie nicht den 29. Februar 2001 als Datum angeben. Da 2001 kein Schaltjahr war, gibt es dieses Datum nicht.

Tabelle 33. CLEM-Datums- und Uhrzeitfunktionen (Forts.).

Funktion	Ergebnis	Beschreibung
date_in_months(DATUM)	reelle Zahl	Ergibt die Zeit in Monaten (als reelle Zahl), die zwischen dem Basisdatum und dem durch <i>DATUM</i> angegebenen Datum oder der Zeitmarke liegt. Dies ist ein Näherungswert, der auf einem Monat mit 30.4375 Tagen beruht. Wenn <i>DATUM</i> vor dem Basisdatum liegt, ergibt diese Funktion eine negative Zahl. Es muss ein gültiges Datum angegeben werden, damit die Berechnung ordnungsgemäß erfolgen kann. Beispielsweise sollten Sie nicht den 29. Februar 2001 als Datum angeben. Da 2001 kein Schaltjahr war, gibt es dieses Datum nicht.
date_in_weeks(DATUM)	reelle Zahl	Ergibt die Zeit in Wochen (als reelle Zahl), die zwischen dem Basisdatum und dem durch <i>DATUM</i> angegebenen Datum oder der Zeitmarke liegt. Der Wert basiert auf einer Woche mit 7,0 Tagen. Wenn <i>DATUM</i> vor dem Basisdatum liegt, ergibt diese Funktion eine negative Zahl. Es muss ein gültiges Datum angegeben werden, damit die Berechnung ordnungsgemäß erfolgen kann. Beispielsweise sollten Sie nicht den 29. Februar 2001 als Datum angeben. Da 2001 kein Schaltjahr war, gibt es dieses Datum nicht.
date_in_years(DATUM)	reelle Zahl	Ergibt die Zeit in Jahren (als reelle Zahl), die zwischen dem Basisdatum und dem durch <i>DATUM</i> angegebenen Datum oder der Zeitmarke liegt. Dies ist ein Näherungswert, der auf einem Jahr mit 365.25 Tagen beruht. Wenn <i>DATUM</i> vor dem Basisdatum liegt, ergibt diese Funktion eine negative Zahl. Es muss ein gültiges Datum angegeben werden, damit die Berechnung ordnungsgemäß erfolgen kann. Beispielsweise sollten Sie nicht den 29. Februar 2001 als Datum angeben. Da 2001 kein Schaltjahr war, gibt es dieses Datum nicht.
date_months_difference (DATUM1, DATUM2)	reelle Zahl	Ergibt die Zeit in Monaten (als reelle Zahl), die zwischen dem durch <i>DATUM1</i> und dem durch <i>DATUM2</i> angegebenen Datum bzw. der Zeitmarke liegt. Dies ist ein Näherungswert, der auf einem Monat mit 30.4375 Tagen beruht. Wenn <i>DATUM2</i> vor <i>DATUM1</i> liegt, ergibt diese Funktion eine negative Zahl.
datetime_date(JAHR, MONAT, TAG)	Datum	Erstellt einen Datumswert für die angegebenen Werte für <i>JAHR</i> , <i>MONAT</i> und <i>TAG</i> . Bei den Argumenten muss es sich um ganze Zahlen handeln.
datetime_day(DATUM)	ganze Zahl	Ergibt den Tag im Monat ausgehend von einem angegebenen <i>DATUM</i> oder einer Zeitmarke. Das Ergebnis ist eine ganze Zahl im Bereich von 1 bis 31.
datetime_day_name(TAG)	Zeichenfolge	Ergibt den vollständigen Namen für den angegebenen <i>TAG</i> . Das Argument muss eine ganze Zahl im Bereich von 1 (Sonntag) bis 7 (Samstag) sein.
datetime_hour(ZEIT)	ganze Zahl	Ergibt die Stunde aus einer <i>ZEIT</i> oder einer Zeitmarke. Das Ergebnis ist eine ganze Zahl im Bereich von 0 bis 23.
datetime_in_seconds(ZEIT)	reelle Zahl	Ergibt den in <i>ZEIT</i> gespeicherten Sekundenabschnitt.
datetime_in_seconds(DATUM), datetime_in_seconds(DATUMZEIT)	reelle Zahl	Ergibt die aus der Differenz zwischen dem aktuellen <i>DATUM</i> oder <i>DATUM/UHRZEIT</i> und dem Basisdatum (1900-01-01) berechneten Gesamtwert (umgewandelt in Sekunden).
datetime_minute(ZEIT)	ganze Zahl	Ergibt die Minute aus einer <i>ZEIT</i> oder einer Zeitmarke. Das Ergebnis ist eine ganze Zahl im Bereich von 0 bis 59.

Tabelle 33. CLEM-Datums- und Uhrzeitfunktionen (Forts.).

Funktion	Ergebnis	Beschreibung
datetime_month(DATUM)	ganze Zahl	Ergibt den Monat aus einem <i>DATUM</i> oder einer Zeitmarke. Das Ergebnis ist eine ganze Zahl im Bereich von 1 bis 12.
datetime_month_name (MONAT)	Zeichenfolge	Ergibt den vollständigen Namen für den angegebenen <i>MONAT</i> . Das Argument muss eine ganze Zahl im Bereich von 1 bis 12 sein.
datetime_now	Zeitmarke	Ergibt die aktuelle Zeit als Zeitmarke.
datetime_second(ZEIT)	ganze Zahl	Ergibt die Sekunde aus einer <i>ZEIT</i> oder einer Zeitmarke. Das Ergebnis ist eine ganze Zahl im Bereich von 0 bis 59.
datetime_day_short_name (TAG)	Zeichenfolge	Ergibt den abgekürzten Namen für den angegebenen <i>TAG</i> . Das Argument muss eine ganze Zahl im Bereich von 1 (Sonntag) bis 7 (Samstag) sein.
datetime_month_short_name (MONAT)	Zeichenfolge	Ergibt den abgekürzten Namen für den angegebenen <i>MONAT</i> . Das Argument muss eine ganze Zahl im Bereich von 1 bis 12 sein.
datetime_time(STUNDE, MINUTE, SEKUNDE)	Zeit	Ergibt den Zeitwert für die angegebene <i>STUNDE</i> , <i>MINUTE</i> und <i>SEKUNDE</i> . Bei den Argumenten muss es sich um ganze Zahlen handeln.
datetime_time(ELEMENT)	Zeit	Ergibt den Zeitwert für das angegebene <i>ELEMENT</i> .
datetime_timestamp(JAHR, MONAT, TAG, STUNDE, MINUTE, SEKUNDE)	Zeitmarke	Ergibt den Zeitmarkenwert für die angegebenen Werte für <i>JAHR</i> , <i>MONAT</i> , <i>TAG</i> , <i>STUNDE</i> , <i>MINUTE</i> und <i>SEKUNDE</i> .
datetime_timestamp(DATUM, ZEIT)	Zeitmarke	Ergibt den Zeitmarkenwert für das angegebene <i>DATUM</i> und die angegebene <i>ZEIT</i> .
datetime_timestamp (ZAHL)	Zeitmarke	Ergibt den Zeitmarkenwert für die angegebene Anzahl von Sekunden.
datetime_weekday(DATUM)	ganze Zahl	Ergibt den Tag in der Woche <i>t</i> ausgehend von einem angegebenen <i>DATUM</i> oder einer Zeitmarke.
datetime_year(DATUM)	ganze Zahl	Ergibt das Jahr aus einem <i>DATUM</i> oder einer Zeitmarke. Das Ergebnis ist eine ganze Zahl, wie beispielsweise 2002.
date_weeks_difference (DATUM1, DATUM2)	reelle Zahl	Ergibt die Zeit in Wochen (als reelle Zahl), die zwischen dem durch <i>DATUM1</i> und dem durch <i>DATUM2</i> angegebenen Datum bzw. der Zeitmarke liegt. Der Wert basiert auf einer Woche mit 7,0 Tagen. Wenn <i>DATUM2</i> vor <i>DATUM1</i> liegt, ergibt diese Funktion eine negative Zahl.
date_years_difference (DATUM1, DATUM2)	reelle Zahl	Ergibt die Zeit in Jahren (als reelle Zahl), die zwischen dem durch <i>DATUM1</i> und dem durch <i>DATUM2</i> angegebenen Datum bzw. der Zeitmarke liegt. Dies ist ein Näherungswert, der auf einem Jahr mit 365,25 Tagen beruht. Wenn <i>DATUM2</i> vor <i>DATUM1</i> liegt, ergibt diese Funktion eine negative Zahl.
date_from_ywd(JAHR, WOCHE, TAG)	ganze Zahl	Wandelt das Jahr, die Woche im Jahr und den Tag in der Woche in ein Datum im Standard ISO 8601 um.
date_iso_day(DATUM)	ganze Zahl	Ergibt den Tag in der Woche aus dem Datum im Standard ISO 8601.
date_iso_week(DATUM)	ganze Zahl	Ergibt die Woche des Jahres aus dem Datum im Standard ISO 8601.
date_iso_year(DATUM)	ganze Zahl	Ergibt das Jahr aus dem Datum im Standard ISO 8601.

Tabelle 33. CLEM-Datums- und Uhrzeitfunktionen (Forts.).

Funktion	Ergebnis	Beschreibung
time_before(ZEIT1, ZEIT2)	boolesch	Ergibt den Wert "Wahr", wenn ZEIT1 eine Uhrzeit oder eine Zeitmarke darstellt, die/der vor der durch ZEIT2 dargestellten Zeit liegt. Andernfalls ergibt die Funktion den Wert 0.
time_hours_difference (ZEIT1, ZEIT2)	reelle Zahl	Ergibt den Zeitunterschied in Stunden (als reelle Zahl) zwischen den durch ZEIT1 und ZEIT2 angegebenen Uhrzeiten oder Zeitmarken. Bei Auswahl von Tage/Minuten übertragen im Dialogfeld "Streameigenschaften" wird ein höherer Wert von ZEIT1 als Verweis auf den vorhergehenden Tag interpretiert. Wenn die Übertragungsoption nicht ausgewählt wird, führt ein höherer Wert von ZEIT1 dazu, dass der ausgegebene Wert negativ ist.
time_in_hours(ZEIT)	reelle Zahl	Ergibt die durch ZEIT angegebene Zeit in Stunden als reelle Zahl. So wird beispielsweise beim Zeitformat HHMM der Ausdruck time_in_hours('0130') als 1,5 ausgewertet. ZEIT kann für eine Uhrzeit oder eine Zeitmarke stehen.
time_in_mins(ZEIT)	reelle Zahl	Ergibt die durch ZEIT angegebene Zeit in Minuten als reelle Zahl. ZEIT kann für eine Uhrzeit oder eine Zeitmarke stehen.
time_in_secs(ZEIT)	ganze Zahl	Ergibt die durch ZEIT angegebene Zeit in Sekunden als ganze Zahl. ZEIT kann für eine Uhrzeit oder eine Zeitmarke stehen.
time_mins_difference(ZEIT1, ZEIT2)	reelle Zahl	Ergibt den Zeitunterschied in Minuten (als reelle Zahl) zwischen den durch ZEIT1 und ZEIT2 angegebenen Uhrzeiten oder Zeitmarken. Bei Auswahl von Tage/Minuten übertragen im Dialogfeld "Streameigenschaften" wird ein höherer Wert von ZEIT1 als Verweis auf den vorhergehenden Tag (bzw. die vorangegangene Stunde, sofern nur Minuten und Sekunden im aktuellen Format angegeben werden) interpretiert. Wenn die Übertragungsoption nicht ausgewählt wird, führt ein höherer Wert von ZEIT1 dazu, dass der ausgegebene Wert negativ ist.
time_secs_difference(ZEIT1, ZEIT2)	ganze Zahl	Ergibt den Zeitunterschied in Sekunden (als ganze Zahl) zwischen den durch ZEIT1 und ZEIT2 angegebenen Uhrzeiten oder Zeitmarken. Bei Auswahl von Tage/Minuten übertragen im Dialogfeld "Streameigenschaften" wird ein höherer Wert von ZEIT1 als Verweis auf den vorhergehenden Tag (bzw. die vorangegangene Stunde, sofern nur Minuten und Sekunden im aktuellen Format angegeben werden) interpretiert. Wenn die Übertragungsoption nicht ausgewählt wird, führt ein höherer Wert von ZEIT1 dazu, dass der ausgegebene Wert negativ ist.

Konvertieren von Datums- und Zeitwerten

Beachten Sie, dass die Konvertierungsfunktionen (und alle anderen Funktionen, für die ein spezieller Eingabetyp, wie beispielsweise ein Wert für Datum oder Uhrzeit, erforderlich ist) von den aktuell im Dialogfeld für die Streamoptionen angegebenen Formaten abhängen. Ein Feld mit der Bezeichnung *DATUM* beispielsweise, das als Zeichenfolge mit den Werten *Jan 2003*, *Feb 2003* usw. gespeichert ist, könnte wie folgt in einen Datumsspeicher konvertiert werden:

```
to_date(DATUM)
```

Damit diese Konvertierung funktionieren kann, müssen Sie das entsprechende Datumsformat **MON JJJJ** als Standarddatumsformat für den Stream festlegen.

Ein Beispiel, bei dem Zeichenfolgewerte mithilfe eines Füllerknotens in Datumswerte konvertiert werden, finden Sie im Stream *broadband_create_models.str*, der im Ordner *\Demos* unter dem Unterordner *streams* installiert ist.

Als Zahlen gespeicherte Datumsangaben. Beachten Sie, dass *DATUM* im vorherigen Beispiel der Name eines Felds ist, während es sich bei *to_date* um eine CLEM-Funktion handelt. Wenn Datumsangaben als Zahlen gespeichert wurden, können Sie sie mithilfe der Funktion *datetime_date* konvertieren. Dabei wird die Zahl als Anfang von Sekunden seit dem Basisdatum (Epoche) interpretiert.

```
datetime_date(DATUM)
```

Indem Sie ein Datum in eine Sekundenzahl (und zurück) konvertieren, können Sie Berechnungen wie die Addition oder Subtraktion einer Anzahl von Tagen zum bzw. vom aktuellen Datum durchführen. Beispiel:

```
datetime_date((date_in_days(DATUM)-7)*60*60*24)
```

Sequenzfunktionen

Bei einigen Operationen ist die Abfolge (Sequenz) der Ereignisse von Bedeutung. Die Anwendung ermöglicht die Arbeit mit folgenden Datensatzsequenzen:

- Sequenzen und Zeitreihen
- Sequenzfunktionen
- Datensatzindizierung
- Durchschnittsbildung, Summierung und Vergleich von Werten
- Überwachen von Änderungen - Differenzierung
- @SINCE
- Offset-Werte
- Zusätzliche Sequenzfunktionen

Bei vielen Anwendungen kann jeder Datensatz, der einen Stream durchläuft, als Einzelfall betrachtet werden, der von allen anderen Fällen unabhängig ist. In solchen Situationen ist die Reihenfolge der Datensätze normalerweise nicht von Bedeutung.

Bei einigen Problemklassen jedoch ist die Datensatzsequenz höchst wichtig. Dies ist üblicherweise bei Zeitreihen der Fall, bei denen die Sequenz der Datensätze eine geordnete Sequenz von Ereignissen oder Vorkommen darstellt. Jeder Datensatz ist eine Momentaufnahme zu einem bestimmten Zeitpunkt; ein Großteil der aussagekräftigsten Informationen sind jedoch möglicherweise nicht in Augenblickswerten enthalten, sondern in der Art und Weise, wie sich diese Werte im Laufe der Zeit ändern und verhalten.

Der entscheidende Parameter muss natürlich nicht die Zeit sein. So könnten die Datensätze Analysen darstellen, die in Abständen entlang einer Linie durchgeführt werden. Doch auch hier gelten dieselben Prinzipien.

Sequenz- und Sonderfunktionen lassen sich sofort an folgenden Merkmalen erkennen:

- Sie enthalten alle das Präfix @.
- Die Namen dieser Funktionen bestehen aus Großbuchstaben.

Sequenzfunktionen können sich auf den derzeit von einem Knoten verarbeiteten Datensatz, auf die Datensätze, die einen Knoten bereits durchlaufen haben, und sogar, in einem Fall, auf Datensätze beziehen, die einen Knoten noch durchlaufen müssen. Sequenzfunktionen können nach Belieben mit anderen Komponenten von CLEM-Ausdrücken gemischt werden; bei einigen gelten jedoch Einschränkungen in Bezug auf die verwendeten Argumente.

Beispiele

Es kann interessant sein zu wissen, wie lange es zurückliegt, dass ein bestimmtes Ereignis eintrat oder eine Bedingung wahr war. Dies kann mit der Funktion @SINCE ermittelt werden. Beispiel:

```
@SINCE(Einkommen > Ausgben)
```

Diese Funktion ergibt das Offset des letzten Datensatzes, bei dem diese Bedingung erfüllt war, also die Anzahl der Datensätze vor dem aktuellen Datensatz, bei denen die Bedingung wahr war. Wenn die Bedingung niemals wahr war, ergibt @SINCE den Wert @INDEX + 1.

Manchmal möchte man sich in dem von @SINCE verwendeten Ausdruck auf einen Wert des aktuellen Datensatzes beziehen. Dies ist mit der Funktion @THIS möglich, die angibt, dass ein Feldname immer für den aktuellen Datensatz gilt. Um den Offset des letzten Datensatzes zu ermitteln, bei dem der Wert des Felds Konzentration mehr als zweimal so hoch war wie der des aktuellen Datensatzes, können Sie folgende Funktion verwenden:

```
@SINCE(Konzentration > 2 * @THIS(Konzentration))
```

In einigen Fällen ist die für @SINCE angegebene Bedingung beim aktuellen Datensatz per definitionem wahr. Beispiel:

```
@SINCE(ID == @THIS(ID))
```

Aus diesem Grund wertet @SINCE die Bedingung nicht für den aktuellen Datensatz aus. Mit einer ähnlichen Funktion @SINCE0 können Sie die Bedingung für den aktuellen Datensatz und die vorangegangenen auswerten; wenn die Bedingung im aktuellen Datensatz wahr ist, ergibt @SINCE0 den Wert 0.

Tabelle 34. CLEM-Sequenzfunktionen

Funktion	Ergebnis	Beschreibung
MEAN(FELD)	reelle Zahl	Ergibt den mittleren Durchschnitt der Werte für das angegebene FELD bzw. die angegebenen FELDER.
@MEAN(FELD, AUSDR)	reelle Zahl	Ergibt den mittleren Durchschnitt der Werte für FELD über die letzten AUSDR-Datensätze, die vom aktuellen Knoten erhalten wurden, einschließlich des aktuellen Datensatzes. FELD muss der Name eines numerischen Felds sein. AUSDR kann jeder Ausdruck sein, dessen Auswertung eine ganze Zahl größer als 0 ergibt. Wenn AUSDR weggelassen wird oder die Anzahl der bisher erhaltenen Datensätze übersteigt, ergibt sich der Durchschnitt aus den bisher erhaltenen Datensätzen.
@MEAN(FELD, AUSDR, GANZZ)	reelle Zahl	Ergibt den mittleren Durchschnitt der Werte für FELD über die letzten AUSDR-Datensätze, die vom aktuellen Knoten erhalten wurden, einschließlich des aktuellen Datensatzes. FELD muss der Name eines numerischen Felds sein. AUSDR kann jeder Ausdruck sein, dessen Auswertung eine ganze Zahl größer als 0 ergibt. Wenn AUSDR weggelassen wird oder die Anzahl der bisher erhaltenen Datensätze übersteigt, ergibt sich der Durchschnitt aus den bisher erhaltenen Datensätzen. GANZZ gibt die maximale Anzahl von zurückliegenden Werten an, die betrachtet werden sollen. Dies ist wesentlich effizienter als die Verwendung von nur zwei Argumenten.
@DIFF1(FELD)	reelle Zahl	Ergibt das erste Differenzial von FELD. Das Format mit einem einzelnen Argument ergibt also einfach die Differenz zwischen dem aktuellen Wert und dem früheren Wert des Feldes. Ergibt \$null\$, wenn die relevanten vorhergehenden Datensätze nicht vorhanden sind.

Tabelle 34. CLEM-Sequenzfunktionen (Forts.)

Funktion	Ergebnis	Beschreibung
@DIFF1(FELD1, FELD2)	reelle Zahl	Das Format mit zwei Argumenten ergibt das erste Differenzial von <i>FELD1</i> in Bezug auf <i>FELD2</i> . Ergibt \$null\$, wenn die relevanten vorhergehenden Datensätze nicht vorhanden sind. Wird berechnet als @DIFF1(FELD1)/@DIFF1(FELD2).
@DIFF2(FELD)	reelle Zahl	Ergibt das zweite Differenzial von <i>FELD</i> . Das Format mit einem einzelnen Argument ergibt also einfach die Differenz zwischen dem aktuellen Wert und dem früheren Wert des Feldes. Ergibt \$null\$, wenn die relevanten vorhergehenden Datensätze nicht vorhanden sind. @DIFF2 wird berechnet als @DIFF(@DIFF(FELD)).
@DIFF2(FELD1, FELD2)	reelle Zahl	Das Format mit zwei Argumenten ergibt das zweite Differenzial von <i>FELD1</i> in Bezug auf <i>FELD2</i> . Ergibt \$null\$, wenn die relevanten vorhergehenden Datensätze nicht vorhanden sind. Dies ist eine komplexe Berechnung - @DIFF1(FELD1)/@DIFF1(FELD2) - @OFFSET(@DIFF1(FELD1),1)/@OFFSET(@DIFF1(FELD2)) / @DIFF1(FELD2).
@INDEX	ganze Zahl	Ergibt den Index des aktuellen Datensatzes. Indizes werden den Datensätzen zugeordnet, wenn sie am aktuellen Knoten ankommen. Der erste Datensatz erhält den Index 1. Dieser Index wird für jeden nachfolgenden Datensatz um den Wert 1 erhöht.
@LAST_NON_BLANK(FELD)	beliebig	Ergibt den letzten nicht leeren Wert für <i>FELD</i> , wie in einem vorgeordneten Quellen- oder Typknoten definiert. Wenn in den bisher gelesenen Datensätzen keine nicht leeren Werte für <i>FELD</i> vorhanden sind, wird \$null\$ zurückgegeben. Leerwerte (auch als benutzerdefiniert fehlende Werte bezeichnet) können für jedes Feld einzeln definiert werden.
@MAX(FELD)	Zahl	Ergibt den Höchstwert für das angegebene <i>FELD</i> .
@MAX(FELD, AUSDR)	Zahl	Ergibt den Höchstwert für <i>FELD</i> über die letzten <i>AUSDR</i> -Datensätze, die bisher erhalten wurden, einschließlich des aktuellen Datensatzes. <i>FELD</i> muss der Name eines numerischen Felds sein. <i>AUSDR</i> kann jeder Ausdruck sein, dessen Auswertung eine ganze Zahl größer als 0 ergibt.
@MAX(FELD, AUSDR, GANZZ)	Zahl	Ergibt den Höchstwert für <i>FELD</i> über die letzten <i>AUSDR</i> -Datensätze, die bisher erhalten wurden, einschließlich des aktuellen Datensatzes. <i>FELD</i> muss der Name eines numerischen Felds sein. <i>AUSDR</i> kann jeder Ausdruck sein, dessen Auswertung eine ganze Zahl größer als 0 ergibt. Wenn <i>AUSDR</i> weggelassen wird oder die Anzahl der bisher erhaltenen Datensätze übersteigt, ergibt sich der Höchstwert aus den bisher erhaltenen Datensätzen. <i>GANZZ</i> gibt die maximale Anzahl von zurückliegenden Werten an, die betrachtet werden sollen. Dies ist wesentlich effizienter als die Verwendung von nur zwei Argumenten.
@MIN(FELD)	Zahl	Ergibt den Mindestwert für das angegebene <i>FELD</i> .
@MIN(FELD, AUSDR)	Zahl	Ergibt den Mindestwert für <i>FELD</i> über die letzten <i>AUSDR</i> -Datensätze, die bisher erhalten wurden, einschließlich des aktuellen Datensatzes. <i>FELD</i> muss der Name eines numerischen Felds sein. <i>AUSDR</i> kann jeder Ausdruck sein, dessen Auswertung eine ganze Zahl größer als 0 ergibt.

Tabelle 34. CLEM-Sequenzfunktionen (Forts.)

Funktion	Ergebnis	Beschreibung
@MIN(FELD, AUSDR, GANZZ)	Zahl	<p>Ergibt den Mindestwert für <i>FELD</i> über die letzten <i>AUSDR</i>-Datensätze, die bisher erhalten wurden, einschließlich des aktuellen Datensatzes. <i>FELD</i> muss der Name eines numerischen Felds sein. <i>AUSDR</i> kann jeder Ausdruck sein, dessen Auswertung eine ganze Zahl größer als 0 ergibt. Wenn <i>AUSDR</i> weggelassen wird oder die Anzahl der bisher erhaltenen Datensätze übersteigt, ergibt sich der Mindestwert aus den bisher erhaltenen Datensätzen. <i>GANZZ</i> gibt die maximale Anzahl von zurückliegenden Werten an, die betrachtet werden sollen. Dies ist wesentlich effizienter als die Verwendung von nur zwei Argumenten.</p>
@OFFSET(FELD, AUSDR)	beliebig	<p>Ergibt den Wert von <i>FELD</i> in dem Datensatz, der gegenüber dem aktuellen Datensatz ein Offset mit dem Wert von <i>AUSDR</i> aufweist. Ein positiver Offset bezieht sich auf einen Datensatz, der den Knoten bereits durchlaufen hat ("Lookback"), ein negativer Offset gibt ein "Lookahead" auf einen Datensatz an, der noch nicht angekommen ist. Beispiel: @OFFSET(Status, 1) ergibt den Wert des Felds Status im vorangegangenen Datensatz, während @OFFSET(Status, -4) vier Datensätze nach vorn in der Sequenz blickt (also auf Datensätze, die den Knoten noch nicht durchlaufen haben), um den Wert zu ermitteln. <i>Beachten Sie, dass ein negatives Offset (Lookahead-Offset) als Konstante angegeben sein muss.</i> Bei positiven Offsets gilt Folgendes: Bei <i>AUSDR</i> kann es sich auch um einen beliebigen CLEM-Ausdruck handeln, der für den aktuellen Datensatz ausgewertet wird, um das Offset zu erhalten. In diesem Fall sollte die Version dieser Funktion mit drei Argumenten die Leistung verbessern (siehe nächste Funktion). Wenn der Ausdruck einen anderen Wert als eine nicht negative ganze Zahl zurückgibt, wird ein Fehler ausgelöst. Berechnete Lookahead-Offsets sind also nicht zulässig. <i>Hinweis:</i> Eine selbstreferenzielle @OFFSET-Funktion kann kein Literal-Lookahead verwenden. In einem Füllerknoten beispielsweise kann der Wert von Feld1 nicht mithilfe eines Ausdrucks wie @OFFSET(Feld1, -2) ersetzt werden. <i>Hinweis:</i> Beim Füllen eines Felds gibt es im Füllerknoten zwei verschiedene Werte für das Feld, nämlich den vorher und den nachher eingefüllten Wert. Wenn @OFFSET auf sich selbst verweist, verweist es auf den nachher eingefüllten Wert. Dieser nachher eingefüllte Wert ist nur für zurückliegende Zeilen vorhanden, daher kann ein auf sich selbst verweisendes @OFFSET nur auf zurückliegende Zeilen verweisen. Da ein auf sich selbst verweisendes @OFFSET nicht auf die Zukunft verweisen kann, führt es die folgenden Prüfungen des Offsets aus:</p> <ul style="list-style-type: none"> • Wenn der Offset ein Literal ist und auf die Zukunft verweist, wird ein Fehler gemeldet, bevor die Ausführung beginnt. • Wenn der Offset ein Ausdruck ist und bei der Auswertung zur Laufzeit einen Wert in der Zukunft ergibt, übergibt @OFFSET den Wert \$null\$. <p><i>Hinweis:</i> Die Verwendung von "lookahead" und "lookback" in einem Knoten wird nicht unterstützt.</p>

Tabelle 34. CLEM-Sequenzfunktionen (Forts.)

Funktion	Ergebnis	Beschreibung
@OFFSET(FELD, AUSDR, GANZZ)	beliebig	<p>Führt die gleiche Operation durch wie @OFFSET, allerdings mit einem zusätzlichen dritten Argument, <i>GANZZ</i>, das die maximale Anzahl der zurückliegenden Werte angibt, die betrachtet werden sollen. Wenn das Offset aus einem Ausdruck berechnet wird, sollte das dritte Argument zu einer Verbesserung der Leistung führen.</p> <p>In einem Ausdruck wie @OFFSET(Dings, Monat, 12) beispielsweise wird das System angewiesen, nur die letzten 12 Werte von Dings zu behalten. Andernfalls müsste sicherheitshalber jeder Wert gespeichert werden. In Fällen, in denen es sich beim Offset-Wert um eine Konstante handelt - einschließlich der negativen "Lookahead"-Offsets, die konstant sein müssen - ist das dritte Argument sinnlos, weshalb hier die Version dieser Funktion mit zwei Argumenten verwendet werden sollte. Siehe auch den Hinweis zu selbstreferenziellen Funktionen für die Version mit zwei Argumenten weiter oben.</p> <p><i>Hinweis:</i> Die Verwendung von "lookahead" und "lookback" in einem Knoten wird nicht unterstützt.</p>
@SDEV(FELD)	reelle Zahl	Ergibt die Standardabweichung der Werte für das angegebene <i>FELD</i> oder die angegebenen <i>FELDER</i> .
@SDEV(FELD, AUSDR)	reelle Zahl	Ergibt die Standardabweichung der Werte für <i>FELD</i> über die letzten <i>AUSDR</i> -Datensätze, die vom aktuellen Knoten erhalten wurden, einschließlich des aktuellen Datensatzes. <i>FELD</i> muss der Name eines numerischen Felds sein. <i>AUSDR</i> kann jeder Ausdruck sein, dessen Auswertung eine ganze Zahl größer als 0 ergibt. Wenn <i>AUSDR</i> weggelassen wird oder die Anzahl der bisher erhaltenen Datensätze übersteigt, ergibt sich die Standardabweichung aus allen bisher erhaltenen Datensätzen.
@SDEV(FELD, AUSDR, GANZZ)	reelle Zahl	Ergibt die Standardabweichung der Werte für <i>FELD</i> über die letzten <i>AUSDR</i> -Datensätze, die vom aktuellen Knoten erhalten wurden, einschließlich des aktuellen Datensatzes. <i>FELD</i> muss der Name eines numerischen Felds sein. <i>AUSDR</i> kann jeder Ausdruck sein, dessen Auswertung eine ganze Zahl größer als 0 ergibt. Wenn <i>AUSDR</i> weggelassen wird oder die Anzahl der bisher erhaltenen Datensätze übersteigt, ergibt sich die Standardabweichung aus allen bisher erhaltenen Datensätzen. <i>GANZZ</i> gibt die maximale Anzahl von zurückliegenden Werten an, die betrachtet werden sollen. Dies ist wesentlich effizienter als die Verwendung von nur zwei Argumenten.
@SINCE(AUSDR)	beliebig	Ergibt die Anzahl der Datensätze, die übergeben wurden, seit <i>AUSDR</i> (ein beliebiger CLEM-Ausdruck) wahr war.
@SINCE(AUSDR, GANZZ)	beliebig	Das zweite Argument, <i>GANZZ</i> , gibt die maximale Anzahl von zurückliegenden Werten an, die betrachtet werden sollen. Wenn <i>AUSDR</i> niemals wahr war, ist <i>GANZZ</i> gleich @INDEX+1.
@SINCE0(AUSDR)	beliebig	Berücksichtigt den aktuellen Datensatz, was bei @SINCE nicht der Fall ist; @SINCE0 ergibt 0, wenn <i>AUSDR</i> für den aktuellen Datensatz wahr ist.

Tabelle 34. CLEM-Sequenzfunktionen (Forts.)

Funktion	Ergebnis	Beschreibung
@SINCE0(AUSDR, GANZZ)	beliebig	Das zweite Argument, <i>GANZZ</i> , gibt die maximale Anzahl von zurückliegenden Werten an, die betrachtet werden sollen.
@SUM(FELD)	Zahl	Ergibt die Summe der Werte für das angegebene <i>FELD</i> bzw. die angegeben <i>FELDER</i> .
@SUM(FELD, AUSDR)	Zahl	Ergibt die Summe der Werte für <i>FELD</i> über die letzten <i>AUSDR</i> -Datensätze, die vom aktuellen Knoten erhalten wurden, einschließlich des aktuellen Datensatzes. <i>FELD</i> muss der Name eines numerischen Felds sein. <i>AUSDR</i> kann jeder Ausdruck sein, dessen Auswertung eine ganze Zahl größer als 0 ergibt. Wenn <i>AUSDR</i> weggelassen wird oder die Anzahl der bisher erhaltenen Datensätze übersteigt, ergibt sich die Summe aus den bisher erhaltenen Datensätzen.
@SUM(FELD, AUSDR, GANZZ)	Zahl	Ergibt die Summe der Werte für <i>FELD</i> über die letzten <i>AUSDR</i> -Datensätze, die vom aktuellen Knoten erhalten wurden, einschließlich des aktuellen Datensatzes. <i>FELD</i> muss der Name eines numerischen Felds sein. <i>AUSDR</i> kann jeder Ausdruck sein, dessen Auswertung eine ganze Zahl größer als 0 ergibt. Wenn <i>AUSDR</i> weggelassen wird oder die Anzahl der bisher erhaltenen Datensätze übersteigt, ergibt sich die Summe aus den bisher erhaltenen Datensätzen. <i>GANZZ</i> gibt die maximale Anzahl von zurückliegenden Werten an, die betrachtet werden sollen. Dies ist wesentlich effizienter als die Verwendung von nur zwei Argumenten.
@THIS(FELD)	beliebig	Ergibt den Wert des Felds namens <i>FELD</i> im aktuellen Datensatz. Wird nur bei @SINCE-Ausdrücken verwendet.

Globale Funktionen

Die Funktionen @MEAN,@SUM, @MIN, @MAX und @SDEV werden maximal auf alle Datensätze bis einschließlich des aktuellen Datensatzes angewendet. In einigen Fällen ist es jedoch von Vorteil, in der Lage zu sein, die Daten im aktuellen Datensatz mit Werten im gesamten Dataset zu vergleichen. Mit einem Globalwerteknoten können Sie Werte im gesamten Dataset generieren, auf die Sie dann mit den globalen Funktionen in einem CLEM-Ausdruck zugreifen können.

Beispiel:

@GLOBAL_MAX(Alter)

ergibt den höchsten Wert von Alter im Dataset, während der Ausdruck

(Wert - @GLOBAL_MEAN(Wert)) / @GLOBAL_SDEV(Wert)

die Differenz zwischen dem Wert dieses Datensatzes und dem globalen Mittelwert als Anzahl der Standardabweichungen angibt. Globale Werte können nur verwendet werden, nachdem sie von einem Globalwerteknoten berechnet wurden. Alle aktuellen Globalwerte können durch Klicken auf die Schaltfläche zum Löschen der Globalwerte auf der Registerkarte "Globalwerte" im Dialogfeld "Streameigenschaften" entfernt werden.

Tabelle 35. Globale Funktionen von CLEM

Funktion	Ergebnis	Beschreibung
@GLOBAL_MAX(FELD)	Zahl	Ergibt den Höchstwert für <i>FELD</i> im gesamten Dataset, das zuvor durch einen Globalwerteknoten erstellt wurde. <i>FELD</i> muss der Name eines numerischen Felds, eines Datums-, Uhrzeit- oder Datums-/Uhrzeitfelds oder eines Zeichenfolgefelds sein. Wenn der entsprechende Globalwert nicht festgelegt wurde, wird ein Fehler ausgegeben.
@GLOBAL_MIN(FELD)	Zahl	Ergibt den Höchstwert für <i>FELD</i> im gesamten Dataset, das zuvor durch einen Globalwerteknoten generiert wurde. <i>FELD</i> muss der Name eines numerischen Felds, eines Datums-, Uhrzeit- oder Datums-/Uhrzeitfelds oder eines Zeichenfolgefelds sein. Wenn der entsprechende Globalwert nicht festgelegt wurde, wird ein Fehler ausgegeben.
@GLOBAL_SDEV(FELD)	Zahl	Ergibt die Standardabweichung der Werte für <i>FELD</i> im gesamten Dataset, das zuvor durch einen Globalwerteknoten generiert wurde. <i>FELD</i> muss der Name eines numerischen Felds sein. Wenn der entsprechende Globalwert nicht festgelegt wurde, wird ein Fehler ausgegeben.
@GLOBAL_MEAN(FELD)	Zahl	Ergibt den mittleren Durchschnitt der Werte für <i>FELD</i> im gesamten Dataset, das zuvor durch einen Globalwerteknoten generiert wurde. <i>FELD</i> muss der Name eines numerischen Felds sein. Wenn der entsprechende Globalwert nicht festgelegt wurde, wird ein Fehler ausgegeben.
@GLOBAL_SUM(FELD)	Zahl	Ergibt die Summe der Werte für <i>FELD</i> im gesamten Dataset, das zuvor durch einen Globalwerteknoten generiert wurde. <i>FELD</i> muss der Name eines numerischen Felds sein. Wenn der entsprechende Globalwert nicht festgelegt wurde, wird ein Fehler ausgegeben.

Funktionen zum Umgang mit Leerstellen und Nullwerten

Mit CLEM können Sie angeben, dass bestimmte Werte in einem Feld als "Leerstellen" oder fehlende Werte betrachtet werden sollen. Die nachstehenden Funktionen können auf Leerstellen angewendet werden.

Tabelle 36. CLEM-Funktionen zum Umgang mit Leerstellen und Nullwerten

Funktion	Ergebnis	Beschreibung
@BLANK(FELD)	boolesch	Ergibt den Wert "Wahr" für alle Datensätze, deren Werte gemäß den Regeln zum Umgang mit Leerstellen, die in einem vorgeordneten Typknoten oder Quellenknoten (Registerkarte "Typen") festgelegt wurden, Leerstellen sind.
@LAST_NON_BLANK(FELD)	beliebig	Ergibt den letzten nicht leeren Wert für <i>FELD</i> , wie in einem vorgeordneten Quellen- oder Typknoten definiert. Wenn in den bisher gelesenen Datensätzen keine nicht leeren Werte für <i>FELD</i> vorhanden sind, wird \$null\$ zurückgegeben. Leerwerte (auch als benutzerdefiniert fehlende Werte bezeichnet) können für jedes Feld einzeln definiert werden.
@NULL(FELD)	boolesch	Ergibt den Wert "Wahr", wenn der Wert von <i>FELD</i> der systemdefiniert fehlende Wert \$null\$ ist. Ergibt den Falschwert für alle anderen Werte, einschließlich benutzerdefinierter Leerstellen. Wenn Sie die Daten auf beides prüfen möchten, verwenden Sie @BLANK(FELD) und @NULL(FELD).

Table 36. CLEM-Funktionen zum Umgang mit Leerstellen und Nullwerten (Forts.)

Funktion	Ergebnis	Beschreibung
undef	<i>beliebig</i>	Wird allgemein in CLEM zur Eingabe eines \$null\$-Wertes verwendet; beispielsweise, um Leerwerte im Füllerknoten mit Nullen aufzufüllen.

Mithilfe des Füllerknotens können leere Felder "aufgefüllt" werden. Bei Füller- und Ableitungsknoten (nur beim Mehrfachmodus) bezieht sich die CLEM-Sonderfunktion @FIELD auf die aktuell untersuchten Felder.

Sonderfelder

Sonderfunktionen dienen zur Kennzeichnung der speziellen untersuchten Felder oder zum Generieren einer Felderliste als Eingabe. Wenn Sie beispielsweise mehrere Felder gleichzeitig ableiten, sollten Sie mit @FIELD angeben, dass diese Ableitungsaktion für die ausgewählten Felder durchgeführt werden soll. Mit dem Ausdruck $\log(@FIELD)$ wird ein neues Log-Feld für jedes ausgewählte Feld abgeleitet.

Table 37. CLEM-Sonderfelder

Funktion	Ergebnis	Beschreibung
@FIELD	<i>beliebig</i>	Führt eine Aktion für alle im Ausdruckskontext angegebenen Felder durch.
@TARGET	<i>beliebig</i>	Wenn ein CLEM-Ausdruck in einer benutzerdefinierten Analysefunktion verwendet wird, steht @TARGET für das Zielfeld bzw. den "korrekten Wert" für das analysierte Paar aus Zielwert und vorhergesagtem Wert. Diese Funktion wird häufig in Analyseknöten verwendet.
@PREDICTED	<i>beliebig</i>	Wenn ein CLEM-Ausdruck in einer benutzerdefinierten Analysefunktion verwendet wird, steht @PREDICTED für den vorhergesagten Wert des analysierten Paares aus Zielwert und vorhergesagtem Wert. Diese Funktion wird häufig in Analyseknöten verwendet.
@PARTITION_FIELD	<i>beliebig</i>	Ersetzt den Namen des aktuellen Partitionsfelds.
@TRAINING_PARTITION	<i>beliebig</i>	Ergibt den Wert der aktuellen Trainingspartition. Beispiel: Wenn Sie einen Auswahlknoten verwenden wollen, um Trainingsdatensätze auszuwählen, verwenden Sie den CLEM-Ausdruck @PARTITION_FIELD = @TRAINING_PARTITION. Dadurch wird gewährleistet, dass der Auswahlknoten immer funktioniert, unabhängig davon, welche Werte zur Darstellung der einzelnen Partitionen in den Daten verwendet werden.
@TESTING_PARTITION	<i>beliebig</i>	Ergibt den Wert der aktuellen Testpartition.
@VALIDATION_PARTITION	<i>beliebig</i>	Ergibt den Wert der aktuellen Validierungspartition.

Tabelle 37. CLEM-Sonderfelder (Forts.)

Funktion	Ergebnis	Beschreibung
@FIELDS_BETWEEN(Anfang, Ende)	<i>beliebig</i>	Ergibt die Liste der Feldnamen, die zwischen dem angegebenen Anfangs- und Endfeld (einschließlich dieser Felder) liegen, und zwar auf der Grundlage der natürlichen Reihenfolge der Felder in den Daten, also gemäß der Einfügereihenfolge.
@FIELDS_MATCHING(Muster)	<i>beliebig</i>	Ergibt eine Liste der Feldnamen, die mit dem angegebenen Muster übereinstimmen. Das Fragezeichen (?) als Platzhalterzeichen kann die Stelle von genau einem beliebigen Zeichen einnehmen, der Stern (*) entspricht keinem, einem oder mehreren Zeichen. Soll ein Fragezeichen oder ein Stern nicht als Platzhalter fungieren, sondern bei der Suche berücksichtigt werden, geben Sie den umgekehrten Schrägstrich (\) als Escapezeichen ein. Anmerkung: Hierfür ist ein Zeichenfolgeliteral als Argument erforderlich. Es kann kein verschachtelter Ausdruck zur Generierung des Arguments verwendet werden.
@MULTI_RESPONSE_SET	<i>beliebig</i>	Ergibt die Liste der Felder im benannten Mehrfachantwortset.

Kapitel 11. Verwenden von IBM SPSS Modeler mit einem Repository

Informationen zu IBM SPSS Collaboration and Deployment Services Repository

Sie können SPSS Modeler zusammen mit einem Repository von IBM SPSS Collaboration and Deployment Services verwenden und dadurch die Lebensdauer von Data-Mining-Modellen und verwandten Vorhersageobjekten verwalten. Dies ermöglicht, dass diese Objekte im Zusammenhang mit Unternehmensanwendungen, Tools und Lösungen verwendet werden. IBM SPSS Modeler-Objekte, die auf diese Weise gemeinsam nutzbar sind, umfassen Streams, Knoten, Streamausgaben, Projekte und Modelle. Die Objekte werden im zentralen Repository gespeichert, wo sie für andere Anwendungen freigegeben und mithilfe von umfassenden Versionserstellungs-, Metadaten- und Suchfunktionen aufgezeichnet werden können.

Bevor Sie SPSS Modeler mit dem Repository verwenden können, müssen Sie einen Adapter auf dem Repository-Host installieren. Ohne diesen Adapter sehen Sie eventuell folgende Nachricht beim Versuch, über bestimmte SPSS Modeler-Knoten oder -Modelle auf Repository-Objekte zuzugreifen:

Das Repository benötigt evtl. Updates zur Unterstützung neuer Knoten-, Modell- und Ausgabetypen.

Anweisungen zum Installieren des Adapters enthält das Installationshandbuch zu SPSS Modeler Deployment, das als PDF-Datei im Rahmen Ihres Produktdownloads verfügbar ist. Genauere Angaben dazu, wie Sie über IBM SPSS Deployment Manager auf IBM SPSS Modeler-Repository-Objekte zugreifen, finden Sie im Bereitstellungshandbuch zu SPSS Modeler.

Die folgenden Abschnitte enthalten Informationen zum Zugriff auf das Repository über SPSS Modeler.

Umfassende Versionserstellung und Unterstützung der Suchfunktion

Das Repository bietet umfassende Funktionen für die Objektversionserstellung und Suchmöglichkeiten. Angenommen, Sie erstellen z. B. einen Stream und speichern ihn im Repository, wo Sie ihn mit Mitarbeitern aus anderen Abteilungen gemeinsam nutzen können. Wenn Sie den Stream später in SPSS Modeler aktualisieren, können Sie die aktualisierte Version zum Repository hinzufügen, ohne die frühere Version zu überschreiben. Es können weiterhin alle Versionen aufgerufen und nach Name, Beschriftung, verwendeten Feldern und anderen Attributen durchsucht werden. So können Sie z. B. alle Modellversionen suchen, die die Nettoeinkünfte als Eingabe verwenden, oder alle von einem bestimmten Autor erstellten Modelle. (Mit einem herkömmlichen Dateisystem müssten Sie jede Version unter einem anderen Dateinamen speichern, und die Beziehungen zwischen den Versionen wären der Software nicht bekannt.)

Single Sign-on

Mit der Single-Sign-on-Funktion können sich Benutzer mit dem Repository verbinden, ohne jedes Mal Benutzernamen und Kennwörter eingeben zu müssen. Die bestehenden Anmeldedaten für das lokale Netzwerk sind als Authentifizierung für IBM SPSS Collaboration and Deployment Services ausreichend. Diese Funktion hängt von folgenden Bedingungen ab:

- IBM SPSS Collaboration and Deployment Services muss für die Verwendung eines Single-Sign-on-Anbieters konfiguriert sein.
- Der Benutzer muss in einem Host angemeldet sein, der mit dem Anbieter kompatibel ist.

Weitere Informationen finden Sie in „Verbinden mit dem Repository“ auf Seite 192.

Speichern und Bereitstellen von Repository-Objekten

Streams, die in IBM SPSS Modeler erstellt werden, lassen sich so, wie sie sind, als Dateien mit der Erweiterung `.str` im Repository **speichern**. Auf diese Weise können mehrere Benutzer im gesamten Unternehmen auf einen einzelnen Stream zugreifen. Weitere Informationen finden Sie in „Speichern von Objekten im Repository“ auf Seite 193.

Es ist auch möglich, einen Stream im Repository bereitzustellen. Bereitgestellte Streams werden als Datei mit zusätzlichen Metadaten gespeichert. Bereitgestellte Streams können die Unternehmensfunktionen von IBM SPSS Collaboration and Deployment Services, wie automatisiertes Scoring und Modellaktualisierung, vollständig nutzen. So kann beispielsweise ein Modell automatisch in regelmäßigen Abständen aktualisiert werden, wenn neue Daten verfügbar werden. Alternativ kann ein Set von Streams zum Zweck der Champion-Challenger-Analyse verwendet werden, in der Streams verglichen werden, um zu ermitteln, welcher das wirksamste Vorhersagemodell enthält.

Anmerkung: Der Assoziationsregelknoten, der STP-Knoten und der TCM-Modellierungsknoten unterstützen die Schritte für die Modellevaluierung oder für Champion-Challenger in IBM SPSS Collaboration and Deployment Services nicht.

Sie können einen Stream (mit der Erweiterung `.str`) bereitstellen. Bei der Bereitstellung als Stream kann der Stream durch die Thin-Client-Anwendung IBM SPSS Modeler Advantage aktiviert werden. Weitere Informationen finden Sie in „Öffnen eines Streams in IBM SPSS Modeler Advantage“ auf Seite 211.

Weitere Informationen finden Sie in „Bereitstellungsoptionen für den Stream“ auf Seite 205.

Andere Bereitstellungsoptionen

IBM SPSS Collaboration and Deployment Services bietet zwar bereits extrem umfassende Funktionen zur Verwaltung von Unternehmensinhalten, es steht aber dennoch eine Reihe weiterer Mechanismen zur Bereitstellung und zum Export von Streams zur Verfügung, darunter:

- Exportieren Sie den Stream zur späteren Verwendung in IBM SPSS Modeler Solution Publisher Runtime.
- Exportieren Sie ein oder mehrere Modelle in PMML, einem XML-basierten Format zur Codierung von Modellinformationen. Weitere Informationen finden Sie in „Importieren und Exportieren von Modellen als PMML“ auf Seite 212.

Verbinden mit dem Repository

1. Um eine Verbindung mit dem Repository herzustellen, klicken Sie im IBM SPSS Modeler-Hauptmenü auf:

Tools > Repository > Optionen...

2. Geben Sie in das Feld **Repository-URL** den Verzeichnispfad zur oder die URL der Repository-Installation ein, auf die Sie zugreifen möchten, oder wählen Sie diesen Verzeichnispfad bzw. diese URL aus. Sie können immer nur die Verbindung zu einem einzigen Repository herstellen.

Die Einstellungen sind für jeden Standort oder jede Installation spezifisch. Für spezielle Anmeldedaten wenden Sie sich an den lokalen Systemadministrator.

Berechtigungs nachweise festlegen. Belassen Sie dieses Kontrollkästchen nicht aktiviert, um die Funktion für *Single Sign-on* zu aktivieren. Diese versucht, Sie mithilfe Ihres lokalen Benutzernamens und Kennworts anzumelden. Falls ein Single Sign-on nicht möglich ist oder falls Sie diese Option zur Inaktivierung des Single Sign-on auswählen (z. B. zur Anmeldung in ein Administratorkonto), wird ein Fenster angezeigt, in dem Sie aufgefordert werden, Ihre Berechtigungs nachweise einzugeben.

Eingabe der Berechtigungsnachweise für das Repository

In Abhängigkeit von Ihren Einstellungen sind möglicherweise folgende Felder im Repository erforderlich: Dialogfeld "Berechtigungsnachweise":

Benutzername und Kennwort. Geben Sie einen gültigen Benutzernamen und ein Kennwort für die Anmeldung an. Bitten Sie, falls erforderlich, Ihren lokalen Administrator um weitere Informationen.

Provider. Wählen Sie zur Authentifizierung einen Sicherheitsanbieter aus. Das Repository kann für die Verwendung verschiedener Sicherheitsanbieter konfiguriert werden. Wenn Sie mehr Hilfe benötigen, wenden Sie sich an Ihren lokalen Administrator.

Repository und Benutzername merken. Speichert die aktuellen Einstellungen als Standard, sodass Sie diese nicht bei jeder neuen Verbindung erneut eingeben müssen.

Suchen von Repository-Berechtigungsnachweisen

Wenn Sie über einen Analytic Server-, Cognos-, ODBC- oder TM1-Quellenknoten eine Verbindung zu einem Repository herstellen, können Sie zuvor aufgezeichnete Berechtigungsnachweise für die Verbindung zu einem Repository auswählen. Diese Berechtigungsnachweise werden im Dialogfeld **Repository-Berechtigungsnachweis auswählen** aufgelistet. Klicken Sie auf **Durchsuchen** neben dem Feld **Berechtigungsnachweis**, um dieses Dialogfeld auszuwählen.

Heben Sie im Dialogfeld **Repository-Berechtigungsnachweis auswählen** die Berechtigungsnachweise in der bereitgestellten Liste hervor und klicken Sie auf **OK**. Wenn die Liste zu umfangreich ist, geben Sie in das Feld **Filter** den Namen oder einen Teil des Namens ein, um die gewünschten Berechtigungsnachweise zu suchen.

Durchsuchen des Repository-Inhalts

Mit dem Repository können Sie die gespeicherten Inhalte ähnlich wie mit Windows Explorer durchsuchen. Außerdem können Sie verschiedene *Versionen* von jedem gespeicherten Objekt durchsuchen.

1. Um das Fenster von IBM SPSS Collaboration and Deployment Services Repository zu öffnen, klicken Sie in den SPSS Modeler-Menüs auf:

Tools > Repository > Explorative Datenanalyse...

1. Geben Sie gegebenenfalls Einstellungen für die Verbindung mit dem Repository ein. Weitere Informationen finden Sie in „Verbinden mit dem Repository“ auf Seite 192. Wenden Sie sich für spezielle Port-, Kennwort- und andere Verbindungsdaten an den lokalen Systemadministrator.

Das Explorer-Fenster zeigt anfangs eine Baumansicht der Ordnerhierarchie. Klicken Sie auf einen Ordnernamen, um den Inhalt des Ordners anzuzeigen.

Objekte, die der aktuellen Auswahl oder Suchkriterien entsprechen, werden zusammen mit detaillierten Informationen zu der ausgewählten Version, die im unteren rechten Fenster angezeigt wird, im rechten Fenster aufgelistet. Die angezeigten Attribute gelten für die aktuellste Version.

Speichern von Objekten im Repository

Sie können Streams, Knoten, Modelle, Modellpaletten, Projekte und Ausgabeobjekte im Repository speichern, von wo andere Benutzer und Anwendungen darauf zugreifen können.

Sie können auch Streamausgaben im Repository in einem Format veröffentlichen, in dem andere Benutzer diese mithilfe von IBM SPSS Collaboration and Deployment Services Deployment Portal über das Internet betrachten können.

Festlegen der Objekteigenschaften

Beim Speichern eines Objekts wird das Dialogfeld "Repository: Speichern" angezeigt, in dem Sie die Werte von mehreren Eigenschaften für das Objekt einstellen können. Sie verfügen über folgende Möglichkeiten:

- Auswählen des Namens und des Repository-Ordners, unter dem das Objekt gespeichert werden soll
- Hinzufügen von Informationen zum Objekt, z. B. Versionsbeschriftungen und andere Eigenschaften, nach denen gesucht werden kann
- Hinzufügen eines oder mehrerer Klassifizierungsthemen zum Objekt
- Festlegen von Sicherheitsoptionen für das Objekt

In den folgenden Abschnitten werden die einstellbaren Eigenschaften beschrieben.

Auswählen des Speicherorts für Objekte

Geben Sie in das Dialogfeld "Repository: Speichern" Folgendes ein.

Speichern in. Zeigt den aktuellen Ordner an, in dem das Objekt gespeichert wird. Doppelklicken Sie auf einen Ordnernamen in der Liste, um diesen Ordner als aktuellen Ordner festzulegen. Mit der Schaltfläche "Ordner nach oben" navigieren Sie zum übergeordneten Ordner. Mit der Schaltfläche "Neuer Ordner" erstellen Sie einen Ordner auf der aktuellen Ebene.

Dateiname. Der Name, unter dem das Objekt gespeichert wird.

Speichern. Speichert das Objekt am aktuellen Speicherort.

Hinzufügen von Informationen zu gespeicherten Objekten

Alle Felder auf der Registerkarte "Informationen" des Dialogfelds "Repository: Speichern" sind optional.

Autor. Der Benutzername des Benutzers, der das Objekt im Repository erstellt. Standardmäßig wird der Benutzername angezeigt, der für die Repository-Verbindung verwendet wurde, Sie können diesen Namen aber hier ändern.

Versionsbeschriftung. Wählen Sie eine Beschriftung aus der Liste aus, um die Objektversion anzugeben, oder klicken Sie auf **Hinzufügen**, um eine neue Beschriftung zu erstellen. Verwenden Sie in der Beschriftung nicht das Zeichen "[". Stellen Sie sicher, dass keine Kontrollkästchen aktiviert sind, wenn Sie dieser Objektversion keine Beschriftung zuweisen möchten. Weitere Informationen finden Sie in „Anzeigen und Bearbeiten von Objekteigenschaften“ auf Seite 203.

Beschreibung. Eine Beschreibung des Objekts. Benutzer können Objekte nach Beschreibung suchen (siehe Hinweis).

Schlüsselwörter. Ein oder mehrere Stichwörter, die sich auf das Objekt beziehen und die zum Suchen verwendet werden können (siehe Hinweis).

Ablaufdatum. Ein Datum, nach dem das Objekt nicht mehr für allgemeine Benutzer sichtbar ist, obwohl es sein Eigentümer und der Repository-Administrator noch sehen können. Zum Festlegen des Ablaufdatums wählen Sie die Option **Datum** und geben Sie das Datum ein oder wählen Sie ein Datum mithilfe der Schaltfläche "Kalender".

Speichern. Speichert das Objekt am aktuellen Speicherort.

Hinweis: Die Informationen in den Feldern **Beschreibung** und **Schlüsselwörter** werden getrennt von den Angaben in SPSS Modeler auf der Registerkarte "Anmerkungen" des Objekts behandelt. Eine Repository-Suche nach Beschreibung oder Schlüsselwort liefert keine Informationen von der Registerkarte "Anmerkungen". Weitere Informationen finden Sie in „Suchen nach Objekten im Repository“ auf Seite 199.

Zuweisen von Themen zu einem gespeicherten Objekt

Themen sind ein hierarchisches Klassifizierungssystem für den Inhalt, der im Repository gespeichert ist. Sie können beim Speichern von Objekten aus den verfügbaren Themen auswählen und Benutzer können auch Objekte nach Thema suchen. Die Liste verfügbarer Themen wird von Repository-Benutzern mit den geeigneten Berechtigungen festgelegt. (Weitere Informationen finden Sie im Benutzerhandbuch zu Deployment Manager).

Über die Registerkarte "Themen" im Dialogfeld "Repository: Speichern" können Sie dem Objekt ein Thema zuweisen:

1. Klicken Sie auf die Schaltfläche **Hinzufügen**.
2. Klicken Sie in der Liste verfügbarer Themen auf einen Themennamen.
3. Klicken Sie auf **OK**.

So entfernen Sie eine Themenzuweisung:

4. Wählen Sie das Thema in der Liste der zugewiesenen Themen aus.
5. Klicken Sie auf **Löschen**.

Festlegen der Sicherheitsoptionen für gespeicherte Objekte

Auf der Registerkarte "Sicherheit" des Dialogfelds "Repository: Speichern" können Sie eine Reihe von Sicherheitsoptionen für ein gespeichertes Objekt festlegen. Für ein oder mehrere **Prinzipale** (d. h. Benutzer oder Benutzergruppen) können Sie:

- dem Objekt Zugriffsrechte zuweisen
- Zugriffsrechte auf das Objekt ändern
- Zugriffsrechte auf das Objekt entfernen

Prinzipal. Der Repository-Benutzername des Benutzers oder der Gruppe mit Zugriffsrechten auf dieses Objekt.

Berechtigungen. Die Zugriffsrechte dieses Benutzers bzw. dieser Gruppe auf das Objekt.

Hinzufügen. Ermöglicht es Ihnen, einen oder mehrere Benutzer zu der Liste der Benutzer mit Zugriffsrechten auf dieses Objekt hinzuzufügen. Weitere Informationen finden Sie in „Hinzufügen eines Benutzers zur Berechtigungsliste“.

Ändern. Ermöglicht es Ihnen, die Zugriffsrechte des ausgewählten Benutzers bzw. der ausgewählten Gruppe auf dieses Objekt zu ändern. Lesezugriff wird standardmäßig gewährt. Mithilfe dieser Option können Sie zusätzliche Zugriffsrechte gewähren, nämlich "Eigentümer", "Schreiben", "Löschen" und "Ändern".

Löschen. Löscht den ausgewählten Benutzer bzw. die ausgewählte Gruppe aus der Berechtigungsliste für dieses Objekt.

Hinzufügen eines Benutzers zur Berechtigungsliste: Folgende Felder sind verfügbar, wenn Sie auf der Registerkarte "Sicherheit" des Dialogfelds "Repository: Speichern" die Option **Hinzufügen** auswählen:

Anbieter auswählen. Wählen Sie zur Authentifizierung einen Sicherheitsanbieter aus. Das Repository kann für die Verwendung verschiedener Sicherheitsanbieter konfiguriert werden. Wenn Sie mehr Hilfe benötigen, wenden Sie sich an Ihren lokalen Administrator.

Suchen. Geben Sie den Repository-Benutzernamen des Benutzers bzw. der Gruppe ein, die Sie hinzufügen möchten, und klicken Sie auf **Suchen**, um diesen Namen in der Benutzerliste anzuzeigen. Um mehrere Benutzernamen gleichzeitig hinzuzufügen, lassen Sie dieses Feld leer und klicken Sie einfach auf **Suchen**, um eine Liste mit allen Repository-Benutzernamen anzuzeigen.

Benutzerliste. Wählen Sie einen oder mehrere Benutzernamen aus der Liste und klicken Sie auf **OK**, um sie der Berechtigungsliste hinzuzufügen.

Ändern der Zugriffsrechte für ein Objekt: Folgende Felder sind verfügbar, wenn Sie auf der Registerkarte "Sicherheit" des Dialogfelds "Repository: Speichern" die Option **Ändern** auswählen.

Eigentümer. Wählen Sie diese Option aus, um diesem Benutzer bzw. dieser Gruppe Zugriffsrechte auf das Objekt zu geben. Der Eigentümer verfügt über die kompletten Zugriffsrechte auf das Objekt einschließlich Lösch- und Änderungsrechten.

Lesen. Standardmäßig hat ein Benutzer bzw. eine Gruppe, die nicht Objekteigentümer ist, nur Lesezugriff für das Objekt. Markieren Sie die geeigneten Kontrollkästchen für Schreiben, Löschen und Ändern, um die entsprechenden Zugriffsrechte für diesen Benutzer bzw. diese Gruppe hinzuzufügen.

Speichern von Streams

Sie können einen Stream als *.str*-Datei im Repository speichern, von wo andere Benutzer darauf zugreifen können.

Hinweis: Weitere Informationen zum Bereitstellen eines Streams, um zusätzliche Repository-Funktionen zu nutzen, finden Sie in „Bereitstellen von Streams“ auf Seite 204.

So speichern Sie den aktuellen Stream:

1. Klicken Sie im Hauptmenü auf Folgendes:
Datei > Speichern > Als Stream speichern...
2. Geben Sie gegebenenfalls Einstellungen für die Verbindung mit dem Repository ein. Weitere Informationen finden Sie in „Verbinden mit dem Repository“ auf Seite 192. Wenden Sie sich für spezielle Port-, Kennwort- und andere Verbindungsdaten an den lokalen Systemadministrator.
3. Wählen Sie im Dialogfeld "Repository: Speichern" den Ordner aus, in dem Sie das Objekt speichern möchten, geben Sie alle weiteren Informationen an, die Sie aufzeichnen möchten, und klicken Sie auf die Schaltfläche **Speichern**. Weitere Informationen finden Sie in „Festlegen der Objekteigenschaften“ auf Seite 194.

Temporäres Speichern von Projekten

Sie können ein vollständiges IBM SPSS Modeler-Projekt als *.cpj*-Datei im Repository speichern, damit andere Benutzer darauf zugreifen können.

Da eine Projektdatei ein Container für andere IBM SPSS Modeler-Objekte ist, müssen Sie angeben, wo IBM SPSS Modeler die Objekte des Projekts im Repository speichern soll. Dies geschieht über eine Einstellung im Dialogfeld "Projekteigenschaften". Weitere Informationen finden Sie in „Festlegen der Projekteigenschaften“ auf Seite 218.

Sobald Sie ein Projekt zum Speichern von Objekten im Repository konfiguriert haben, werden Sie immer, wenn Sie ein neues Objekt zum Projekt hinzufügen, automatisch von IBM SPSS Modeler aufgefordert, das Objekt zu speichern.

Nach Abschluss der IBM SPSS Modeler-Sitzung müssen Sie eine neue Version der Projektdatei speichern, damit Ihre Änderungen erhalten bleiben. Die Projektdatei enthält automatisch die aktuellsten Versionen ihrer Objekte (und ruft sie ab). Wenn Sie einem Projekt während einer IBM SPSS Modeler-Sitzung keine Objekte hinzugefügt haben, brauchen Sie die Projektdatei nicht wiederherzustellen. Sie müssen jedoch neuere Versionen der Projektobjekte (Streams, Ausgaben usw.), die Sie geändert haben, temporär speichern.

So speichern Sie ein Projekt:

1. Wählen Sie das Projekt auf der Registerkarte "CRISP-DM" oder "Klassen" im Managerbereich in IBM SPSS Modeler aus. Klicken Sie dann im Hauptmenü auf:
Datei > Projekt > Projekt temporär speichern...
2. Geben Sie gegebenenfalls Einstellungen für die Verbindung mit dem Repository ein. Weitere Informationen finden Sie in „Verbinden mit dem Repository“ auf Seite 192. Wenden Sie sich für spezielle Port-, Kennwort- und andere Verbindungsdaten an den lokalen Systemadministrator.
3. Wählen Sie im Dialogfeld "Repository: Speichern" den Ordner aus, in dem Sie das Objekt speichern möchten, geben Sie alle weiteren Informationen an, die Sie aufzeichnen möchten, und klicken Sie auf die Schaltfläche **Speichern**. Weitere Informationen finden Sie in „Festlegen der Objekteigenschaften“ auf Seite 194.

Speichern von Knoten

Sie können eine individuelle Knotendefinition aus dem aktuellen Stream als *.nod*-Datei im Repository speichern, von wo andere Benutzer darauf zugreifen können.

So speichern Sie einen Knoten:

1. Klicken Sie im Streamerstellungsbereich mit der rechten Maustaste auf den Knoten und klicken Sie auf **Knoten temporär speichern**.
2. Geben Sie gegebenenfalls Einstellungen für die Verbindung mit dem Repository ein. Weitere Informationen finden Sie in „Verbinden mit dem Repository“ auf Seite 192. Wenden Sie sich für spezielle Port-, Kennwort- und andere Verbindungsdaten an den lokalen Systemadministrator.
3. Wählen Sie im Dialogfeld "Repository: Speichern" den Ordner aus, in dem Sie das Objekt speichern möchten, geben Sie alle weiteren Informationen an, die Sie aufzeichnen möchten, und klicken Sie auf die Schaltfläche **Speichern**. Weitere Informationen finden Sie in „Festlegen der Objekteigenschaften“ auf Seite 194.

Speichern von Ausgabeobjekten

Sie können ein Ausgabeobjekt aus dem aktuellen Stream als *.cou*-Datei im Repository speichern, von wo andere Benutzer darauf zugreifen können.

So speichern Sie ein Ausgabeobjekt:

1. Klicken Sie auf der Registerkarte "Ausgaben" im Managerbereich in SPSS Modeler auf das Objekt und klicken Sie dann im Hauptmenü auf:
Datei > Ausgaben > Ausgabe speichern...
2. Klicken Sie als Alternative mit der rechten Maustaste auf ein Objekt auf der Registerkarte "Ausgaben" und klicken Sie auf **Speichern**.
3. Geben Sie gegebenenfalls Einstellungen für die Verbindung mit dem Repository ein. Weitere Informationen finden Sie in „Verbinden mit dem Repository“ auf Seite 192. Wenden Sie sich für spezielle Port-, Kennwort- und andere Verbindungsdaten an den lokalen Systemadministrator.
4. Wählen Sie im Dialogfeld "Repository: Speichern" den Ordner aus, in dem Sie das Objekt speichern möchten, geben Sie alle weiteren Informationen an, die Sie aufzeichnen möchten, und klicken Sie auf die Schaltfläche **Speichern**. Weitere Informationen finden Sie in „Festlegen der Objekteigenschaften“ auf Seite 194.

Speichern von Modellen und Modellpaletten

Sie können ein einzelnes Modell als *.gm*-Datei im Repository speichern, von wo andere Benutzer darauf zugreifen können. Sie können auch den vollständigen Inhalt der Modellpalette als *.gen*-Datei im Repository speichern.

Speichern eines Modells:

1. Klicken Sie in der Modellpalette in SPSS Modeler auf das Objekt und klicken Sie dann im Hauptmenü auf:
Datei > Modelle > Modell speichern...
2. Klicken Sie als Alternative mit der rechten Maustaste auf ein Objekt in der Modellpalette und klicken Sie dann auf **Modell speichern**.
3. Fahren Sie mit "Abschließen des Speichervorgangs" fort.

Speichern einer Modellpalette:

1. Klicken Sie mit der rechten Maustaste in den Hintergrund der Modellpalette.
2. Klicken Sie im Popup-Menü auf **Palette speichern**.
3. Fahren Sie mit "Abschließen des Speichervorgangs" fort.

Abschließen des Speichervorgangs:

1. Geben Sie gegebenenfalls Einstellungen für die Verbindung mit dem Repository ein. Weitere Informationen finden Sie in „Verbinden mit dem Repository“ auf Seite 192. Wenden Sie sich für spezielle Port-, Kennwort- und andere Verbindungsdaten an den lokalen Systemadministrator.
2. Wählen Sie im Dialogfeld "Repository: Speichern" den Ordner aus, in dem Sie das Objekt speichern möchten, geben Sie alle weiteren Informationen an, die Sie aufzeichnen möchten, und klicken Sie auf die Schaltfläche **Speichern**. Weitere Informationen finden Sie in „Festlegen der Objekteigenschaften“ auf Seite 194.

Abrufen von Objekten aus dem Repository

Sie können Streams, Modelle, Modellpaletten, Knoten, Projekte und Ausgabeobjekte abrufen, die temporär im Repository gespeichert wurden.

Hinweis: Sie können neben der hier beschriebenen Verwendung der Menüoptionen diese auch nutzen, um Streams, Ausgabeobjekte, Modelle und Modellpaletten abzurufen, indem Sie mit der rechten Maustaste in die jeweilige Registerkarte des Managerbereichs oben rechts im SPSS Modeler-Fenster klicken.

1. Zum Abrufen eines Streams klicken Sie im IBM SPSS Modeler-Hauptmenü auf:
Datei > Stream abrufen...
2. Um ein Modell, eine Modellpalette, ein Projekt oder ein Ausgabeobjekt abzurufen, klicken Sie im IBM SPSS Modeler-Hauptmenü auf:
Datei > Modelle > Modell abrufen...
oder
Datei > Modelle > Modellpalette abrufen...
oder
Datei > Projekte > Projekt abrufen...
oder
Datei > Ausgaben > Ausgabe abrufen...
3. Klicken Sie als Alternative mit der rechten Maustaste in den Manager- oder Projektbereich und klicken Sie im Popup-Menü auf **Abrufen**.
4. Zum Abrufen eines Knotens klicken Sie im IBM SPSS Modeler-Hauptmenü auf:
Einfügen > Knoten (oder Superknoten) aus Repository...
 - a. Geben Sie gegebenenfalls Einstellungen für die Verbindung mit dem Repository ein. Weitere Informationen finden Sie in „Verbinden mit dem Repository“ auf Seite 192. Wenden Sie sich für spezielle Port-, Kennwort- und andere Verbindungsdaten an den lokalen Systemadministrator.
5. Navigieren Sie im Dialogfeld "Repository: Abrufen" zum Objekt, wählen Sie es aus und klicken Sie auf die Schaltfläche **Abrufen**.

Auswählen eines abzurufenden Objekts

Folgende Felder sind im Dialogfeld "Repository: Abrufen/Suchen" verfügbar:

Suchen in: Zeigt die Ordnerhierarchie für den aktuellen Ordner. Um zu einem anderen Ordner zu navigieren, wählen Sie einen aus dieser Liste aus, um direkt zu dem Ordner zu wechseln, oder navigieren Sie mithilfe der Objektliste unter diesem Feld.

Schaltfläche "Ordner nach oben". Wechselt in der Hierarchie auf die Ebene über dem aktuellen Ordner.

Schaltfläche "Neuer Ordner". Erstellt auf der aktuellen Ebene in der Hierarchie einen neuen Ordner.

Dateiname. Der Repository-Dateiname des ausgewählten Objekts. Klicken Sie zum Abrufen dieses Objekts auf **Abrufen**.

Dateien des Typs. Der Typ des Objekts, das Sie abrufen möchten. Nur Objekte dieses Typs werden zusammen mit Ordnern in der Objektliste angezeigt. Um Objekte eines anderen Typs zum Abrufen anzuzeigen, wählen Sie den Objekttyp aus der Liste aus.

Gesperrt öffnen. Ein Objekt wird beim Abrufen standardmäßig im Repository gesperrt, damit andere es nicht aktualisieren können. Wenn das Objekt beim Abrufen nicht gesperrt werden soll, inaktivieren Sie dieses Kontrollkästchen.

Beschreibung, Schlüsselwörter. Wenn beim Speichern des Objekts zusätzliche Details über das Objekt definiert wurden, werden diese Details hier angezeigt. Weitere Informationen finden Sie in „Hinzufügen von Informationen zu gespeicherten Objekten“ auf Seite 194.

Version. Um eine andere Objektversion als die aktuellste abzurufen, klicken Sie auf diese Schaltfläche. Es werden detaillierte Informationen für alle Versionen angezeigt, damit Sie die gewünschte Version auswählen können.

Auswählen einer Objektversion

So wählen Sie im Dialogfeld "Repository: Version auswählen" eine bestimmte Version eines Repository-Objekts aus:

1. (Optional) Sortieren Sie die Liste nach Version, Beschriftung, Größe, Erstellungsdatum oder Ersteller, indem Sie auf den Titel der entsprechenden Spalte doppelklicken.
2. Wählen Sie die Objektversion aus, mit der Sie arbeiten möchten.
3. Klicken Sie auf "Weiter".

Suchen nach Objekten im Repository

Sie können Objekten anhand von Namen, Ordner, Typ, Beschriftung, Datum oder anderen Kriterien suchen.

Suchen nach Objekten nach Name

1. Klicken Sie im IBM SPSS Modeler-Hauptmenü auf:
Tools > Repository > Explorative Datenanalyse...
 - a. Geben Sie gegebenenfalls Einstellungen für die Verbindung mit dem Repository ein. Weitere Informationen finden Sie in „Verbinden mit dem Repository“ auf Seite 192. Wenden Sie sich für spezielle Port-, Kennwort- und andere Verbindungsdaten an den lokalen Systemadministrator.
2. Klicken Sie auf die Registerkarte **Suchen**.
3. Geben Sie im Feld **Suchen nach Objekten mit Namen** den Namen des zu suchenden Objekts an.

Wenn Sie anhand des Namens nach Objekten suchen, kann ein Stern (*) als Platzhalterzeichen für eine beliebige Zeichenfolge oder ein Fragezeichen (?) als Platzhalterzeichen für ein einzelnes Zeichen verwendet werden. So werden beispielsweise mit *cluster* alle Objekte gefunden, die irgendwo im Namen die Zeichenfolge cluster enthalten. Mit der Suchzeichenfolge m0?_* werden beispielsweise die Namen M01_cluster.str und M02_cluster.str, nicht jedoch M01a_cluster.str gefunden. Bei der Suche wird nicht zwischen Groß- und Kleinschreibung unterschieden (cluster entspricht Cluster und CLUSTER).

Anmerkung: Bei einer großen Anzahl von Objekten kann der Suchvorgang einige Zeit in Anspruch nehmen.

Suchen nach anderen Kriterien

Sie können die Suche auf der Basis von Titel, Beschriftung, Datum, Autor, Schlüsselwort, Indexinhalt oder Beschreibung ausführen. Es werden lediglich Objekte gefunden, die *allen* angegebenen Suchkriterien entsprechen. Sie könnten beispielsweise alle Streams suchen, die mindestens ein Clustering-Modell enthalten und auf die außerdem eine bestimmte Beschriftung angewendet wurde und die nach einem bestimmten Datum bearbeitet wurden.

Objekttypen. Sie können die Suche auf Modelle, Streams, Ausgaben, Knoten, Superknoten, Projekte, Modellpaletten oder andere Objekttypen einschränken.

- **Modelle.** Sie können anhand der Kategorie (Klassifizierung, Näherung, Clusterbildung usw.) eines bestimmten Modellierungsalgorithmus, wie beispielsweise Kohonen, nach Modellen suchen.

Sie können die Suche auch nach verwendeten Feldern durchführen, z. B. nach allen Modellen, die ein Feld *Einkommen* als Eingabe- oder Ausgabefeld (Zielfeld) verwenden.

- **Streams.** Für Streams können Sie die Suche nach verwendeten Feldern oder Modelltyp (Kategorie oder Algorithmus) im Stream beschränken.

Themen. Sie können nach Modellen suchen, die bestimmten Themen aus einer Liste zugeordnet sind, die von Repository-Benutzern mit den entsprechenden Berechtigungen festgelegt wurde. (Weitere Informationen finden Sie im Benutzerhandbuch zu Deployment Manager). Um die Liste abzurufen, aktivieren Sie dieses Kontrollkästchen. Klicken Sie dann auf die Schaltfläche "Themen hinzufügen", die eingeblendet wird, wählen Sie ein oder mehrere Themen in der Liste aus und klicken Sie auf **OK**.

Beschriftung. Beschränkt die Suche auf bestimmte Objektversionsbeschriftungen.

Daten. Sie können ein Erstellungs- oder Änderungsdatum angeben und nach Objekten vor, nach oder innerhalb des angegebenen Datumsbereichs suchen.

Autor. Beschränkt die Suche auf Objekte, die von einem bestimmten Benutzer erstellt wurden.

Schlüsselwörter. Suche nach bestimmten Schlüsselwörtern. In IBM SPSS Modeler werden Schlüsselwörter auf der Registerkarte "Anmerkung" für einen Stream, ein Modell oder ein Ausgabeobjekt angegeben.

Beschreibung. Suche nach bestimmten Termen im Beschreibungsfeld. In IBM SPSS Modeler wird die Beschreibung auf der Registerkarte "Anmerkung" für einen Stream, ein Modell oder ein Ausgabeobjekt angegeben. Mehrere Suchausdrücke können durch Semikolons voneinander getrennt werden, beispielsweise Einkommen; Feldfruchttyp; Subventionshöhe. (Beachten Sie, dass innerhalb von Suchausdrücken, Leerzeichen von Bedeutung sind. Zum Beispiel sind Höhe der Subventionen mit einem Leerzeichen vor "der" und Höhe der Subventionen mit zwei Leerzeichen vor "der" nicht dasselbe.)

Ändern von Repository-Objekten

Sie können bestehende Objekte im Repository direkt von SPSS Modeler aus ändern. Sie verfügen über folgende Möglichkeiten:

- Erstellen, Umbenennen oder Löschen von Ordnern
- Sperren oder Entsperren von Objekten
- Löschen von Objekten

Erstellen, Umbenennen und Löschen von Ordnern

1. Um Operationen an Ordnern im Repository auszuführen, klicken Sie im SPSS Modeler-Hauptmenü auf:
Tools > Repository > Explorative Datenanalyse...
 - a. Geben Sie gegebenenfalls Einstellungen für die Verbindung mit dem Repository ein. Weitere Informationen finden Sie in „Verbinden mit dem Repository“ auf Seite 192. Wenden Sie sich für spezielle Port-, Kennwort- und andere Verbindungsdaten an den lokalen Systemadministrator.
2. Stellen Sie sicher, dass die Registerkarte **Ordner** aktiv ist.
3. Um einen neuen Ordner zu erstellen, klicken Sie mit der rechten Maustaste auf den übergeordneten Ordner und klicken Sie dann auf **Neuer Ordner**.
4. Um einen Ordner umzubenennen, klicken Sie mit der rechten Maustaste darauf und klicken Sie dann auf **Ordner umbenennen**.
5. Um einen Ordner zu löschen, klicken Sie mit der rechten Maustaste darauf und klicken Sie dann auf **Ordner löschen**.

Sperren und Entsperren von Repository-Objekten

Sie können ein Objekt sperren, um zu verhindern, dass andere Benutzer seine bestehenden Versionen aktualisieren oder neue Versionen erstellen. Ein gesperrtes Objekt wird durch ein Vorhängeschlosssymbol auf dem Objektsymbol gekennzeichnet.



Abbildung 17. Gesperrtes Objekt

So sperren Sie ein Objekt:

1. Klicken Sie im Repository-Explorer-Fenster mit der rechten Maustaste auf das gewünschte Objekt.
2. Klicken Sie auf **Sperren**.

So entsperren Sie ein Objekt:

1. Klicken Sie im Repository-Explorer-Fenster mit der rechten Maustaste auf das gewünschte Objekt.
2. Klicken Sie auf **Entsperren**.

Löschen von Repository-Objekten

Bevor Sie ein Objekt aus dem Repository löschen, müssen Sie entscheiden, ob Sie alle Versionen oder nur eine bestimmte Version des Objekts löschen möchten.

So löschen Sie alle Versionen eines Objekts:

1. Klicken Sie im Repository-Explorer-Fenster mit der rechten Maustaste auf das gewünschte Objekt.
2. Klicken Sie auf **Objekte löschen**.

So löschen Sie die aktuellste Version eines Objekts:

1. Klicken Sie im Repository-Explorer-Fenster mit der rechten Maustaste auf das gewünschte Objekt.

2. Klicken Sie auf **Löschen**.

So löschen Sie eine frühere Version eines Objekts:

1. Klicken Sie im Repository-Explorer-Fenster mit der rechten Maustaste auf das gewünschte Objekt.
2. Klicken Sie auf **Versionen löschen**.
3. Wählen Sie die zu löschenden Versionen aus und klicken Sie auf **OK**.

Verwalten der Eigenschaften von Repository-Objekten

Sie können zahlreiche Objekteigenschaften von SPSS Modeler aus steuern. Sie verfügen über folgende Möglichkeiten:

- Anzeigen der Ordneigenschaften
- Anzeigen und Bearbeiten von Objekteigenschaften
- Erstellen, Zuweisen und Löschen von Versionsbeschriftungen für ein Objekt

Anzeigen von Ordneigenschaften

Um Eigenschaften für einen beliebigen Ordner im Repository-Fenster anzuzeigen, klicken Sie mit der rechten Maustaste auf den gewünschten Ordner. Klicken Sie auf **Ordneigenschaften**.

Registerkarte "Allgemein"

Diese Registerkarte zeigt den Ordnernamen, das Erstellungsdatum und das Änderungsdatum an.

Registerkarte "Berechtigungen"

Auf dieser Registerkarte geben Sie Lese- und Schreibberechtigungen für den Ordner an. Es werden alle Benutzer und Gruppen mit Zugriff auf den übergeordneten Ordner aufgelistet. Es gibt eine Hierarchie der Berechtigungen. Wenn Sie beispielsweise nicht über eine Leseberechtigung verfügen, können Sie auch nicht über eine Schreibberechtigung verfügen. Wenn Sie nicht über eine Schreibberechtigung verfügen, können Sie auch nicht über eine Löschberechtigung verfügen.

Benutzer und Gruppen. Listet die Repository-Benutzer und -Gruppen auf, die mindestens über Lesezugriff für diesen Ordner verfügen. Aktivieren Sie die Kontrollkästchen für Schreiben und Löschen, um die entsprechenden Zugriffsrechte für einen bestimmten Benutzer bzw. eine bestimmte Gruppe hinzuzufügen. Klicken Sie auf das Symbol **Benutzer/Gruppen hinzufügen** rechts auf der Registerkarte "Berechtigungen", um weiteren Benutzern und Gruppen den Zugriff zuzuweisen. Die Liste mit den verfügbaren Benutzern und Gruppen wird vom Administrator gesteuert.

Berechtigungen weiterleiten. Wählen Sie eine Option aus, um zu steuern, wie Änderungen am aktuellen Ordner seinen untergeordneten Ordnern (falls vorhanden) zugewiesen werden.

- **Alle Berechtigungen weiterleiten.** Leitet die Berechtigungseinstellungen vom aktuellen Ordner zu allen untergeordneten Ordnern weiter. Dies stellt eine schnelle Möglichkeit dar, Berechtigungen für mehrere Ordner gleichzeitig festzulegen. Legen Sie die Berechtigungen für den übergeordneten Ordner wie gewünscht fest und leiten Sie diese dann nach Bedarf weiter.
- **Nur Änderungen weiterleiten.** Leitet nur Änderungen weiter, die seit den letzten Änderungen vorgenommen wurden. Wenn z. B. eine neue Gruppe hinzugefügt wurde und sie dieser Gruppe Zugriff auf alle Ordner unter "Umsatz" zuweisen möchten, können Sie der Gruppe Zugriff auf den Stammordner "Umsatz" geben und die Änderung an alle Unterordner weiterleiten. Alle anderen Berechtigungen für bereits vorhandene Unterordner bleiben wie zuvor bestehen.
- **Nicht weiterleiten.** Alle vorgenommenen Änderungen gelten lediglich für den aktuellen Ordner und werden nicht an die untergeordneten Ordner weitergeleitet.

Anzeigen und Bearbeiten von Objekteigenschaften

Im Dialogfeld "Objekteigenschaften" können Sie Eigenschaften anzeigen und bearbeiten. Zwar können einige Eigenschaften nicht geändert werden, Sie können jedoch immer ein Objekt aktualisieren, indem Sie eine neue Version hinzufügen.

1. Klicken Sie im Repository-Fenster mit der rechten Maustaste auf das gewünschte Objekt.
2. Klicken Sie auf **Objekteigenschaften**.

Registerkarte "Allgemein"

Name. Der Name des Objekts, wie im Repository angezeigt.

Erstellt am: Datum, an dem das Objekt (nicht die Version) erstellt wurde.

Zuletzt geändert am: Datum, an dem die aktuellste Version geändert wurde.

Autor. Der Anmeldenname des Benutzers.

Beschreibung. Standardmäßig enthält diese Option die Beschreibung, die in SPSS Modeler auf der Registerkarte "Anmerkung" angegeben ist.

Verknüpfte Themen. Mit dem Repository können Modelle und verwandte Objekte nach Themen geordnet werden. Die Liste verfügbarer Themen wird von Repository-Benutzern mit den geeigneten Berechtigungen festgelegt. (Weitere Informationen finden Sie im Benutzerhandbuch zu Deployment Manager).

Schlüsselwörter. Schlüsselwörter werden auf der Registerkarte "Anmerkung" für einen Stream, ein Modell oder ein Ausgabeobjekt angegeben. Mehrere Schlüsselwörter sollten durch Leerzeichen voneinander getrennt werden, insgesamt sind bis zu 255 Zeichen zulässig. (Wenn Schlüsselwörter Leerzeichen enthalten, müssen Sie zur Trennung Ausrufezeichen verwenden.)

Registerkarte "Versionen"

Die im Repository gespeicherten Objekte weisen möglicherweise mehrere Versionen auf. Auf der Registerkarte "Versionen" werden Informationen zu jeder Version angezeigt.

Die folgenden Eigenschaften können für bestimmte Versionen eines gespeicherten Objekts angegeben oder geändert werden:

Version. Eindeutige ID für die Version, generiert basierend auf der Uhrzeit, zu der die Version gespeichert wurde.

Beschriftung. Aktuelle Beschriftung für die Version, falls vorhanden. Im Gegensatz zur Versions-ID können die Beschriftungen von einer Version eines Objekts zu einer anderen verschoben werden.

Die Dateigröße, das Erstellungsdatum und der Autor werden ebenfalls für jede Version angezeigt.

Beschriftungen bearbeiten. Klicken Sie auf das Symbol **Beschriftungen bearbeiten** rechts oben auf der Registerkarte "Versionen", um Beschriftungen für gespeicherte Objekte zu definieren, anzuwenden oder zu entfernen. Weitere Informationen finden Sie in „Verwalten von Objektversionsbeschriftungen“ auf Seite 204.

Registerkarte "Berechtigungen"

Auf der Registerkarte **Berechtigungen** können Sie Lese- und Schreibberechtigungen für das Objekt festlegen. Es werden alle Benutzer und Gruppen mit Zugriff auf das aktuelle Objekt aufgelistet. Es gibt eine Hierarchie der Berechtigungen. Wenn Sie beispielsweise nicht über eine Leseberechtigung verfügen, kön-

nen Sie auch nicht über eine Schreibberechtigung verfügen. Wenn Sie nicht über eine Schreibberechtigung verfügen, können Sie auch nicht über eine Löschberechtigung verfügen.

Benutzer und Gruppen. Listet die Repository-Benutzer und -Gruppen auf, die mindestens über Lesezugriff für dieses Objekt verfügen. Aktivieren Sie die Kontrollkästchen für Schreiben und Löschen, um die entsprechenden Zugriffsrechte für einen bestimmten Benutzer bzw. eine bestimmte Gruppe hinzuzufügen. Klicken Sie auf das Symbol **Benutzer/Gruppen hinzufügen** rechts auf der Registerkarte "Berechtigungen", um weiteren Benutzern und Gruppen den Zugriff zuzuweisen. Die Liste mit den verfügbaren Benutzern und Gruppen wird vom Administrator gesteuert.

Verwalten von Objektversionsbeschriftungen

Im Dialogfeld "Versionsbeschriftungen bearbeiten" haben Sie folgende Möglichkeiten:

- Anwenden von Beschriftungen auf das ausgewählte Objekt
- Entfernen von Beschriftungen vom ausgewählten Objekt
- Definieren einer neuen Beschriftung und Anwenden auf das Objekt

So wenden Sie Beschriftungen auf ein Objekt an:

1. Wählen Sie in der Liste **Verfügbare Beschriftungen** eine oder mehrere Beschriftungen aus.
2. Klicken Sie auf die Schaltfläche mit dem Rechtspfeil, um die ausgewählten Beschriftungen in die Liste **Angewendete Beschriftungen** zu verschieben.
3. Klicken Sie auf **OK**.

So entfernen Sie Beschriftungen von einem Objekt:

1. Wählen Sie in der Liste **Angewendete Beschriftungen** eine oder mehrere Beschriftungen aus.
2. Klicken Sie auf die Schaltfläche mit dem Linkspfeil, um die ausgewählten Beschriftungen in die Liste **Verfügbare Beschriftungen** zu verschieben.
3. Klicken Sie auf **OK**.

So definieren Sie eine neue Beschriftung und wenden sie auf das Objekt an:

1. Geben Sie den Beschriftungsnamen in das Feld **Neue Beschriftung** ein.
2. Klicken Sie auf die Schaltfläche mit dem Rechtspfeil, um die neue Beschriftung in die Liste **Angewendete Beschriftungen** zu verschieben.
3. Klicken Sie auf **OK**.

Bereitstellen von Streams

Damit ein Stream mit der Thin-Client-Anwendung IBM SPSS Modeler Advantage verwendet werden kann, muss er als Stream (.str-Datei) im Repository bereitgestellt werden.

Anmerkung: Sie können keinen Stream bereitstellen, der mehrere Quellenknoten in der Scoring-Verzweigung hat.

Streams können die Unternehmensfunktionen von IBM SPSS Collaboration and Deployment Services vollständig nutzen. Weitere Informationen finden Sie in „Speichern und Bereitstellen von Repository-Objekten“ auf Seite 192.

So stellen Sie den aktuellen Stream bereit (Verfahren mit dem Menü "Datei"):

1. Klicken Sie im Hauptmenü auf Folgendes:
Datei > Bereitstellen
2. Wählen Sie den Bereitstellungstyp aus und füllen Sie das übrige Dialogfeld wie erforderlich aus.

3. Klicken Sie auf **Als Stream bereitstellen**, um den Stream für die Verwendung mit IBM SPSS Modeler Advantage oder IBM SPSS Collaboration and Deployment Services bereitzustellen.
4. Klicken Sie auf **Speichern**. Wenn Sie Informationen wünschen, klicken Sie auf **Hilfe**.
5. Fahren Sie mit "Abschließen des Bereitstellungsvorgangs" fort.

So stellen Sie den aktuellen Stream bereit (Verfahren mit dem Menü "Tools"):

1. Klicken Sie im Hauptmenü auf Folgendes:
Tools > Streameigenschaften > Bereitstellung
2. Wählen Sie den Bereitstellungstyp aus, füllen Sie den Rest der Registerkarte "Bereitstellung" wie erforderlich aus und klicken Sie auf **Speichern**. Weitere Informationen finden Sie in „Bereitstellungsoptionen für den Stream“.

Abschließen des Bereitstellungsvorgangs

1. Geben Sie gegebenenfalls Einstellungen für die Verbindung mit dem Repository ein. Weitere Informationen finden Sie in „Verbinden mit dem Repository“ auf Seite 192. Wenden Sie sich für spezielle Port-, Kennwort- und andere Verbindungsdaten an den lokalen Systemadministrator.
2. Wählen Sie im Dialogfeld **Repository: Speichern** den Ordner aus, in dem Sie das Objekt speichern möchten, geben Sie alle weiteren Informationen an, die Sie aufzeichnen möchten, und klicken Sie auf **Speichern**. Weitere Informationen finden Sie in „Festlegen der Objekteigenschaften“ auf Seite 194.

Bereitstellungsoptionen für den Stream

Auf der Registerkarte "Bereitstellung" im Dialogfeld "Streamoptionen" können Sie Optionen für die Bereitstellung des Streams festlegen.

Wenn Sie einen Stream bereitstellen, können Sie den Stream in der Thin-Client-Anwendung IBM SPSS Modeler Advantage öffnen und ändern. Der Stream wird im Repository als Datei mit der Erweiterung `.str` gespeichert.

Bei der Bereitstellung eines Streams können Sie die zusätzliche Funktionalität nutzen, die in IBM SPSS Collaboration and Deployment Services zur Verfügung steht, z. B. Mehrbenutzerzugriff, automatisches Scoring, Modellaktualisierung und Champion-Challenger-Analyse.

Anmerkung: Der Assoziationsregelknoten, der STP-Knoten und der TCM-Modellierungsknoten unterstützen die Schritte für die Modellevaluierung oder für Champion-Challenger in IBM SPSS Collaboration and Deployment Services nicht.

Über die Registerkarte "Bereitstellung" können Sie auch eine Vorschau der Streambeschreibung anzeigen, die IBM SPSS Modeler für den Stream erstellt. Weitere Informationen finden Sie in „Streambeschreibungen“ auf Seite 57.

Bereitstellungstyp. Wählen Sie aus, wie der Stream bereitgestellt werden soll. Alle Streams benötigen einen speziellen Scoring-Knoten, bevor sie bereitgestellt werden können; weitere Voraussetzungen und Optionen hängen vom Bereitstellungstyp ab.

Anmerkung: Der Assoziationsregelknoten, der STP-Knoten und der TCM-Modellierungsknoten unterstützen die Schritte für die Modellevaluierung oder für Champion-Challenger in IBM SPSS Collaboration and Deployment Services nicht.

- **<keine>**. Der Stream wird nicht für das Repository bereitgestellt. Alle Optionen mit Ausnahme der Vorschau der Streambeschreibung sind inaktiviert.
- **Nur Scoring.** Der Stream wird dem Repository bereitgestellt, wenn Sie auf die Schaltfläche **Speichern** klicken. Daten können mithilfe des Knotens gescort werden, den Sie im Feld **Scoring-Knoten** angeben.
- **Modellaktualisierung.** Wie "Nur Scoring", jedoch kann das Modell zusätzlich im Repository mit Objekten aktualisiert werden, die Sie in den Feldern **Modellierungsknoten** und **Modellnutzget** festlegen. Be-

achten Sie, dass die automatische Modellaktualisierung in IBM SPSS Collaboration and Deployment Services nicht standardmäßig unterstützt wird. Daher müssen Sie diesen Bereitstellungstyp auswählen, wenn Sie diese Funktion beim Ausführen eines Streams aus dem Repository verwenden wollen. Weitere Informationen finden Sie in „Modellaktualisierung“ auf Seite 208.

Scoring-Knoten. Wählen Sie ein Diagramm, einen Ausgabe- oder einen Exportknoten aus, um die Streamverzweigung zu identifizieren, die zum Scoring der Daten verwendet werden soll. Der Stream kann zwar eine beliebige Anzahl von gültigen Verzweigungen, Modellen und Endknoten enthalten, zu Bereitstellungszwecken muss jedoch genau eine Scoring-Verzweigung angegeben werden. Dies ist die grundlegendste Voraussetzung für die Bereitstellung eines Streams.

Scoring-Parameter. Dient zur Angabe von Parametern, die bei der Ausführung der Scoring-Verzweigung geändert werden können. Weitere Informationen finden Sie in „Scoring und Modellierung von Parametern“.

Modellierungsknoten. Gibt für die Modellaktualisierung den Modellierungsknoten an, der zum Generieren bzw. zur Aktualisierung des Modells im Repository verwendet wird. Dies muss ein Modellierungsknoten desselben Typs sein, der für **Modellnugget** angegeben wurde.

Modellerstellungsparameter. Dient zur Angabe von Parametern, die bei der Ausführung des Modellierungsknotens geändert werden können. Weitere Informationen finden Sie in „Scoring und Modellierung von Parametern“.

Modellnugget. Gibt zum Zwecke der Modellaktualisierung das Modellnugget an, das bei jeder Aktualisierung des Streams im Repository aktualisiert bzw. neu generiert werden soll (üblicherweise im Rahmen einer geplanten Aufgabe). Das Modell muss sich in der Scoring-Verzweigung befinden. In der Scoring-Verzweigung können mehrere Modelle vorhanden sein, es kann jedoch nur eines angegeben werden. Beachten Sie: Bei der erstmaligen Erstellung des Streams kann es sich hierbei im Grunde um ein Platzhaltermodell handeln, das immer dann aktualisiert bzw. neu generiert wird, wenn neue Daten verfügbar sind.

Als Stream bereitstellen. Klicken Sie auf diese Option, wenn Sie den Stream mit IBM SPSS Modeler Advantage oder IBM SPSS Collaboration and Deployment Services verwenden möchten.

Überprüfen. Klicken Sie auf diese Schaltfläche, um zu prüfen, ob es sich um einen gültigen Stream zur Bereitstellung handelt. Alle Streams müssen über einen speziellen Scoring-Knoten verfügen, bevor sie bereitgestellt werden können. Wenn diese Bedingungen nicht erfüllt sind, werden Fehlernachrichten angezeigt.

Speichern. Stellt den Stream bereit, sofern dieser gültig ist. Andernfalls wird eine Fehlernachricht angezeigt. Klicken Sie auf die Schaltfläche **Korrigieren**, um den Fehler zu korrigieren und es erneut zu versuchen.

Vorschau der Streambeschreibung anzeigen. Ermöglicht die Anzeige des Inhalts einer Streambeschreibung, die IBM SPSS Modeler für den Stream erstellt. Weitere Informationen finden Sie in „Streambeschreibungen“ auf Seite 57.

Anmerkung: Der Assoziationsregelknoten, der STP-Knoten und der TCM-Modellierungsknoten unterstützen die Schritte für die Modellevaluierung oder für Champion-Challenger in IBM SPSS Collaboration and Deployment Services nicht.

Scoring und Modellierung von Parametern

Bei der Bereitstellung eines Streams für IBM SPSS Collaboration and Deployment Services können Sie auswählen, welche Parameter angezeigt bzw. bearbeitet werden können, wenn das Modell aktualisiert oder gescort wird. Sie können beispielsweise Mindest- und Höchstwerte oder einen anderen Wert angeben, der sich bei jeder Ausführung eines Jobs ändern kann.

1. Um einen Parameter sichtbar zu machen, sodass er nach der Bereitstellung des Streams angezeigt bzw. bearbeitet werden kann, müssen Sie ihn aus der Liste im Dialogfeld "Scoring-Parameter" auswählen.

Die Liste der verfügbaren Parameter wird im Dialogfeld für die Streameigenschaften auf der Registerkarte "Parameter" definiert. Weitere Informationen finden Sie in „Festlegen der Stream- und Sitzungsparameter“ auf Seite 53.

Scoring-Verzweigung

Wenn Sie einen Stream bereitstellen, muss eine Verzweigung des Streams als **Scoring-Verzweigung** (d. h. die Verzweigung, die den Scoring-Knoten enthält) festgelegt werden. Wenn Sie eine Verzweigung als Scoring-Verzweigung festlegen, wird im Streamerstellungsbereich diese Verzweigung sowie die Modellverknüpfung zum Nugget auf der Scoring-Verzweigung hervorgehoben. Die visuelle Repräsentation ist besonders nützlich in komplexen Streams mit mehreren Verzweigungen, in denen die Scoring-Verzweigung nicht sofort erkennbar ist.

Hinweis: Nur eine Streamverzweigung kann als Scoring-Verzweigung festgelegt werden.

Wenn für den Stream bereits eine Scoring-Verzweigung definiert wurde, wird diese durch die neu festgelegte Verzweigung als Scoring-Verzweigung ersetzt. Sie können die Farbe der Scoring-Verzweigung mithilfe einer Option für benutzerdefinierte Farben festlegen. Weitere Informationen finden Sie in „Festlegen von Anzeigeeoptionen“ auf Seite 226.

Sie können die Anzeige der Scoring-Verzweigung mithilfe der Symbolleistschaltfläche zum Ein-/Ausblenden von Stream-Markups ein- und ausblenden.



Abbildung 18. Symbolleistschaltfläche zum Ein-/Ausblenden von Stream-Markups

Angeben der Scoring-Verzweigung für die Bereitstellung

Sie können die Scoring-Verzweigung über das Popup-Menü eines Endknotens oder über das Menü "Tools" festlegen. Bei Verwendung des Popup-Menüs wird der Scoring-Knoten automatisch auf der Registerkarte "Bereitstellung" der Streameigenschaften festgelegt.

So legen Sie eine Verzweigung als die Scoring-Verzweigung fest (Popup-Menü):

1. Verbinden Sie das Modellnugget mit einem Endknoten (einem dem Nugget nachgeordneten Verarbeitungs- oder Ausgabeknoten).
2. Klicken Sie mit der rechten Maustaste auf den Endknoten.
3. Klicken Sie im Menü auf die Option **Als Scoring-Verzweigung verwenden**.

So legen Sie eine Verzweigung als die Scoring-Verzweigung fest (Menü "Tools"):

1. Verbinden Sie das Modellnugget mit einem Endknoten (einem dem Nugget nachgeordneten Verarbeitungs- oder Ausgabeknoten).
2. Klicken Sie im Hauptmenü auf Folgendes:
Tools > Streameigenschaften > Bereitstellung
3. Klicken Sie in der Liste **Bereitstellungstyp** wie gewünscht auf die Option **Nur Scoring** oder **Modellaktualisierung**. Weitere Informationen finden Sie in „Bereitstellungsoptionen für den Stream“ auf Seite 205.
4. Klicken Sie auf das Feld **Scoring-Knoten** und wählen Sie einen Endknoten aus der Liste.
5. Klicken Sie auf **OK**.

Modellaktualisierung

Eine Modellaktualisierung ist der Vorgang, bei dem ein vorhandenes Modell in einem Stream anhand von neueren Daten neu aufgebaut wird. Der Stream selbst ändert sich im Repository nicht. Beispielsweise bleiben Algorithmustyp und streamspezifische Einstellungen gleich, aber das Modell wird mit neuen Daten neu trainiert und dann aktualisiert, wenn die neue Version des Modells besser als die alte Version funktioniert.

Nur ein Modellnugget in einem Stream kann auf Aktualisierung eingestellt werden. Dies ist das sogenannte **Aktualisierungsmodell**. Wenn Sie auf die Option **Modellaktualisierung** auf der Registerkarte "Bereitstellung" klicken (siehe „Bereitstellungsoptionen für den Stream“ auf Seite 205), wird das zu diesem Zeitpunkt festgelegte Modellnugget zu Ihrem Aktualisierungsmodell. Sie können auch aus dem Pop-up-Menü eines Modellnuggets ein Modell als Aktualisierungsmodell festlegen. Das Nugget muss sich dafür bereits auf der Scoring-Verzweigung befinden.

Wenn Sie den Modellaktualisierungsstatus eines Nuggets inaktivieren, hat dies dieselbe Wirkung wie das Einstellen des Bereitstellungstyps des Streams auf "Nur Scoring" und die Registerkarte "Bereitstellung" des Dialogfelds für die Streameigenschaften wird entsprechend aktualisiert. Sie können diesen Status mit der Option **Als Aktualisierungsmodell verwenden** im Pop-up-Menü des Nuggets auf der aktuellen Scoring-Verzweigung ein- und ausschalten.

Beim Entfernen der Modellverknüpfung auf der Scoring-Verzweigung wird auch der Aktualisierungsmodellstatus des Nuggets entfernt. Sie können das Entfernen der Modellverknüpfung über das Menü "Bearbeiten" oder die Symbolleiste widerrufen, aber damit wird auch der Aktualisierungsmodellstatus wieder eingestellt.

Auswählen des Aktualisierungsmodells

Wie die Scoring-Verzweigung wird auch die Verknüpfung zum Aktualisierungsmodell im Stream hervorgehoben. Das als Aktualisierungsmodell gewählte Modellnugget und die hervorgehobene Verknüpfung hängen davon ab, wie viele Nuggets sich im Stream befinden.

Einzelnes Modell im Stream

Wenn ein einzelnes, verknüpftes Modellnugget auf einer Scoring-Verzweigung als solches identifiziert wird, wird dieses Nugget das Aktualisierungsmodell des Streams.

Mehrere Modelle im Stream

Falls der Stream mehrere verknüpfte Nuggets enthält, wird das Aktualisierungsmodell wie folgt ausgewählt.

Wenn ein Modellnugget auf der Registerkarte "Bereitstellung" des Dialogfelds für die Streameigenschaften definiert wurde und sich im Stream befindet, wird dieses Nugget zum Aktualisierungsmodell.

Wenn auf der Registerkarte "Bereitstellung" kein Nugget definiert wurde oder wenn einer definiert wurde, sich aber nicht auf der Scoring-Verzweigung befindet, wird das Nugget, das sich am nächsten zum Endknoten befindet, zum Aktualisierungsmodell.

Wenn Sie anschließend die Auswahl aller Modellverknüpfungen als Aktualisierungsverknüpfungen aufheben, wird nur die Scoring-Verzweigung hervorgehoben, nicht die Verknüpfungen. Der Bereitstellungstyp wird auf "Nur Scoring" eingestellt.

Hinweis: Sie können bei Bedarf für eine der Verknüpfungen den Status "Ersetzen" festlegen, nicht aber für die andere. In diesem Fall ist das als Aktualisierungsmodell gewählte Modellnugget dasjenige, das über die Aktualisierungsverknüpfung verfügt und sich am nächsten zum Endknoten befindet, wenn die Scoring-Verzweigung festgelegt wird.

Keine Modelle im Stream

Wenn der Stream keine Modelle oder nur Modelle ohne Modellverknüpfungen enthält, ist der Bereitstellungstyp auf "Nur Scoring" eingestellt.

Überprüfen einer Scoring-Verzweigung auf Fehler

Wenn Sie die Scoring-Verzweigung festlegen, wird sie auf Fehler überprüft.

Bei Auffinden eines Fehlers wird die Scoring-Verzweigung in der entsprechenden Fehlerfarbe markiert und eine Fehlernachricht angezeigt. Sie können die Fehlerfarbe mithilfe einer Option für benutzerdefinierte Farben festlegen. Weitere Informationen finden Sie in „Festlegen von Anzeigoptionen“ auf Seite 226.

Gehen Sie bei Auffinden eines Fehlers wie folgt vor:

1. Korrigieren Sie den Fehler gemäß dem Inhalt der Fehlernachricht.
2. Klicken Sie im Hauptmenü auf Folgendes:
Tools > Streameigenschaften > Bereitstellung
und klicken Sie auf **Überprüfen**.
3. Wiederholen Sie diesen Vorgang, bis keine Fehler mehr gefunden werden.

Kapitel 12. Exportieren in externe Anwendungen

Exportieren in externe Anwendungen

IBM SPSS Modeler bietet eine Reihe von Mechanismen zum Exportieren des gesamten Data-Mining-Prozesses in externe Anwendungen, damit Sie die durchgeführten Arbeiten zur Vorbereitung von Daten und zum Erstellen von Modellen auch außerhalb von IBM SPSS Modeler gewinnbringend einsetzen können.

Im vorherigen Abschnitt wurde gezeigt, wie Sie Streams für ein Repository von IBM SPSS Collaboration and Deployment Services bereitstellen können, um dessen Mehrbenutzerzugriff, Aufgabenplanung und andere Funktionen zu nutzen. Ebenso lassen sich IBM SPSS Modeler-Streams zusammen mit folgenden Anwendungen nutzen:

- IBM SPSS Modeler Advantage
- Anwendungen, die Dateien im PMML-Format importieren und exportieren können

Weitere Informationen zum Verwenden von Streams in IBM SPSS Modeler Advantage finden Sie in „Öffnen eines Streams in IBM SPSS Modeler Advantage“.

Informationen zum Exportieren und Importieren von Modellen als PMML-Dateien, um Modelle mit beliebigen anderen Anwendungen, die dieses Format unterstützen, gemeinsam zu nutzen, finden Sie in „Importieren und Exportieren von Modellen als PMML“ auf Seite 212.

Öffnen eines Streams in IBM SPSS Modeler Advantage

IBM SPSS Modeler-Streams können zusammen mit der Thin-Client-Anwendung IBM SPSS Modeler Advantage verwendet werden. Es ist zwar möglich, benutzerdefinierte Anwendungen komplett in IBM SPSS Modeler Advantage zu erstellen, jedoch können Sie auch einen bereits in IBM SPSS Modeler erstellten Stream als Basis eines Anwendungsablaufs verwenden.

So öffnen Sie einen Stream in IBM SPSS Modeler Advantage:

1. Stellen Sie den Stream im Repository von IBM SPSS Collaboration and Deployment Services bereit, wobei die Option **Als Stream bereitstellen** aktiviert sein muss. Weitere Informationen finden Sie in „Bereitstellen von Streams“ auf Seite 204.
2. Klicken Sie in der Symbolleiste auf die Schaltfläche "In IBM SPSS Modeler Advantage öffnen" oder klicken Sie im Hauptmenü auf:

Datei > In IBM SPSS Modeler Advantage öffnen

1. Geben Sie gegebenenfalls Einstellungen für die Verbindung mit dem Repository ein. Weitere Informationen finden Sie in „Verbinden mit dem Repository“ auf Seite 192. Wenden Sie sich für spezielle Port-, Kennwort- und andere Verbindungsdaten an den lokalen Systemadministrator.

Hinweis: Auf dem Repository-Server muss die IBM SPSS Modeler Advantage-Software ebenfalls installiert sein.

1. Wählen Sie im Dialogfeld "Repository: Speichern" den Ordner aus, in dem Sie das Objekt speichern möchten, geben Sie alle weiteren Informationen an, die Sie aufzeichnen möchten, und klicken Sie auf die Schaltfläche **Speichern**. Weitere Informationen finden Sie in „Festlegen der Objekteigenschaften“ auf Seite 194.

Dadurch wird IBM SPSS Modeler Advantage gestartet, wobei der Stream bereits geöffnet ist. Der Stream wird in IBM SPSS Modeler geschlossen.

Importieren und Exportieren von Modellen als PMML

PMML (Predictive Model Markup Language) ist ein XML-Format zur Beschreibung von Data-Mining-Modellen und statistischen Modellen, einschließlich der Eingaben zu den Modellen, der zur Vorbereitung der Daten für das Data-Mining verwendeten Transformationen sowie der Parameter, die die Modelle selbst definieren. IBM SPSS Modeler kann PMML importieren und exportieren, wodurch es möglich ist, Modelle mit anderen Anwendungen zu teilen, die dieses Format unterstützen, wie beispielsweise IBM SPSS Statistics.

Weitere Informationen zu PMML finden Sie auf der Website der Data Mining Group (<http://www.dmg.org>).

So exportieren Sie ein Modell:

Der PMML-Export wird für die meisten der in IBM SPSS Modeler erstellten Modelltypen unterstützt. Weitere Informationen finden Sie in „Modelltypen, die PMML unterstützen“ auf Seite 213.

1. Klicken Sie mit der rechten Maustaste auf ein Modellnugget in der Modellpalette. (Alternativ können Sie auf ein Modellnugget im Erstellungsbereich klicken und das Menü "Datei" auswählen.)
2. Klicken Sie im Menü auf **PMML exportieren**.
3. Geben Sie im Dialogfeld "Exportieren" (oder "Speichern") ein Zielverzeichnis und einen eindeutigen Namen für das Modell an.

Anmerkung:

Im Dialogfeld "Benutzeroptionen" können Sie Optionen für den PMML-Export ändern. Klicken Sie im Hauptmenü auf Folgendes:

Tools > Optionen > Benutzeroptionen

Klicken Sie auf die Registerkarte "PMML".

Weitere Informationen finden Sie in „Festlegen von PMML-Exportoptionen“ auf Seite 228.

So importieren Sie ein als PMML gespeichertes Modell:

Modelle, die aus IBM SPSS Modeler oder einer anderen Anwendung als PMML exportiert wurden, können in die Modellpalette importiert werden. Weitere Informationen finden Sie in „Modelltypen, die PMML unterstützen“ auf Seite 213.

1. Klicken Sie in der Modellpalette mit der rechten Maustaste auf die Palette und wählen Sie im Menü die Option **PMML importieren** aus.
2. Wählen Sie die zu importierende Datei aus und geben Sie nach Bedarf Optionen für Variablenbeschriftungen an.
3. Klicken Sie auf **Öffnen**.

Variablenbeschriftungen verwenden, sofern im Modell vorhanden. Im PMML-Code können sowohl die Variablennamen als auch die Variablenbeschriftungen (beispielsweise "Referrer ID" für *RefID*) für Variablen im Datenwörterbuch gefunden. Wählen Sie diese Option aus, um Variablenbeschriftungen zu verwenden, wenn diese im ursprünglich exportierten PMML-Code vorhanden sind.

Wenn Sie die Optionen zur Variablenbeschriftung ausgewählt haben, im PMML-Code jedoch keine Variablenbeschriftungen vorhanden sind, werden die Variablennamen wie üblich verwendet.

Modelltypen, die PMML unterstützen

PMML-Export

IBM SPSS Modeler-Modelle. Die folgenden in IBM SPSS Modeler erstellten Modelle können als PMML 4.0 exportiert werden:

- C&R-Baum
- QUEST
- CHAID
- Neuronales Netz
- C5.0
- Logistische Regression
- Genlin
- SVM
- Apriori
- Carma
- K-Means
- Kohonen
- TwoStep
- TwoStep-AS
- GLMM (PMML wird für alle GLMM-Modelle exportiert, PMML hat jedoch nur feste Effekte)
- Entscheidungsliste
- Cox
- Sequenz (Scoring für PMML-Sequenzmodelle wird nicht unterstützt)
- Random Trees
- Tree-AS
- Linear
- Linear-AS
- Regression
- Logistisch
- GLE
- LSVM
- Anomalieerkennung
- KNN
- Assoziationsregeln

Datenbankeigene Modelle. Bei Modellen, die mithilfe von datenbankeigenen Algorithmen generiert wurden, ist der PMML-Export nicht verfügbar. Modelle, die mithilfe von Analysis Services von Microsoft oder mit Oracle Data Miner erstellt wurden, können nicht exportiert werden.

PMML-Import

IBM SPSS Modeler kann PMML-Modelle importieren und scoren, die von aktuellen Versionen aller IBM SPSS Statistics-Produkte erstellt wurden, darunter Modelle, die aus IBM SPSS Modeler exportiert wurden, sowie Modell- bzw. Transformations-PMML, die von IBM SPSS Statistics 17.0 oder höher generiert wurde. Dies gilt also im Grunde für jegliche PMML, die die Scoring-Engine scoren kann - mit folgenden Ausnahmen:

- Apriori- und CARMA-Modelle sowie Anomalieerkennungsmodelle, Sequenzmodelle und Assoziationsregelmodelle können nicht importiert werden.

- PMML-Modelle können nach dem Import in IBM SPSS Modeler nicht durchsucht werden, obwohl sie für das Scoring verwendet werden können. (Dies gilt auch für Modelle, die ursprünglich aus IBM SPSS Modeler exportiert wurden. Um diese Einschränkung zu vermeiden, sollten Sie das betreffende Modell als generierte Modelldatei (*.gm) und nicht als PMML exportieren.
- Eine eingeschränkte Validierung findet beim Import statt, aber die vollständige Validierung erfolgt beim Versuch, das Modell zu scoren. Daher kann der Import erfolgreich durchgeführt werden, das Scoring aber fehlschlagen oder falsche Ergebnisse erzeugen.

Anmerkung: Bei PMML eines anderen Anbieters, die in IBM SPSS Modeler importiert wird, versucht IBM SPSS Modeler ein Scoring für die gültige PMML durchzuführen, die erkannt und gescort werden kann. Es ist jedoch nicht garantiert, dass für die gesamte PMML ein Scoring durchgeführt wird oder dass das Scoring dasselbe Ergebnis bringt wie das Scoring der Anwendung, aus der sie generiert wurde.

Kapitel 13. Projekte und Berichte

Einführung in Projekte

Ein **Projekt** ist eine Gruppe von Dateien, die in Bezug zu einer Data-Mining-Aufgabe stehen. Zu Projekten gehören Datenstreams, Diagramme, generierte Modelle, Berichte und alles andere, was Sie in IBM SPSS Modeler erstellt haben. Auf den ersten Blick kann der Eindruck entstehen, IBM SPSS Modeler-Projekte seien einfach eine Methode zur Organisation von Ausgaben, doch sie haben noch viele weitere Funktionen. Mit Projekten haben Sie folgende Möglichkeiten:

- Erstellen von Anmerkungen für jedes Objekt in der Projektdatei.
- Verwenden der CRISP-DM-Methodologie als Orientierung beim Data-Mining. Projekte können außerdem ein CRISP-DM-Hilfesystem enthalten, das Einzelheiten und reale Beispiele für Data-Mining mit CRISP-DM bietet.
- Hinzufügen von nicht aus IBM SPSS Modeler stammenden Objekten zum Projekt, beispielsweise einer PowerPoint-Präsentation zur Vorstellung Ihrer Data-Mining-Ziele oder White Papers für die Algorithmen, die Sie verwenden möchten.
- Erstellen von umfassenden und gleichzeitig einfachen Aktualisierungsberichten auf der Grundlage Ihrer Anmerkungen. Diese Berichte können in HTML generiert werden, um eine einfache Veröffentlichung im Intranet Ihrer Organisation zu ermöglichen.

Hinweis: Wenn der Projektbereich im IBM SPSS Modeler-Fenster nicht sichtbar ist, klicken Sie auf die Option **Projekt** im Menü "Ansicht".

Sie haben zwei Möglichkeiten, die einem Projekt hinzugefügten Objekte anzuzeigen: in der Ansicht **Klassen** und in der Ansicht **CRISP-DM**. Alles, was Sie zu einem Projekt hinzufügen, wird zu beiden Ansichten hinzugefügt. Sie können zwischen den Ansichten umschalten, um die Organisation zu erstellen, die am besten geeignet ist.

CRISP-DM-Ansicht

Durch Unterstützung von CRISP-DM (Cross-Industry Standard Process for Data Mining) bieten IBM SPSS Modeler-Projekte eine in der Branche bewährte, nicht gesetzlich geschützte Methode zur Organisation der einzelnen Teile Ihrer Data-Mining-Arbeit. CRISP-DM verwendet sechs Phasen zur Beschreibung des Prozesses vom Start (Zusammenstellen der Geschäftsanforderungen) bis zum Ende (Bereitstellen der Ergebnisse). Auch wenn einige Phasen bei IBM SPSS Modeler normalerweise keine Arbeit beinhalten, umfasst der Projektbereich alle sechs Phasen, sodass alle zum Projekt gehörenden Materialien an einer zentralen Stelle gespeichert und verfolgt werden können. Zum Geschäftsverständnis gehören üblicherweise das Sammeln von Anforderungen sowie Besprechungen mit Kollegen zur Bestimmung der Ziele und weniger die Arbeit mit den Daten in IBM SPSS Modeler. Mit dem Projektbereich können Sie Ihre Aufzeichnungen aus solchen Besprechungen im Ordner *Geschäftsverständnis* speichern, um sie später nachlesen oder in Berichte aufnehmen zu können.

Die CRISP-DM-Ansicht im Projektbereich ist mit einem eigenen Hilfesystem ausgestattet, das Sie durch den Data-Mining-Zyklus führt. In IBM SPSS Modeler ist der Zugriff auf diese Hilfe durch Klicken auf **CRISP-DM-Hilfe** im Hilfemenü möglich.

Hinweis: Wenn der Projektbereich nicht im Fenster sichtbar ist, klicken Sie auf die Option **Projekt** im Menü "Ansicht".

Einrichten der Standardprojektphase

Die zu einem Projekt hinzugefügten Objekte werden zu einer Standardphase von CRISP-DM hinzugefügt. Daher müssen Sie die Objekte manuell entsprechend der Data-Mining-Phase organisieren, in der Sie sie verwendet haben. Es ist sinnvoll, den Standardordner auf die Phase einzustellen, in der Sie gerade arbeiten.

So legen Sie fest, welche Phase als Standard verwendet wird:

1. Klicken Sie in der CRISP-DM-Ansicht mit der rechten Maustaste auf den Ordner für die Phase, die als Standard festgelegt werden soll.
2. Klicken Sie im Menü auf **Als Standard festlegen**.

Der Standardordner wird fett gedruckt angezeigt.

Klassenansicht

Mit der Ansicht "Klassen" im Projektbereich können Sie Ihre Arbeit in IBM SPSS Modeler nach den erstellten Objekttypen in Kategorien organisieren. Die gespeicherten Objekte können zu einer der folgenden Kategorien hinzugefügt werden:

- Streams
- Knoten
- Modelle
- Tabellen, Diagramme, Berichte
- Andere (Nicht-IBM SPSS Modeler-Dateien, wie beispielsweise Präsentationen oder White Papers, die für Ihre Data-Mining-Arbeit relevant sind)

Wenn Objekte zur Klassenansicht hinzugefügt werden, werden sie auch zum Standardphasenordner in der CRISP-DM-Ansicht hinzugefügt.

Hinweis: Wenn der Projektbereich nicht im Fenster sichtbar ist, klicken Sie auf die Option **Projekt** im Menü "Ansicht".

Erstellen von Projekten

Bei einem Projekt handelt es sich im Wesentlichen um eine Datei mit Verweisen auf alle Dateien, die Sie dem Projekt zuordnen. Projektelemente werden also sowohl einzeln als auch als Verweis in der Projektdatei (.cpj) gespeichert. Aufgrund dieser referenziellen Struktur sollten Sie auf folgende Punkte achten:

- Projektelemente müssen zunächst einzeln gespeichert werden, bevor sie zu einem Projekt hinzugefügt werden. Wenn ein Element nicht gespeichert ist, werden Sie aufgefordert, es zu speichern, bevor Sie es zum aktuellen Projekt hinzufügen.
- Einzeln aktualisierte Objekte, beispielsweise Streams, werden auch in der Projektdatei aktualisiert.
- Beim manuellen Verschieben oder Löschen von Objekten (beispielsweise Streams, Knoten und Ausgabeobjekte) aus dem Dateisystem werden die Verknüpfungen in der Projektdatei ungültig.

Erstellen eines neuen Projekts

Neue Projekte können problemlos im IBM SPSS Modeler-Fenster erstellt werden. Sie können entweder mit dem Erstellen eines Projekts beginnen, wenn noch keines geöffnet ist, oder Sie können ein bestehendes Objekt schließen und von vorne anfangen.

Klicken Sie im Hauptmenü auf Folgendes:

Datei > Projekt > Neues Projekt...

Hinzufügen zu einem Projekt

Nach dem Erstellen oder Öffnen von Projekten stehen Ihnen mehrere Methoden zur Verfügung, um Objekte, wie beispielsweise Datenstreams, Knoten und Berichte, hinzuzufügen.

Hinzufügen von Objekten aus den Managern

Mithilfe der Manager in der rechten oberen Ecke des IBM SPSS Modeler-Fensters können Sie Streams oder Ausgaben hinzufügen.

1. Wählen Sie ein Objekt, beispielsweise eine Tabelle oder einen Stream auf einer der Manager-Registerkarten aus.
2. Klicken Sie mit der rechten Maustaste und klicken Sie auf die Option **Zu Projekt hinzufügen**.
Wenn das Objekt zuvor gespeichert wurde, wird es automatisch zum entsprechenden Objektordner (in der Klassenansicht) oder zum Ordner der Standardphase (in der CRISP-DM-Ansicht) hinzugefügt.
3. Alternativ können Sie durch Ziehen und Ablegen Objekte aus den Managern in den Projektbereich verschieben.

Hinweis: Sie werden möglicherweise dazu aufgefordert, das Objekt zuerst zu speichern. Beim Speichern muss **Datei zu Projekt hinzufügen** im Dialogfeld "Speichern" ausgewählt sein. Dadurch wird das Objekt nach dem Speichern automatisch zum Projekt hinzugefügt.

Hinzufügen von Knoten aus dem Erstellungsbereich

Mit dem Dialogfeld "Speichern" können Sie einzelne Knoten aus dem Streamerstellungsbereich hinzufügen.

1. Wählen Sie einen Knoten im Erstellungsbereich aus.
2. Klicken Sie mit der rechten Maustaste und klicken Sie auf die Option **Knoten speichern**. Klicken Sie alternativ im Hauptmenü auf Folgendes:
Bearbeiten > Knoten > Knoten speichern...
3. Wählen Sie im Dialogfeld "Speichern" die Option **Datei zum Projekt hinzufügen** aus.
4. Erstellen Sie einen Namen für den Knoten und klicken Sie auf **Speichern**.

Dadurch wird die Datei gespeichert und zum Projekt hinzugefügt. Knoten werden in der Klassenansicht zum Ordner *Knoten* und in der CRISP-DM-Ansicht zum Ordner der Standardphase hinzugefügt.

Hinzufügen externer Dateien

Sie können eine große Vielzahl von Nicht-IBM SPSS Modeler-Objekten zu einem Projekt hinzufügen. Dies ist sinnvoll, wenn der gesamte Data-Mining-Vorgang innerhalb von IBM SPSS Modeler verwaltet werden soll. Sie können beispielsweise Verknüpfungen mit Daten, Notizen, Präsentationen und Grafiken in einem Objekt speichern. In der CRISP-DM-Ansicht können externe Dateien zum Ordner Ihrer Wahl hinzugefügt werden. In der Klassenansicht können externe Dateien nur im Ordner *Andere* gespeichert werden.

So fügen Sie externe Dateien zu einem Projekt hinzu:

1. Ziehen Sie die Dateien vom Desktop auf das Projekt.
oder
2. Klicken Sie mit der rechten Maustaste auf den Zielordner in der CRISP-DM- oder Klassen-Ansicht.
3. Klicken Sie im Menü auf die Option **Zu Ordner hinzufügen**.
4. Wählen Sie eine Datei im Dialogfeld aus und klicken Sie auf **Öffnen**.

Dadurch wird ein Verweis zum ausgewählten Objekt innerhalb der IBM SPSS Modeler-Projekte hergestellt.

Übertragen von Projekten an IBM SPSS Collaboration and Deployment Services Repository

Sie können ein gesamtes Projekt, einschließlich aller Komponentendateien, in einem Schritt an IBM SPSS Collaboration and Deployment Services Repository übertragen. Alle Objekte, die sich bereits am Zielspeicherort befinden, werden nicht verschoben. Die Funktion lässt sich auch umgekehrt nutzen: Sie können gesamte Projekte von IBM SPSS Collaboration and Deployment Services Repository in Ihr lokales Dateisystem übertragen.

Übertragen von Projekten

Stellen Sie sicher, dass das zu übertragende Projekt im Projektbereich geöffnet ist.

So übertragen Sie ein Projekt:

1. Klicken Sie mit der rechten Maustaste auf den Projektordner und klicken Sie auf die Option **Projekt übertragen**.
2. Wenn Sie dazu aufgefordert werden, melden Sie sich bei IBM SPSS Collaboration and Deployment Services Repository an.
3. Geben Sie den neuen Speicherort für das Projekt an und klicken Sie auf **OK**.

Festlegen der Projekteigenschaften

Sie können den Inhalt und die Dokumentation eines Projekts mithilfe des Dialogfelds "Projekteigenschaften" anpassen. So können Sie auf die Projekteigenschaften zugreifen:

1. Klicken Sie mit der rechten Maustaste auf ein Objekt oder einen Ordner im Projektbereich und klicken Sie auf die Option **Projekteigenschaften**.
2. Klicken Sie auf die Registerkarte **Projekt**, um grundlegende Projektinformationen anzugeben.

Erstellt. Zeigt das Erstellungsdatum des Projekts an (nicht bearbeitbar).

Zusammenfassung. Sie können eine Übersicht für Ihr Data-Mining-Projekt eingeben, die im Projektbericht angezeigt werden soll.

Inhalt. Listet Typ und Anzahl der von der Projektdatei referenzierten Komponenten auf (nicht bearbeitbar).

Ungespeicherte Objekte speichern als. Gibt an, ob ungespeicherte Objekte im lokalen Dateisystem oder im Repository gespeichert werden sollen. Weitere Informationen finden Sie in „Informationen zu IBM SPSS Collaboration and Deployment Services Repository“ auf Seite 191.

Objektverweise beim Laden des Projekts aktualisieren. Wählen Sie diese Option aus, um die Verweise des Projekts auf seine Komponenten zu aktualisieren. *Hinweis:* Die zu einem Projekt hinzugefügten Dateien werden nicht in der Projektdatei selbst gespeichert. Stattdessen wird ein Verweis auf die Dateien im Projekt gespeichert. Das bedeutet, dass durch Verschieben oder Löschen einer Datei das betreffende Objekt aus dem Projekt entfernt wird.

Erstellen von Anmerkungen für ein Projekt

Der Projektbereich bietet eine Reihe von Möglichkeiten, um Anmerkungen für Ihre Data-Mining-Arbeit zu erstellen. Anmerkungen auf der Projektebene werden häufig verwendet, um weiter reichende Ziele und Entscheidungen aufzuzeichnen, während die Anmerkungen auf der Ordner- oder Knotenebene zusätzliche Details bieten. Die Registerkarte "Anmerkungen" bietet genügend Raum zum Dokumentieren von Informationen für die Projektebene, wie beispielsweise den Ausschluss von Daten mit nicht wiederherstellbaren fehlenden Daten oder vielversprechenden Hypothesen, die während der Datenexploration aufgestellt wurden.

So erstellen Sie Anmerkungen zu einem Projekt:

1. Wählen Sie den Projektordner entweder in der CRISP-DM- oder in der Klassenansicht aus.
2. Klicken Sie mit der rechten Maustaste auf den Ordner und klicken Sie auf die Option **Projekteigenschaften**.
3. Klicken Sie auf die Registerkarte **Anmerkungen**.
4. Geben Sie Schlüsselwörter und Text zur Beschreibung des Projekts ein.

Ordneigenschaften und Anmerkungen

Einzelne Projektordner (in der CRISP-DM-Ansicht und der Klassenansicht) können mit Anmerkungen versehen werden. In der CRISP-DM-Ansicht kann dies eine sehr effektive Methode sein, um die Ziele Ihrer Organisation für die einzelnen Phasen des Data-Mining zu dokumentieren. Beispielsweise können Sie mit dem Anmerkungsstool im Ordner *Geschäftsverständnis* Hinweise aufnehmen wie "Das Geschäftsziel für diese Studie besteht darin, die Abwanderung hochwertiger Kunden zu reduzieren". Dieser Text könnte dann durch Auswahl der Option **In Bericht einschließen** automatisch in den Projektbericht aufgenommen werden.

So erstellen Sie Anmerkungen zu einem Ordner:

1. Wählen Sie einen Ordner im Projektbereich aus.
2. Klicken Sie mit der rechten Maustaste auf den Ordner und klicken Sie auf die Option **Ordneigenschaften**.

In der CRISP-DM-Ansicht bestehen die Anmerkungen für die Ordner aus einer Übersicht über den Zweck der einzelnen Phasen sowie in Hinweisen zur Durchführung der relevanten Data-Mining-Aufgaben. Sie können diese Anmerkungen nach Belieben entfernen oder bearbeiten.

Name. In diesem Bereich wird der Name des ausgewählten Felds angezeigt.

QuickInfo-Text. Dient zum Erstellen benutzerdefinierter QuickInfos, die angezeigt werden, wenn der Mauszeiger über einen Projektordner geführt wird. Dies ist vor allem in der CRISP-DM-Ansicht sinnvoll, beispielsweise um einen schnellen Überblick über die Ziele der einzelnen Phasen zu bieten oder den Status einer Phase (z. B. "In Arbeit" oder "Abgeschlossen") anzugeben.

Anmerkungsfeld. Dieses Feld ist für längere Anmerkungen gedacht, die im Projektbericht zusammengestellt werden können. Die CRISP-DM-Ansicht enthält eine Beschreibung der einzelnen Data-Mining-Phasen in der Anmerkung, Sie können diese jedoch für Ihr Projekt nach Belieben anpassen.

In Bericht einschließen. Um die Anmerkung in die Berichte einzuschließen, wählen Sie **In Bericht einschließen** aus.

Objekteigenschaften

Sie können Objekteigenschaften anzeigen und auswählen, ob einzelne Objekte in den Projektbericht aufgenommen werden sollen. So können Sie auf die Objekteigenschaften zugreifen:

1. Klicken Sie mit der rechten Maustaste auf ein Objekt im Projektbereich.
2. Klicken Sie im Menü auf die Option **Objekteigenschaften**.

Name. In diesem Bereich wird der Name des gespeicherten Objekts aufgeführt.

Pfad. In diesem Bereich wird der Speicherort des gespeicherten Objekts aufgeführt.

In Bericht einschließen. Wählen Sie diese Option aus, um die Objektdetails in einem generierten Bericht einzuschließen.

Schließen von Projekten

Wenn Sie IBM SPSS Modeler beenden oder ein neues Projekt öffnen, wird die vorhandene Projektdatei (.cpj) geschlossen.

Einige Dateien, die mit dem Projekt assoziiert sind (z. B. Streams, Knoten oder Grafiken), sind möglicherweise noch geöffnet. Wenn diese Dateien geöffnet bleiben sollen, antworten Sie mit Nein auf die Nachricht ... **Sollen diese Dateien gespeichert und geschlossen werden?**

Wenn Sie zugehörige Dateien nach dem Schließen eines Projekts bearbeiten und speichern, werden diese aktualisierten Versionen beim nächsten Öffnen des Projekts in das Projekt aufgenommen. Um dies zu verhindern, müssen Sie die betreffende Datei aus dem Projekt entfernen oder unter einem anderen Dateinamen speichern.

Generieren von Berichten

Eine der nützlichsten Funktionen von Projekten ist die Möglichkeit, Berichte auf der Grundlage der Projektelemente und der Anmerkungen zu erstellen. Dies ist eine äußerst wichtige Komponente des effektiven Data-Mining, wie in der gesamten CRISP-DM-Methodologie. Sie können einen Bericht direkt in einem von mehreren Dateitypen erstellen oder in einem Ausgabefenster auf dem Bildschirm, um ihn sofort betrachten zu können. Von dort können Sie den Bericht drucken, speichern oder in einem Web-Browser anzeigen. Sie können gespeicherte Berichte an andere Personen in Ihrer Organisation verteilen.

Berichte werden häufig mehrmals während des Data-Mining-Prozesses aus den Berichtdateien generiert, um sie an die am Projekt beteiligten Personen weiterzugeben. Der Bericht entnimmt Informationen zu den referenzierten Objekten aus der Projektdatei sowie aus den erstellten Anmerkungen (sofern vorhanden). Sie können die Berichte entweder auf der Grundlage der Klassenansicht oder der CRISP-DM-Ansicht erstellen.

So generieren Sie einen Bericht:

1. Wählen Sie den Projektordner entweder in der CRISP-DM- oder in der Klassenansicht aus.
2. Klicken Sie mit der rechten Maustaste auf den Ordner und klicken Sie auf die Option **Projektbericht**.
3. Geben Sie die Berichtsoptionen an und klicken Sie auf **Bericht erstellen**.

Die Optionen im Berichtdialogfeld bieten mehrere Möglichkeiten zur Generierung des erforderlichen Berichtstyps:

Ausgabename. Geben Sie den Namen des Ausgabefensters an, wenn die Ausgabe des Berichts an den Bildschirm gesendet werden soll. Sie können einen benutzerdefinierten Namen angeben oder IBM SPSS Modeler das Fenster automatisch benennen lassen.

Ausgabe auf Bildschirm. Wählen Sie diese Option aus, um den Bericht zu generieren und in einem Ausgabefenster anzuzeigen. Beachten Sie, dass Sie die Möglichkeit haben, den Bericht aus dem Ausgabefenster in verschiedene Dateitypen zu exportieren.

Ausgabe in Datei. Wählen Sie diese Option aus, um den Bericht zu erstellen und als Datei mit dem in der Liste "Dateityp" angegebenen Typ zu speichern.

Dateiname. Geben Sie einen Dateinamen für den generierten Bericht an. Die Dateien werden standardmäßig im Verzeichnis `\bin` von IBM SPSS Modeler gespeichert. Mit der Auslassungsschaltfläche (...) können Sie einen anderen Speicherort angeben.

Dateityp. Die folgenden Dateitypen stehen zur Auswahl:

- **HTML-Dokument.** Der Bericht wird als einzelne HTML-Datei gespeichert. Wenn Ihr Bericht Diagramme enthält, werden diese als PNG-Dateien gespeichert und von der HTML-Datei referenziert. Wenn Sie Ihren Bericht im Internet veröffentlichen, müssen Sie die HTML-Datei und alle von ihr referenzierten Bilder hochladen.
- **Textdokument.** Der Bericht wird als einzelne Textdatei gespeichert. Wenn der Bericht Diagramme enthält, werden nur die Referenzen auf Dateinamen und Pfad in den Bericht aufgenommen.
- **Microsoft Word-Dokument.** Der Bericht wird als Einzeldokument gespeichert. Alle Diagramme sind direkt in das Dokument eingebettet.
- **Microsoft Excel-Dokument.** Der Bericht wird als einzelnes Arbeitsblatt gespeichert. Alle Diagramme sind direkt in das Arbeitsblatt eingebettet.
- **Microsoft PowerPoint-Dokument.** Jede Phase wird auf einer Folie angezeigt. Alle Diagramme sind direkt in die PowerPoint-Folien eingebettet.
- **Ausgabeobjekt.** Wenn diese Datei (.cou) in IBM SPSS Modeler geöffnet wird, stimmt sie mit der oben genannten Option **Ausgabe auf Bildschirm** in der Gruppe **Berichtformat** überein.

Hinweis: Zum Exportieren einer Microsoft Office-Datei muss die entsprechende Anwendung installiert sein.

Titel. Geben Sie einen Titel für den Bericht an.

Berichtsstruktur. Wählen Sie **CRISP-DM** oder **Klassen** aus. Die CRISP-DM-Ansicht bietet einen Statusbericht mit überblickartigen Synapsen sowie Einzelheiten zu den einzelnen Data-Mining-Phasen. Bei der Ansicht "Klassen" handelt es sich um eine objektbasierte Ansicht, die besser für die interne Verfolgung von Daten und Streams geeignet ist.

Autor. Der Standardbenutzername wird angezeigt. Sie können ihn jedoch ändern.

Bericht enthält. Wählen Sie eine Methode für die Aufnahme von Objekten in den Bericht aus. Wählen Sie **alle Ordner und Objekte** aus, um alle zur Projektdatei hinzugefügten Elemente einzuschließen. Sie können auch nur die Elemente einschließen lassen, bei denen in den Objekteigenschaften **In Bericht einschließen** ausgewählt ist. Alternativ können Sie, um ungemeldete Elemente zu überprüfen, festlegen, dass nur Elemente aufgenommen werden sollen, die für den Ausschluss markiert sind (bei denen **In Bericht einschließen** inaktiviert ist).

Auswählen. Mit dieser Option können Sie Projektaktualisierungen bereitstellen, indem Sie nur Elemente der Kategorie **zuletzt verwendete Elemente** in den Bericht aufnehmen. Alternativ können Sie ältere und vielleicht ungelöste Probleme nachverfolgen, indem Sie die Parameter für **alte Elemente** festlegen. Wählen Sie **alle Elemente** aus, um die Zeit als Berichtsparameter zu ignorieren.

Sortieren nach. Sie können eine Kombination der folgenden Objektmerkmale auswählen, um sie innerhalb eines Ordners zu sortieren:

- **Typ.** Objekte werden nach Typ gruppiert.
- **Name.** Objekte werden alphabetisch angeordnet.
- **Einfügedatum.** Objekte werden nach dem Datum ihrer Hinzufügung zum Projekt sortiert.

Speichern und Exportieren von generierten Berichten

Ein auf dem Bildschirm generierter Bericht wird in einem neuen Ausgabefenster angezeigt. Alle im Bericht enthaltenen Diagramme werden als Inlinebilder angezeigt.

Berichtsterminologie

Im Bericht wird die Gesamtzahl der Knoten in den einzelnen Streams aufgeführt. Die Zahlen werden unter folgenden Überschriften angegeben, wobei IBM SPSS Modeler-Terminologie, nicht CRISP-DM-Terminologie, verwendet wird:

- **Datenleser.** Quellenknoten.
- **Datenschreiber.** Exportknoten.
- **Modellersteller.** Erstellungs- bzw. Modellierungsknoten.
- **Modellanwender.** Generierte Modelle, auch als "Nuggets" bezeichnet.
- **Ausgabeersteller.** Diagramm- oder Ausgabeknoten.
- **Andere.** Alle anderen Knoten, die zum Projekt gehören. Dazu gehören beispielsweise die Knoten, die auf der Registerkarte "Feldoperationen" oder "Datensatzoperationen" auf der Knotenpalette verfügbar sind.

So speichern Sie einen Bericht:

1. Klicken Sie im Menü "Datei" auf die Option **Speichern**.

2. Geben Sie einen Dateinamen an.

Der Bericht wird als Ausgabeobjekt gespeichert.

So exportieren Sie einen Bericht:

3. Klicken Sie im Menü "Datei" auf die Option **Exportieren** und wählen Sie den zu exportierenden Dateityp aus.

4. Geben Sie einen Dateinamen an.

Der Bericht wird im gewählten Format gespeichert.

Ein Export ist in folgende Dateitypen möglich:

- HTML
- Text
- Microsoft Word
- Microsoft Excel
- Microsoft PowerPoint

Hinweis: Zum Exportieren einer Microsoft Office-Datei muss die entsprechende Anwendung installiert sein.

Mit den Schaltflächen oben im Fenster haben Sie folgende Möglichkeiten:

- Drucken Sie den Bericht.
- Zeigen Sie den Bericht in einem externen Web-Browser als HTML an.

Kapitel 14. Anpassen von IBM SPSS Modeler

Anpassen von IBM SPSS Modeler-Optionen

Es gibt eine Reihe von Operationen, mit denen Sie IBM SPSS Modeler an Ihre Bedürfnisse anpassen können. Diese Anpassung besteht hauptsächlich aus der Festlegung bestimmter Benutzeroptionen, wie Arbeitsspeicherzuordnung, Standardverzeichnisse und die Verwendung von Klängen und Farben. Außerdem können Sie die Knotenpalette anpassen, die sich unten im IBM SPSS Modeler-Fenster befindet.

Festlegen von IBM SPSS Modeler-Optionen

Es gibt verschiedene Möglichkeiten, Optionen für IBM SPSS Modeler anzupassen und festzulegen:

- Legen Sie Systemoptionen, wie Speicherbelegung und Ländereinstellung fest, durch Klicken auf **Systemoptionen** im Menü **Tools > Optionen**.
- Legen Sie Benutzeroptionen, wie Schriftarten und Farben für die Anzeige fest, durch Klicken auf **Benutzeroptionen** im Menü **Tools > Optionen**.
- Legen Sie das Verzeichnis für die Anwendungen fest, die mit IBM SPSS Modeler verwendet werden können, durch Klicken auf **Hilfsanwendungen** im Menü **Tools > Optionen**.
- Legen Sie die in IBM SPSS Modeler verwendeten Standardverzeichnisse fest durch Klicken auf **Verzeichnis festlegen** oder **Serververzeichnis festlegen** im Menü "Datei".

Sie können auch Optionen festlegen, die für einige Ihrer Streams oder für alle Streams gelten. Weitere Informationen finden Sie in „Festlegen von Optionen für Streams“ auf Seite 44.

Systemoptionen

Sie können die bevorzugte Sprache oder Ländereinstellung für IBM SPSS Modeler festlegen, indem Sie im Menü **Tools > Optionen** auf **Systemoptionen** klicken. Hier können Sie auch die maximale Speicherbelegung für SPSS Modeler festlegen und angeben, wie oft Datenströme automatisch gespeichert werden sollen. Beachten Sie, dass die in diesem Dialogfeld vorgenommenen Änderungen so lange unwirksam sind, bis Sie SPSS Modeler erneut starten.

Maximale Speichergröße. Wählen Sie diese Option aus, um einen Grenzwert in Megabyte für die Speicherbelegung von IBM SPSS Modeler festzulegen. Auf einigen Plattformen beschränkt SPSS Modeler die Prozessgröße, um die Belastung der Computer mit begrenzten Ressourcen oder hoher Auslastung zu reduzieren. Wenn Sie große Datenmengen bearbeiten, führt dies eventuell zu einer Fehlermeldung, dass nicht genügend Speicher vorhanden ist. Sie können die Speicherauslastung verbessern, indem Sie einen neuen Schwellenwert angeben.

Wenn z. B. versucht wird, einen sehr großen Entscheidungsbaum anzuzeigen, kann dies zu einem Speicherfehler führen. In diesem Fall wird empfohlen, dass Sie den Speicher auf den Maximalwert von 4096 MB vergrößern. Beenden Sie in Fällen wie diesen, wo Sie vermutlich sehr große Datenmengen verarbeiten, SPSS Modeler nach dem Erhöhen des zugeordneten Speichers und starten Sie das Programm aus der Befehlszeile, um sicherzustellen, dass die maximale Speicherkapazität für die Verarbeitung Ihrer Daten verwendet wird.

Geben Sie Folgendes in ein Fenster mit einer Eingabeaufforderung ein, um die Software in einer Befehlszeile zu starten (wenn SPSS Modeler in der Standardposition installiert ist):

```
C:\Programme\IBM\SPSS\Modeler\18.2.1\bin\modelerclient.exe" -J-Xss4096M
```

Ländereinstellungen aus System verwenden. Diese Option wird standardmäßig ausgewählt und auf "Englisch" (USA) gesetzt. Inaktivieren Sie diese Option, um eine andere Sprache aus der Liste mit verfügbaren Sprachen und Ländereinstellungen auszuwählen.

Intervall für automatische Streamspeicherung (Minuten). Geben Sie an, wie oft SPSS Modeler Datenströme automatisch speichern soll. Der Höchstwert ist 60 Minuten, der Mindestwert ist 1 Minute und der Standardwert ist 5 Minuten.

Speicherverwaltung

Zusätzlich zur Einstellung **Maximale Speichergröße**, die Sie im Dialogfeld "Systemoptionen" angegeben haben, gibt es mehrere Möglichkeiten zur Optimierung der Speicherbelegung:

- Ändern Sie die Option **Maximale Elemente für nominale Felder** im Dialogfeld für die Streameigenschaften. Diese Option legt eine maximale Anzahl Mitglieder für nominale Felder fest, nach denen das Messniveau des Felds zu *ohne Typ* wird. Weitere Informationen finden Sie in „Festlegen allgemeiner Optionen für Streams“ auf Seite 44.
- Zwingen Sie IBM SPSS Modeler, Speicher freizugeben, indem Sie in die rechte untere Ecke des Fensters klicken, in der der von IBM SPSS Modeler belegte Speicher und die zugeordnete Speichergröße angezeigt werden (*xx MB/xx MB*). Wenn Sie in diesen Bereich klicken, wird dieser dunkler und die Zahlen für die Speicherzuordnung werden niedriger. Sobald der Bereich erneut die reguläre Farbe annimmt, hat IBM SPSS Modeler den gesamten möglichen Speicher freigegeben.

Festlegen von Standardverzeichnissen

Sie können das für Dateibrowser und die Ausgabe verwendete Standardverzeichnis angeben, indem Sie im Menü "Datei" **Verzeichnis festlegen** oder **Serververzeichnis festlegen** auswählen.

- **Verzeichnis festlegen.** Sie können diese Option verwenden, um das Arbeitsverzeichnis festzulegen. Das Standardarbeitsverzeichnis basiert auf dem Installationspfad Ihrer IBM SPSS Modeler-Version oder auf dem Befehlszeilenpfad, der zum Aufrufen von IBM SPSS Modeler verwendet wird. Im lokalen Modus ist das Arbeitsverzeichnis der Pfad, der für alle clientseitigen Operationen und Ausgabedateien verwendet wird (falls diese mit relativen Pfaden angegeben werden).
- **Serververzeichnis festlegen.** Die Option "Serververzeichnis festlegen" im Menü "Datei" wird aktiviert, wenn eine Remote-Serververbindung hergestellt wird. Verwenden Sie diese Option, um das Standardverzeichnis für alle Server- und Datendateien festzulegen, die für die Ein- und Ausgabe angegeben sind. Das Serverstandardverzeichnis lautet *\$CLEO/data*, wobei *\$CLEO* das Verzeichnis ist, in dem die Serverversion von IBM SPSS Modeler installiert ist. Mithilfe der Befehlszeile können Sie diesen Standardwert auch überschreiben, indem Sie das Flag `-server_directory` mit dem Befehlszeilenargument `modelerclient` verwenden.

Festlegen von Benutzeroptionen

Sie können allgemeine Optionen für IBM SPSS Modeler festlegen, indem Sie im Menü **Tools > Optionen** den Eintrag **Benutzeroptionen** auswählen. Diese Optionen gelten für alle in IBM SPSS Modeler verwendeten Streams.

Die folgenden Optionen können festgelegt werden, indem Sie auf die entsprechende Registerkarte klicken:

- Benachrichtigungsoptionen, wie Modellüberschreibung und Fehlernachrichten.
- Anzeigeoptionen, wie Diagramm- und Hintergrundfarben.
- Syntaxfarbanzeigeoptionen.
- PMML-Exportoptionen, die beim Export von Modellen in Predictive Model Markup Language (PMML) verwendet werden.
- Informationen zum Benutzer bzw. Autor, wie Name, Initialen und E-Mail-Adresse. Diese Informationen können auf der Registerkarte "Anmerkungen" für Knoten und andere von Ihnen erstellte Objekte angezeigt werden.
- Umschalten zwischen traditionellem Modus und Analytic Server-Modus.

Um streamspezifische Optionen, wie Dezimaltrennzeichen, Uhrzeit- und Datumsformate, Optimierung, Stream-Layout sowie Stream-Scripts, festzulegen, verwenden Sie das Dialogfeld "Streameigenschaften", das über die Menüs "Datei" und "Tools" aufgerufen werden kann.

Festlegen von Benachrichtigungsoptionen

Mithilfe der Registerkarte "Benachrichtigungen" im Dialogfeld "Benutzeroptionen" können Sie verschiedene Optionen bzgl. des Auftretens und des Typs von Warnungen und Bestätigungsfenstern in IBM SPSS Modeler festlegen. Sie können auch das Verhalten der Registerkarten "Ausgaben" und "Modelle" im Managerbereich festlegen, wenn eine neue Ausgabe und neue Modelle generiert werden.

Feedbackdialogfeld für Streamausführung anzeigen Wählen Sie diese Option aus, um ein Dialogfeld mit Fortschrittsanzeige einzublenden, wenn ein Stream seit drei Sekunden ausgeführt wird. Das Dialogfeld enthält auch Einzelheiten zu den Ausgabeobjekten, die vom Stream erstellt werden.

- **Dialogfeld nach Abschluss schließen** Das Dialogfeld wird standardmäßig am Ende der Streamausführung geschlossen. Inaktivieren Sie dieses Kontrollkästchen, wenn das Dialogfeld sichtbar bleiben soll, nachdem der Stream beendet ist.

Warnen, wenn eine Datei durch einen Knoten überschrieben wird Wählen Sie diese Option aus, falls eine Fehlermeldung angezeigt werden soll, wenn eine vorhandene Datei durch Knotenoperationen überschrieben wird.

Warnen, wenn eine Datenbanktabelle durch einen Knoten überschrieben wird Wählen Sie diese Option aus, falls eine Fehlermeldung angezeigt werden soll, wenn eine vorhandene Datenbanktabelle durch Knotenoperationen überschrieben wird.

Akustische Benachrichtigungen

Verwenden Sie die Liste, um festzulegen, ob Sie bei Ereignissen und Fehlern durch ein Tonsignal benachrichtigt werden möchten. Es stehen mehrere Klänge zur Verfügung. Mit der Wiedergabeschaltfläche (Lautsprecher) geben Sie einen ausgewählten Klang wieder. Mit der Schaltfläche mit den Auslassungspunkten (...) wählen Sie einen Klang aus.

Anmerkung: Die .wav-Dateien, mit denen in IBM SPSS Modeler Töne erstellt werden, sind im Verzeichnis /media/sounds Ihrer Installation gespeichert.

- **Alle Klänge stummschalten** Wählen Sie diese Option aus, um die akustische Benachrichtigung für alle Ereignisse zu inaktivieren.

Visuelle Benachrichtigungen

Die Optionen in dieser Gruppe werden verwendet, um das Verhalten der Manager-Registerkarten "Ausgaben" und "Modelle" oben rechts beim Generieren neuer Elemente festzulegen. Wählen Sie **Neues Modell** oder **Neue Ausgabe** in der Liste aus, um das Verhalten der entsprechenden Registerkarte festzulegen.

Die folgende Option ist für **Neues Modell** verfügbar:

Bisheriges Modell ersetzen Wenn diese Option ausgewählt ist (Standard), wird ein bestehendes Modell auf der Registerkarte "Modelle" und im Streamerstellungsbereich aus diesem Stream überschrieben. Wenn dieses Kontrollkästchen inaktiviert ist, wird das Modell den vorhandenen Modellen auf der Registerkarte und im Erstellungsbereich hinzugefügt. Beachten Sie, dass diese Einstellung durch die Modellersetzungseinstellung bei einer Modellverknüpfung überschrieben wird.

Die folgende Option ist für **Neue Ausgabe** verfügbar:

Warnung, wenn Ausgabe größer als: [n] Wählen Sie aus, ob eine Warnung angezeigt werden soll, wenn die Anzahl der Elemente auf der Registerkarte "Ausgaben" einen zuvor angegebenen Wert übersteigt. Der Standardwert ist 20; Sie können diesen Wert jedoch bei Bedarf ändern.

Die folgenden Optionen sind in allen Fällen verfügbar:

Registerkarte auswählen Wählen Sie aus, ob Sie zur Registerkarte "Ausgaben" oder "Modelle" wechseln möchten, wenn das entsprechende Objekt während der Streamausführung generiert wird.

- Wählen Sie **Immer** aus, um zur entsprechenden Registerkarte im Managerbereich zu wechseln.
- Wählen Sie **Wenn durch aktuellen Stream generiert** aus, um nur bei Objekten, die von dem Stream generiert wurden, der derzeit im Erstellungsbereich sichtbar ist, zur entsprechenden Registerkarte zu wechseln.
- Wählen Sie **Nie** aus, um die Software so zu beschränken, dass nicht zu der entsprechenden Registerkarte gewechselt wird, um Sie über generierte Ausgaben oder Modelle zu benachrichtigen.

Registerkarte blinken lassen Wählen Sie aus, ob die Registerkarte "Ausgaben" oder "Modelle" im Managerbereich blinken soll, wenn neue Ausgaben bzw. Modelle generiert wurden.

- Wählen Sie **Wenn nicht ausgewählt** aus, damit die entsprechende Registerkarte blinkt (falls nicht bereits ausgewählt), sobald neue Objekte im Managerbereich generiert wurden.
- Wählen Sie **Nie** aus, um die Software so zu beschränken, dass die entsprechende Registerkarte nicht blinkt, um Sie über generierte Objekte zu benachrichtigen.

Sichtbarmachen durch Scrollen in Palette (nur Neues Modell). Legen Sie fest, ob Sie durch die Registerkarte "Modelle" im Managerbereich automatisch blättern möchten, damit das aktuellste Modell sichtbar wird.

- Wählen Sie **Immer** aus, um das Scrollen zu aktivieren.
- Wählen Sie **Wenn durch aktuellen Stream generiert** aus, um nur nach Objekten zu suchen, die von dem Stream generiert wurden, der derzeit im Erstellungsbereich sichtbar ist.
- Wählen Sie **Nie** aus, um die Software so zu beschränken, dass durch die Registerkarte "Modelle" nicht automatisch geblättert wird.

Fenster öffnen (nur Neue Ausgabe). Legen Sie fest, ob nach der Generierung automatisch ein Ausgabefenster geöffnet werden soll.

- Wählen Sie **Immer** aus, um immer ein neues Ausgabefenster zu öffnen.
- Wählen Sie **Wenn durch aktuellen Stream generiert** aus, um ein neues Fenster für die Ausgabe zu öffnen, die von dem Stream generiert wurde, der derzeit im Erstellungsbereich sichtbar ist.
- Wählen Sie **Nie** aus, um die Software so zu beschränken, dass keine neuen Fenster für die generierte Ausgabe automatisch geöffnet werden.

Klicken Sie auf **Standardwerte**, um auf die Systemstandardwerte für diese Registerkarte zurückzusetzen.

Festlegen von Anzeigoptionen

Auf der Registerkarte "Anzeige" im Dialogfeld "Benutzeroptionen" können Sie die Optionen für die Anzeige von Schriftarten und Farben in IBM SPSS Modeler festlegen.

Begrüßungsdialogfeld beim Start anzeigen. Wählen Sie diese Option aus, um das Begrüßungsdialogfeld beim Start anzuzeigen. Das Begrüßungsdialogfeld enthält Optionen, mit denen das Lernprogramm mit Anwendungsbeispielen gestartet, ein Demostream bzw. ein bestehender Stream oder ein bestehendes Projekt geöffnet oder ein neuer Stream erstellt werden kann.

Stream und Superknotenmarkups anzeigen. Wählen Sie diese Option aus, damit vorhandene Markups zu Streams und Superknoten standardmäßig angezeigt werden. Markup umfasst Streamkommentare, Modellverknüpfungen und die Hervorhebung von Scoring-Verzweigungen.

Standardschriftarten und -farben (Umstellung beim Neustart). Die Optionen in diesem Kontrollfeld werden verwendet, um das Bildschirmdesign und das Farbschema von IBM SPSS Modeler sowie die Größe der angezeigten Schriftarten anzugeben. Die hier ausgewählten Optionen werden erst wirksam, wenn Sie IBM SPSS Modeler schließen und neu starten.

- **Erscheinungsbild.** Wählen Sie ein Standardfarbschema und ein Bildschirmdesign aus. Sie haben folgende Möglichkeiten:
 - **SPSS Standard**, das Standarddesign.
 - **SPSS Classic**, ein Design, mit dem Benutzer früherer SPSS Modeler-Versionen vertraut sind.
 - **Windows**, ein Windows-Design, das u. U. hilfreich für einen höheren Kontrast im Streamerstellungsbereich und in den Paletten ist.
 - **Analytics Carbon**, ein modernes Design mit glänzenden Symbolen und Farben.
- **Standardschriftgröße für Knoten.** Geben Sie eine Schriftgröße an, die in den Knotenpaletten verwendet werden soll, sowie für Knoten, die im Streamerstellungsbereich angezeigt werden.
- **Schriftart fester Breite angeben.** Wählen Sie dieses Kontrollkästchen aus, wenn Sie eine Schriftart fester Breite und die zugehörige **Größe** der Schriftart zur Verwendung beim Scripting und bei Steuerungen von CLEM-Ausdrücken auswählen wollen. Die Standardschriftart ist `Monospace plain`. Klicken Sie auf **Ändern...**, wenn Sie eine Liste anderer Schriftarten anzeigen wollen, die Sie auswählen können.

Anmerkung: Sie können die Größe der Knotensymbole für einen Stream im Fenster "Layout" der Registerkarte "Optionen" des Dialogfelds mit den Streameigenschaften festlegen. Wählen Sie im Hauptmenü **Tools > Streameigenschaften > Optionen > Layout** aus.

Benutzerdefinierte Farben. In dieser Tabelle werden die aktuell ausgewählten Farben aufgelistet, die für verschiedene Anzeigeelemente verwendet werden. Sie können für jedes in der Tabelle aufgeführte Element die aktuelle Farbe ändern, indem Sie in die entsprechende Zeile in der Spalte **Farbe** doppelklicken und eine Farbe aus der Liste auswählen. Um eine benutzerdefinierte Farbe festzulegen, blättern Sie in der Liste nach unten und klicken Sie auf den Eintrag **Farbe...**

Reihenfolge für Diagrammkategorienfarbe. In dieser Tabelle werden die aktuell ausgewählten Farben aufgelistet, die in den neu erstellten Grafiken verwendet werden. Die Reihenfolge der Farben spiegelt die Reihenfolge wider, in der sie im Diagramm verwendet werden. Wenn z. B. ein als eine Farbüberlagerung verwendetes nominales Feld vier eindeutige Werte enthält, werden lediglich die ersten vier der hier aufgelisteten Farben verwendet. Sie können für jedes in der Tabelle aufgeführte Element die aktuelle Farbe ändern, indem Sie in die entsprechende Zeile in der Spalte **Farbe** doppelklicken und eine Farbe aus der Liste auswählen. Um eine benutzerdefinierte Farbe festzulegen, blättern Sie in der Liste nach unten und klicken Sie auf den Eintrag **Farbe...** Hier vorgenommene Änderungen wirken sich nicht auf zuvor erstellte Grafiken aus.

Klicken Sie auf **Standardwerte**, um auf die Systemstandardwerte für diese Registerkarte zurückzusetzen.

Festlegen von Optionen für die Syntaxanzeige

Auf der Registerkarte "Syntax" im Dialogfeld "Benutzeroptionen" können Sie die Optionen für die Schriftartattribute und Anzeigefarben in Scripts festlegen, die Sie in IBM SPSS Modeler erstellen.

Syntaxhervorhebung. In dieser Tabelle werden die aktuell ausgewählten Farben aufgelistet, die für verschiedene Syntaxelemente verwendet werden, einschließlich der Schriftart und des Fensters, in dem sie angezeigt wird. Sie können für jedes in der Tabelle aufgeführte Element die Farbe ändern, indem Sie auf die entsprechende Dropdown-Liste in der Zeile klicken und eine Farbe aus der Liste auswählen. Darüber hinaus können Sie für Schriftartelemente eine Hervorhebung durch Fettdruck oder Kursivdruck auswählen.

Vorschau. Diese Tabelle enthält ein Beispiel einer Syntaxanzeige, die die Farben und Schriftartattribute verwendet, die Sie in der Tabelle **Syntaxhervorhebung** auswählen. Diese Vorschau wird aktualisiert, sobald Sie eine Auswahl ändern.

Klicken Sie auf **Standardwerte**, um auf die Systemstandardwerte für diese Registerkarte zurückzusetzen.

Festlegen von PMML-Exportoptionen

Auf der Registerkarte "PMML" können Sie steuern, wie IBM SPSS Modeler Modelle in die Predictive Model Markup Language (PMML) exportiert. Weitere Informationen finden Sie in „Importieren und Exportieren von Modellen als PMML“ auf Seite 212.

PMML exportieren. Hier können Sie Variationen von PMML konfigurieren, die am besten mit Ihrer Zielanwendung funktionieren.

- Wählen Sie **mit Erweiterungen** aus, um PMML-Erweiterungen für Sonderfälle zuzulassen, für die es keine Standard-PMML-Entsprechung gibt. Beachten Sie, dass dies in den meisten Fällen zu demselben Ergebnis führt wie Standard-PMML.
- Wählen Sie **Als Standard-PMML...** aus, um PMML zu exportieren, die so nah wie möglich am PMML-Standard bleibt.

PMML-Standardoptionen. Wenn oben **Als Standard-PMML...** ausgewählt ist, können Sie eine von zwei gültigen Möglichkeiten auswählen, um lineare und logistische Regressionsmodelle zu exportieren:

- Als **PMML <GeneralRegression>-Modelle**
- Als **PMML <Regression>-Modelle**

Weitere Informationen zu PMML finden Sie auf der Website der Data Mining Group unter <http://www.dmg.org>.

Festlegen von Benutzerinformationen

Benutzer-/Autoreninformationen. Hier eingegebene Informationen können auf der Registerkarte "Anmerkungen" für Knoten und andere von Ihnen erstellte Objekte angezeigt werden.

Festlegen des Modus

Modeler-Moduseinstellungen. Auf der Registerkarte **Modus** können Sie unter den folgenden Modi wählen:

- **Traditioneller SPSS Modeler-Modus** zeigt alle verfügbaren Knoten und Ausdrücke in der Benutzerschnittstelle an.
- **Analytic Server-Modus** zeigt nur die Knoten und Ausdrücke an, die von Analytic Server unterstützt werden. Beachten Sie jedoch, dass einige Knoten und CLEM-Ausdrücke angezeigt werden, auch wenn sie nicht *vollständig* durch Analytic Server unterstützt werden. Die folgende Tabelle stellt allgemeine Informationen bereit, welche Knoten durch Analytic Server unterstützt, teilweise unterstützt und nicht unterstützt werden.

Weitere Details finden Sie in [Unterstützte Knoten](#).

Weitere Informationen zu Analytic Server finden Sie in der [Analytic Server-Dokumentation](#).

Beachten Sie, dass ein Moduswechsel sich nicht auf die Datenbankknoten der Datenbankmodellierungspalette auswirkt, wenn Sie SPSS Modeler für die Anzeige von Datenbankknoten in dieser Palette konfigurieren. Die Datenbankknoten werden immer angezeigt. Wenn Sie IBM Db2 for z/OS- oder IBM Netezza-Integration verwenden, werden die Knoten in einigen Fällen nach dem Wechsel in den Analytic Server-Modus nicht mehr in der Datenbankmodellierungspalette angezeigt. Ist dies der Fall, rufen Sie **Tools > Optionen > Hilfsanwendungen** auf und setzen Sie die Kontrollkästchen zurück.

Tabelle 38. Knotenunterstützung

Knotentyp (Palettenname)	Unterstützt durch Analytic Server	Teilweise unterstützt durch Analytic Server	Nicht unterstützt durch Analytic Server
Quellen	<ul style="list-style-type: none"> • Analytic Server-Quellenknoten 		<ul style="list-style-type: none"> • Datenbank • Variable Datei • Datei (fest) • Statistikdatei • Data Collection • IBM Cognos • TWC-Import • TM1-Importknoten • SAS-Dateiknoten • Excel • XML-Benutzereingabe • simgen • Datenansicht • Georäumlich • Objektspeicher • Erweiterungsimport • R-Import • SNA (Diffusionsanalyse und Gruppenanalyse)
Datensatzoperationen	<ul style="list-style-type: none"> • Auswählen • Sortieren • Balance • Duplikat • RFM-Aggregat • Anhängen • Streaming-ZR • Erweiterungstransformation • Streaming-TCM 	<ul style="list-style-type: none"> • Stichprobe (unterstützt nur zufälligen Ziehungen (%) für die Methode "Einfach". "Komplex" wird nicht unterstützt.) • Zusammenführen (unterstützt nur Join nach Schlüssel und Bedingungen) • Aggregieren (1. Quantil, 3. Quantil und Median werden durch Analytic Server nicht unterstützt) 	<ul style="list-style-type: none"> • Space-Time-Boxes • CPLEX-Optimierung

Tabelle 38. Knotenunterstützung (Forts.)

Knotentyp (Palettenname)	Unterstützt durch Analytic Server	Teilweise unterstützt durch Analytic Server	Nicht unterstützt durch Analytic Server
Feldoperationen	<ul style="list-style-type: none"> • Typ • Filter • Ableitung • Füller • Umcodierung • Ensemble • Dichotom • Umstrukturierung • Felder umordnen • Reprojizieren • Zeitintervalle 	<ul style="list-style-type: none"> • Automatische Datenaufbereitung (unterstützt nur Transformieren) • Klassierung (Klassierungsmethode equalFreq wird nicht unterstützt, wenn Bindungseinstellung In aktuellem beibehalten ausgewählt ist) • RFM-Analyse (Klassierungsmethode Tiles wird nicht unterstützt, wenn Bindungseinstellung In aktuellem beibehalten ausgewählt ist) • Partition (nicht durch Analytic Server unterstützt, sofern nicht ein eindeutiges Feld verwendet wird, um wiederholt Zeilen zu Partitionen zuzuweisen) 	<ul style="list-style-type: none"> • Anonymisierung • Verlauf • Transponieren
Grafiken	<ul style="list-style-type: none"> • Diagramm • Multiplot • Zeitdiagramm • Verteilung • Histogramm • Sammlung • Internet • Evaluierung • Kartenvisualisierung • E-Plot (Beta) • t-SNE 	<ul style="list-style-type: none"> • Diagrammtafel (unterstützt nur die Aggregationsmodusfunktion für Felder mit dem Messniveau "kategorial", "nominal", "ordinal" oder "Flags") 	

Tabelle 38. Knotenunterstützung (Forts.)

Knotentyp (Palettenname)	Unterstützt durch Analytic Server	Teilweise unterstützt durch Analytic Server	Nicht unterstützt durch Analytic Server
Modellierung	<ul style="list-style-type: none"> • Zeitreihen • TCM • Isotonisch-AS • Random Trees • Tree-AS • Linear-AS • GLE • LSVM • STP • TwoStep-AS • Assoziationsregeln • XGBoost-AS • K-Means-AS 	<ul style="list-style-type: none"> • Automatisches Klassifikationsmerkmal • Autonumerisch (die beiden Knoten unterstützten nur "Aufteilung" und ein Feld mit der Rolle "Aufteilung" muss bereitgestellt werden, wenn die Option Auf Analytic Server ausführen (Aufteilungen aktiviert) des automatischen Klassifikationsmerkmals verwendet wird) • Erweiterung (das R-Syntax-Erstellungsmodell wird nicht durch Analytic Server unterstützt) • Die folgenden Knoten unterstützen nur Aufteilung und PSM: <ul style="list-style-type: none"> – C&R-Baum – Linear – Neuronales Netz – CHAID – Quest 	<ul style="list-style-type: none"> • Autom. Cluster • Entscheidungsliste • C5.0 • Regression • Faktor/PCA • Merkmalauswahl • Diskriminanz • Logistisch • GenLin • GLMM • Bayes-Netz • Apriori • Carma • Sequenz • K-Means • Kohonen • TwoStep • Anomalie • KNN • R • Random Forest • Die folgenden Knoten verwenden ASL, jedoch nur Schreib-Lese-ASL: <ul style="list-style-type: none"> – Cox – SVM – SLRM
Ausgabe	<ul style="list-style-type: none"> • Tabelle • Matrix • Analyse • Data Audit • Transformieren • Statistik • Mittelwert • Bericht • Globalwerte 	<ul style="list-style-type: none"> • Erweiterungsausgabe • R-Syntax-Ausgabe 	<ul style="list-style-type: none"> • Simulationsevaluierung • Simulationsanpassung • R-Ausgabe

Tabelle 38. Knotenunterstützung (Forts.)

Knotentyp (Palettenname)	Unterstützt durch Analytic Server	Teilweise unterstützt durch Analytic Server	Nicht unterstützt durch Analytic Server
Export	<ul style="list-style-type: none"> Analytic Server-Exportknoten 	<ul style="list-style-type: none"> Erweiterungsexport R-Syntax-Export 	<ul style="list-style-type: none"> Datenbank Flatfile Statistikexport Data Collection Excel IBM Cognos-Export TM1-Export SAS XML-Export Objektspeicher R-Export
IBM SPSS Statistics			<ul style="list-style-type: none"> Statistikdatei Statistics-Transformation Statistics-Modell Statistikausgabe Statistikexport
IBM SPSS Text Analytics	<ul style="list-style-type: none"> Textlinkanalyse Textmining Sprachknoten 		<ul style="list-style-type: none"> Dateiliste Web-Feed Textlinkanalyse Übersetzung Textmining Dateianzeigefunktion
Python			<ul style="list-style-type: none"> SMOTE One-Class SVM XGBoost Tree XGBoost Linear t-SNE Random Forest HDBSCAN
Spark	Alle unterstützt		

Anpassen der Knotenpalette

Streams werden mithilfe von Knoten erstellt. Die Knotenpalette unten im IBM SPSS Modeler-Fenster enthält alle Knoten, die bei der Streamerstellung verwendet werden können. Weitere Informationen finden Sie in „Knotenpalette“ auf Seite 15.

Sie haben zwei Möglichkeiten für die Neuordnung der Knotenpalette:

- Anpassen des Palettenmanagers. Weitere Informationen finden Sie in „Anpassen des Palettenmanagers“ auf Seite 233.

- Ändern der Art und Weise, wie Palettenregisterkarten, die Unterpaletten enthalten, auf der Knotenpalette angezeigt werden. Weitere Informationen finden Sie in „Erstellen einer Unterpalette“ auf Seite 235.

Anpassen des Palettenmanagers

Der Palettenmanager kann entsprechend Ihrer Nutzung von IBM SPSS Modeler angepasst werden. Wenn Sie z. B. häufig Zeitreihendaten aus einer Datenbank analysieren, möchten Sie u. U. sicherstellen, dass der Datenbankquellenknoten, der Zeitintervallknoten, der Zeitreihenknoten und der Zeitdiagrammknoten gemeinsam auf einer speziellen Registerkarte der Palette zur Verfügung stehen. Im Palettenmanager können Sie diese Anpassungen problemlos vornehmen, indem Sie auf der Knotenpalette benutzerdefinierte Palettenregisterkarten erstellen.

Mit dem Palettenmanager können Sie verschiedene Aufgaben durchführen:

- Festlegen, welche Palettenregisterkarten auf der Knotenpalette unterhalb des Streamerstellungsbereichs angezeigt werden sollen.
- Ändern der Reihenfolge, in der die Palettenregisterkarten auf der Knotenpalette angezeigt werden.
- Erstellen und Bearbeiten eigener Palettenregisterkarten und gegebenenfalls zugehöriger Unterpaletten.
- Bearbeiten der Standardknotenauswahl auf Ihren Registerkarten.

Um auf den Palettenmanager zuzugreifen, klicken Sie im Menü "Tools" auf **Paletten verwalten**.

Palettenname. Jede verfügbare Palettenregisterkarte wird aufgeführt, unabhängig davon, ob sie auf der Knotenpalette angezeigt wird oder nicht. Dies umfasst gegebenenfalls auch alle von Ihnen selbst erstellten Palettenregisterkarten. Weitere Informationen finden Sie in „Erstellen einer Palettenregisterkarte“.

Anzahl der Knoten. Die Anzahl der auf den einzelnen Palettenregisterkarten angezeigten Knoten. Ein hoher Wert bedeutet, dass es praktischer sein kann, Unterpaletten zu erstellen, um die Knoten auf der Registerkarte zu unterteilen. Weitere Informationen finden Sie in „Erstellen einer Unterpalette“ auf Seite 235.

Angezeigt? Wählen Sie dieses Feld aus, um die Palettenregisterkarte auf der Knotenpalette anzuzeigen. Weitere Informationen finden Sie in „Anzeigen von Palettenregisterkarten auf der Knotenpalette“ auf Seite 234.

Unterpaletten. Um Unterpaletten zur Anzeige auf einer Palettenregisterkarte auszuwählen, markieren Sie den entsprechenden **Palettennamen** und klicken Sie auf diese Schaltfläche, um die Registerkarte "Unterpaletten" anzuzeigen. Weitere Informationen finden Sie in „Erstellen einer Unterpalette“ auf Seite 235.

Standardeinstellungen wiederherstellen. Durch Klicken auf diese Schaltfläche können Sie alle Änderungen und Zusätze, die Sie an den Paletten und Unterpaletten vorgenommen haben, vollständig entfernen und zu den Standardeinstellungen für die Palette zurückkehren.

Erstellen einer Palettenregisterkarte

So erstellen Sie eine benutzerdefinierte Palettenregisterkarte:

1. Öffnen Sie im Menü "Tools" die Option "Palettenmanager".
2. Klicken Sie rechts neben der Spalte *Angezeigt?* auf die Schaltfläche zum Hinzufügen von Paletten; das Dialogfeld "Palette erstellen/bearbeiten" wird angezeigt.
3. Geben Sie einen eindeutigen **Palettennamen** ein.
4. Wählen Sie im Bereich **Verfügbare Knoten** den Knoten aus, der der Palettenregisterkarte hinzugefügt werden soll.
5. Klicken Sie unter "Knoten hinzufügen" auf die Schaltfläche mit Rechtspfeil, um den hervorgehobenen Knoten zum Bereich **Ausgewählte Knoten** hinzuzufügen. Wiederholen Sie den Vorgang, bis Sie alle gewünschten Knoten hinzugefügt haben.

Nachdem alle erforderlichen Knoten hinzugefügt wurden, können Sie die Reihenfolge ändern, in der sie auf der Palettenregisterkarte angezeigt werden:

- Mithilfe der einfachen Pfeilschaltflächen können Sie einen Knoten eine Zeile nach oben bzw. unten verschieben.
- Mithilfe der Schaltflächen mit Pfeil und Balken können Sie einen Knoten ganz nach unten oder oben in der Liste verschieben.
- Um einen Knoten aus einer Palette zu entfernen, markieren Sie den Knoten und klicken Sie auf die Schaltfläche "Löschen" rechts neben dem Bereich **Ausgewählte Knoten**.

Anzeigen von Palettenregisterkarten auf der Knotenpalette

Es kann in IBM SPSS Modeler Optionen geben, die Sie nie verwenden. In diesem Fall können Sie die Registerkarten mit den entsprechenden Knoten mithilfe des Palettenmanagers ausblenden.

So können Sie auswählen, welche Registerkarten auf der Knotenpalette angezeigt werden sollen:

- Öffnen Sie im Menü "Tools" die Option "Palettenmanager".
- Mithilfe der Kontrollkästchen in der Spalte *Angezeigt?* können Sie auswählen, ob die einzelnen Palettenregisterkarten ein- oder ausgeblendet werden sollen.

Wenn Sie eine Palettenregisterkarte dauerhaft aus der Knotenpalette entfernen möchten, müssen Sie den Knoten markieren und rechts neben der Spalte *Angezeigt?* auf die Schaltfläche "Löschen" klicken. Gelöschte Palettenregisterkarten können nicht wiederhergestellt werden.

Hinweis: Mit Ausnahme der Registerkarte "Favoriten" können die im Lieferumfang von IBM SPSS Modeler enthaltenen Standardpalettenregisterkarten nicht gelöscht werden.

Ändern der Anzeigereihenfolge auf der Knotenpalette

Nachdem Sie ausgewählt haben, welche Palettenregisterkarten angezeigt werden sollen, können Sie die Reihenfolge ändern, in der sie auf der Knotenpalette angezeigt werden:

- Mithilfe der einfachen Pfeilschaltflächen können Sie eine Palettenregisterkarte eine Zeile nach oben bzw. unten verschieben. Beim Verschieben nach oben werden sie links neben die Knotenpalette verschoben und umgekehrt.
- Mithilfe der Schaltflächen mit Pfeil und Balken können Sie eine Palettenregisterkarte ganz nach unten oder oben in der Liste verschieben. Die Elemente oben in der Liste werden links neben der Knotenpalette angezeigt.

Anzeigen von Unterpaletten auf einer Palettenregisterkarte

So wie Sie festlegen können, welche Palettenregisterkarten auf der Knotenpalette angezeigt werden sollen, so können Sie auch festlegen, welche Unterpaletten auf der zugehörigen übergeordneten Palettenregisterkarte verfügbar sein sollen.

So wählen Sie die Unterpaletten aus, die auf einer Palettenregisterkarte angezeigt werden sollen:

- Öffnen Sie im Menü "Tools" die Option "Palettenmanager".
- Wählen Sie die erforderliche Palette aus.
- Klicken Sie auf die Schaltfläche "Unterpaletten". Das Dialogfeld "Unterpaletten" wird angezeigt.
- Mithilfe der Kontrollkästchen in der Spalte *Angezeigt?* können Sie auswählen, ob die einzelnen Unterpaletten auf der Palettenregisterkarte erscheinen sollen. Die Unterpalette **Alle** wird immer angezeigt und kann nicht gelöscht werden.
- Wenn Sie eine Unterpalette dauerhaft von der Registerkarte "Palette" entfernen möchten, müssen Sie die Unterpalette markieren und rechts neben der Spalte *Angezeigt?* auf die Schaltfläche "Löschen" klicken.

Hinweis: Die im Lieferumfang der Palettenregisterkarte "Modellierung" enthaltenen Standardunterpaletten können nicht gelöscht werden.

Ändern der Anzeigereihenfolge auf der Palettenregisterkarte

Nachdem Sie ausgewählt haben, welche Unterpaletten angezeigt werden sollen, können Sie die Reihenfolge ändern, in der sie auf der übergeordneten Palettenregisterkarte angezeigt werden:

1. Mithilfe der einfachen Pfeilschaltflächen können Sie eine Unterpalette eine Zeile nach oben bzw. unten verschieben.
2. Mithilfe der Schaltflächen mit Pfeil und Balken können Sie eine Unterpalette ganz nach unten oder oben in der Liste verschieben.

Die von Ihnen erstellten Unterpaletten werden auf der Knotenpalette angezeigt, wenn Sie die zugehörige übergeordnete Palettenregisterkarte auswählen. Weitere Informationen finden Sie in „Ändern der Ansicht einer Palettenregisterkarte“.

Erstellen einer Unterpalette

Da Sie in die von Ihnen erstellten benutzerdefinierten Palettenregisterkarten jeden beliebigen bestehenden Knoten aufnehmen können, ist es möglich, dass Sie mehr Knoten auswählen, als ohne Bildlauf auf dem Bildschirm angezeigt werden können. Um die Notwendigkeit eines Bildlaufs zu vermeiden, können Sie Unterpaletten erstellen, in die Sie die Knoten aufnehmen können, die Sie für die Palettenregisterkarte ausgewählt haben. Wenn Sie beispielsweise eine Palettenregisterkarte erstellt haben, die die Knoten enthält, die Sie am häufigsten zum Erstellen von Streams verwenden, können Sie vier Unterpaletten erstellen, die die Auswahl nach Quellenknoten, Feldoperationen, Modellierung und Ausgabe aufgliedern.

Hinweis: Als Knoten für die Unterpaletten stehen nur diejenigen Knoten zur Auswahl, die zur übergeordneten Palettenregisterkarte hinzugefügt wurden.

So erstellen Sie eine Unterpalette:

1. Öffnen Sie im Menü "Tools" die Option "Palettenmanager".
2. Wählen Sie die Palette aus, zu der Sie Unterpaletten hinzufügen möchten.
3. Klicken Sie auf die Schaltfläche "Unterpaletten". Das Dialogfeld "Unterpaletten" wird angezeigt.
4. Klicken Sie rechts neben der Spalte *Angezeigt?* auf die Schaltfläche zum Hinzufügen von Unterpaletten. Das Dialogfeld "Unterpalette anzeigen/bearbeiten" wird angezeigt.
5. Geben Sie einen eindeutigen **Unterpalettennamen** ein.
6. Wählen Sie im Bereich **Verfügbare Knoten** den Knoten aus, der zur Unterpalette hinzugefügt werden soll.
7. Klicken Sie unter "Knoten hinzufügen" auf den nach rechts weisenden Pfeil, um den ausgewählten Knoten zum Bereich **Ausgewählte Knoten** hinzuzufügen.
8. Klicken Sie nach dem Hinzufügen der erforderlichen Knoten auf **OK**, um zum Dialogfeld "Unterpaletten" zurückzukehren.

Die von Ihnen erstellten Unterpaletten werden auf der Knotenpalette angezeigt, wenn Sie die zugehörige übergeordnete Palettenregisterkarte auswählen. Weitere Informationen finden Sie in „Ändern der Ansicht einer Palettenregisterkarte“.

Ändern der Ansicht einer Palettenregisterkarte

Aufgrund der großen Anzahl der Knoten, die in IBM SPSS Modeler verfügbar sind, sind möglicherweise auf kleinen Bildschirmen nicht alle Knoten sichtbar, ohne dass auf der Knotenpalette ein Bildlauf nach links oder rechts durchgeführt wird. Dieses Problem tritt verstärkt auf der Palettenregisterkarte "Modellierung" auf. Um die Notwendigkeit eines Bildlaufs zu verringern, können Sie auswählen, dass nur die auf einer Unterpalette (sofern verfügbar) enthaltenen Knoten angezeigt werden sollen. Weitere Informationen finden Sie in „Erstellen einer Unterpalette“.

Um die auf einer Palettenregisterkarte angezeigten Knoten zu ändern, wählen Sie die Palettenregisterkarte aus und legen dann im Menü auf der linken Seite fest, ob alle Knoten angezeigt werden sollen oder nur diejenigen, die sich auf einer bestimmten Unterpalette befinden.

Kapitel 15. Leistungsaspekte für Streams und Knoten

Sie können Ihre Streams so entwerfen, dass die Leistungsfähigkeit maximiert wird, indem Sie die Knoten in der effizientesten Konfiguration anordnen, an den entsprechenden Stellen Knotencaches aktivieren und andere Punkte beachten, die in diesem Abschnitt ausgeführt sind.

Neben den hier erörterten Punkten können weitere und deutlichere Leistungsverbesserungen normalerweise durch die effektive Nutzung der Datenbank, insbesondere durch SQL-Optimierung, erzielt werden.

Knotenreihenfolge

Selbst wenn Sie keine SQL-Optimierung verwenden, kann sich die Reihenfolge der Knoten in einem Stream auf die Leistung auswirken. Das allgemeine Ziel besteht in der Minimierung des Verarbeitungsaufwands weiter unten im Stream. Daher sollten Knoten, die die Datenmenge verringern, sofern vorhanden, relativ an den Anfang des Streams gestellt werden. IBM SPSS Modeler Server kann während der Kompilierung automatisch einige Neuordnungsregeln anwenden, um bestimmte Knoten an eine Stelle weiter oben im Stream zu versetzen, wenn nachgewiesen werden kann, dass dieser Vorgang keine Probleme mit sich bringt. (Diese Funktion ist standardmäßig aktiviert. Fragen Sie bei Ihrem Systemadministrator nach, ob diese Funktion in Ihrer Installation tatsächlich aktiviert ist.)

Durch die Verwendung der SQL-Optimierung sollen Verfügbarkeit und Effizienz maximiert werden. Da die Optimierung anhält, wenn ein Stream eine Operation enthält, die nicht in der Datenbank durchgeführt werden kann, sollten SQL-optimierte Operationen zusammen am Anfang des Streams gruppiert werden. Diese Strategie sorgt dafür, dass ein größerer Teil der Verarbeitung in der Datenbank erfolgt und somit weniger Daten in IBM SPSS Modeler eingebracht werden.

Die folgenden Operationen können in den meisten Datenbanken ausgeführt werden. Versuchen Sie, sie an den *Anfang* des Streams zu stellen:

- Zusammenführen (Merge) nach Schlüssel (join)
- Auswählen
- Aggregieren
- Sortieren
- Stichprobe
- Anhängen
- Verschiedene Vorgänge im Modus *Einschließen*, in dem alle Felder ausgewählt werden
- Füller-Operationen
- Grundlegende Ableitungsoperationen unter Verwendung von Standardarithmetik oder Zeichenfolgenrebearbeitung (je nachdem, welche Operationen von der Datenbank unterstützt werden)
- Dichotomie

Die folgenden Operationen können in den meisten Datenbanken nicht ausgeführt werden. Sie sollten im Stream *hinter* den Operationen in der vorstehenden Liste platziert werden:

- Operationen für andere Daten als Datenbankdaten, beispielsweise für Flatfiles
- Zusammenführen (Merge) nach Reihenfolge
- Balance
- Bestimmte Operationen im Modus *Verwerfen* oder wenn nur ein Subset von Feldern als unterschiedlich ausgewählt wurde
- Alle Operationen, bei denen auf Daten zugegriffen werden muss, die sich in einem anderen Datensatz als dem gerade verarbeiteten befinden

- Ableitungen von Status- und Anzahlfeldern
- Verlaufsknotenoperationen
- Operationen, die Zeitreihenfunktionen (mit "@") beinhalten
- Die Typprüfungsmodi *Warnen* und *Abbrechen*
- Modellerstellung, -anwendung und -analyse

Hinweis: Entscheidungsbäume, Regelsets, lineare Regression und faktorgenerierte Modelle können SQL generieren und daher per Pushback an die Datenbank zurückübertragen werden.

- Datenausgabe an einen anderen Ort als die Datenbank, die die Daten verarbeitet

Knotencaches

Um die Streamausführung zu optimieren, können Sie einen *Cache* für Knoten einrichten, die keine Endknoten sind. Wenn Sie einen Cache für einen Knoten einrichten, wird der Cache mit den Daten gefüllt, die bei der nächsten Ausführung des Datenstreams durch den Knoten laufen. Ab diesem Zeitpunkt werden die Daten aus dem Cache (der sich in einem temporären Verzeichnis auf dem Datenträger befindet) und nicht mehr aus der Datenquelle eingelesen.

Caching ist am nützlichsten im Anschluss an zeitaufwendige Operationen, wie Sortieren, Zusammenführen oder Aggregation. Angenommen, es ist beispielsweise ein Quellenknoten vorhanden, der zum Lesen von Umsatzdaten aus einer Datenbank eingerichtet ist, und ein Aggregatknoten, der den Umsatz nach Standort zusammenfasst. Sie können einen Cache an dem Aggregatknoten einrichten anstatt an dem Quellenknoten, weil Sie möchten, dass der Cache die aggregierten Daten und nicht das gesamte Dataset speichert.

Anmerkung: Ein Caching bei Quellenknoten, bei denen einfach eine Kopie der ursprünglichen Dateien, die in IBM SPSS Modeler eingelesen wurden, gespeichert wird, bewirkt meist keine Leistungsverbesserung.

Knoten mit aktiviertem Caching werden mit einem kleinen Dokumentsymbol in der oberen rechten Ecke angezeigt. Wenn die Daten am Knoten zwischengespeichert werden, ist das Dokumentsymbol grün.

So aktivieren Sie einen Cache:

1. Klicken Sie im Streamerstellungsbereich mit der rechten Maustaste auf den Knoten und klicken Sie im Menü auf die Option **Cache**.
2. Klicken Sie im Untermenü auf die Option **Aktivieren**.
3. Sie können den Cache inaktivieren, indem Sie mit der rechten Maustaste auf den Knoten klicken und im Untermenü auf **Inaktivieren** klicken.

Caching von Knoten in einer Datenbank

Bei in einer Datenbank ausgeführten Streams muss nicht das Dateisystem als Cache verwendet werden, sondern die Daten können mitten im Stream in einer temporären Tabelle in der Datenbank zwischengespeichert werden. Bei einer Kombination mit SQL-Optimierung kann dies zu einer erheblichen Leistungssteigerung führen. So kann beispielsweise die Ausgabe aus einem Stream, der mehrere Tabellen zusammenfasst, um eine Data-Mining-Ansicht zu erstellen, im Cache gespeichert und bei Bedarf wieder verwendet werden. Durch die automatische Generierung von SQL für alle nachgeordneten Knoten lässt sich die Leistung weiter erhöhen.

Um das Caching von Datenbanken nutzen zu können, müssen sowohl SQL-Optimierung als auch Datenbankcaching aktiviert sein. Die Optimierungseinstellungen auf dem Server haben Vorrang vor jenen auf dem Client. Weitere Informationen finden Sie in „Festlegen von Optimierungsoptionen für Streams“ auf Seite 47.

Klicken Sie bei aktiviertem Datenbankcaching einfach mit der rechten Maustaste auf einen Knoten (keinen Endknoten), um die Daten an dieser Stelle im Cache zu speichern, und der Cache wird bei der nächsten Ausführung des Streams automatisch unmittelbar in der Datenbank erstellt. Wenn Datenbankcaching oder SQL-Optimierung nicht aktiviert ist, wird der Cache stattdessen in das Dateisystem geschrieben.

Anmerkung: Die folgenden Datenbanken unterstützen temporäre Tabellen zu Caching-Zwecken: Db2, Oracle, SQL Server und Teradata. Andere Datenbanken, beispielsweise Netezza, verwenden für das Datenbankcaching eine normale Tabelle. Der SQL-Code kann für bestimmte Datenbanken angepasst werden. Wenden Sie sich an den Kundendienst, wenn Sie Unterstützung benötigen.

Leistung: Prozessknoten

Sortieren. Der Sortierknoten muss das gesamte Eingabedataset lesen, bevor es sortiert werden kann. Die Daten werden bis zu einer gewissen Obergrenze im Arbeitsspeicher gespeichert, der Rest wird auf den Datenträger geschrieben. Der Sortieralgorithmus ist ein Kombinationsalgorithmus: Daten werden bis zur angegebenen Grenze in den Arbeitsspeicher eingelesen und mithilfe eines schnellen hybriden Quick-Sort-Algorithmus sortiert. Wenn alle Daten in den Arbeitsspeicher passen, ist der Sortiervorgang abgeschlossen. Andernfalls wird ein Merge-Sort-Algorithmus angewendet. Die sortierten Daten werden in eine Datei geschrieben und die nächste Datenmenge wird in den Arbeitsspeicher eingelesen, sortiert und auf den Datenträger geschrieben. Dieser Vorgang wird wiederholt, bis alle Daten gelesen wurden; anschließend werden die sortierten Datenmengen zusammengeführt (Merge). Für die Zusammenführung können mehrere Durchläufe über die auf dem Datenträger gespeicherten Daten erforderlich sein. Bei Spitzennutzung hat der Sortierknoten zwei vollständige Kopien des Datasets auf dem Datenträger gespeichert, eine sortierte und eine unsortierte.

Die Gesamtlaufzeit des Algorithmus liegt in der Größenordnung von $N \cdot \log(N)$. Dabei ist N die Anzahl der Datensätze. Die Sortierung im Arbeitsspeicher ist schneller als die Zusammenführung über den Datenträger, sodass sich die tatsächliche Laufzeit durch die Zuordnung von mehr Arbeitsspeicher zum Sortiervorgang reduzieren lässt. Der Algorithmus ordnet sich selbst einen Teil des physischen RAM zu, der von der IBM SPSS Modeler Server-Konfigurationsoption *Multiplikator für Arbeitsspeicherverwendung* gesteuert wird. Um den für die Sortierung verwendeten Arbeitsspeicher zu erhöhen, müssen Sie mehr physischen RAM bereitstellen oder diesen Wert erhöhen. Beachten Sie: Wenn der Anteil des verwendeten Arbeitsspeichers die Arbeitsmenge des Prozesses überschreitet, sodass ein Teil des Arbeitsspeichers auf den Datenträger ausgelagert wird, lässt die Leistung nach, da das Muster für den Arbeitsspeicherzugriff des arbeitsspeicherinternen Sortieralgorithmus zufällig ist und zu übermäßigem Paging führen kann. Der Sortieralgorithmus wird außer vom Sortierknoten noch von mehreren anderen Knoten verwendet; es gelten jedoch dieselben Leistungsregeln.

Klassieren. Der Klassierknoten liest das gesamte Eingabedataset ein, um die Klassengrenzen zu berechnen, bevor er Datensätze zu Klassen zuordnet. Das Dataset wird während der Berechnung der Grenzen im Cache gespeichert; anschließend wird es für die Zuordnung erneut überprüft. Bei der Klassiermethode *Feste Breite* oder *Mittelwert/Standardabweichung* wird das Dataset direkt auf dem Datenträger im Cache gespeichert. Diese Methoden weisen eine lineare Laufzeit auf und erfordern so viel Speicherplatz auf dem Datenträger, dass das gesamte Dataset gespeichert werden kann. Bei der Klassiermethode *Ränge* oder *n-tile* wird das Dataset mithilfe des zuvor beschriebenen Sortieralgorithmus sortiert und das sortierte Dataset wird als Cache verwendet. Die Sortierung führt bei diesen Methoden zu einer Laufzeit von $M \cdot N \cdot \log(N)$. Dabei ist M die Anzahl der klassierten Felder und N ist die Anzahl der Datensätze. Auf dem Datenträger muss das Doppelte der Datasetgröße an Speicherplatz verfügbar sein.

Das Generieren eines Ableitungsknotens auf der Grundlage von generierten Klassen verbessert die Leistung in den nachfolgenden Durchgängen. Ableitungsoperationen sind wesentlich schneller als das Klassieren.

Zusammenführen (Merge) nach Schlüssel (join). Bei der Zusammenführungsmethode *Schlüssel* (entspricht einem Datenbankjoin) sortiert der Zusammenführungsknoten alle Eingabedatasets anhand der

Schlüsselfelder. Dieser Teil des Verfahrens weist eine Laufzeit von $M*N*\log(N)$ auf. Dabei ist M die Anzahl der Eingaben und N die Anzahl der Datensätze in der größten Eingabe: Es erfordert genügend Datenträgerspeicher, um alle Eingabedatasets sowie eine zweite Kopie des größten Datensets zu speichern. Die Laufzeit der Zusammenführung selbst ist proportional zur Größe des Ausgabedatasets, die von der Häufigkeit übereinstimmender Schlüssel abhängt. Im schlechtesten Fall, bei dem die Ausgabe das kartesische Produkt der Eingaben ist, kann sich die Laufzeit dem Wert NM nähern. Dies kommt jedoch nur selten vor. Die meisten Joins weisen deutlich weniger übereinstimmende Schlüssel auf. Wenn ein Dataset größer ist als die anderen oder wenn die eingehenden Daten bereits durch ein Schlüsselfeld sortiert wurden, können Sie mithilfe der Registerkarte "Optimierung" die Leistung dieses Knotens verbessern.

Aggregieren. Wenn die Option *Schlüssel sind zusammenhängend* nicht ausgewählt ist, liest dieser Knoten das gesamte Eingabedataset ein (speichert es jedoch nicht), bevor eine aggregierte Ausgabe erstellt wird. In extremeren Situationen, wenn die Größe der aggregierten Daten einen Grenzwert erreicht (von der Konfigurationsoption *Multiplikator für Arbeitsspeicherverwendung* von IBM SPSS Modeler Server festgelegt), wird der Rest des Datensets so sortiert und verarbeitet, als wäre die Option *Schlüssel sind zusammenhängend* ausgewählt worden. Wenn diese Option nicht ausgewählt wurde, werden keine Daten gespeichert, weil die aggregierten Ausgabedatensätze beim Lesen der Eingabedaten produziert werden.

Duplikat. Der Duplikatknoten speichert alle eindeutigen Schlüsselfelder im Eingabedataset. In Fällen, in denen alle Felder Schlüsselfelder und alle Datensätze eindeutig sind, wird das gesamte Dataset gespeichert. Standardmäßig sortiert der Duplikatknoten die Daten in den Schlüsselfeldern und wählt (oder verwirft) dann den ersten eindeutigen Datensatz aus jeder Gruppe. Für kleinere Datensets mit einer geringen Anzahl eindeutiger Schlüssel oder für diejenigen, die vorsortiert wurden, können Sie Optionen wählen, um Geschwindigkeit und Effizienz der Verarbeitung zu steigern.

Typ. In einigen Fällen speichert der Typknoten beim Lesen der Werte die Eingabedaten im Cache. Der Cache wird für die nachfolgende Verarbeitung verwendet. Für den Cache muss genügend Plattenspeicherplatz zum Speichern des gesamten Datensets verfügbar sein. Dafür wird jedoch die Verarbeitung beschleunigt.

Evaluierung. Der Evaluierungsknoten muss die Eingabedaten sortieren, um n -tile zu berechnen. Der Sortiervorgang wird für jedes evaluierte Modell wiederholt, da die Scores und folglich die Datensatzreihenfolge in jedem Fall unterschiedlich sind. Die Laufzeit beträgt $M*N*\log(N)$. Dabei ist M die Anzahl der Modelle und N die Anzahl der Datensätze.

Leistung: Modellierungsknoten

Neuronales Netz und Kohonen. Trainingsalgorithmen für neuronale Netze (einschließlich des Kohonen-Algorithmus) führen zahlreiche Durchläufe über die Trainingsdaten durch. Die Daten werden bis zu einer gewissen Obergrenze im Arbeitsspeicher gespeichert, der Rest wird auf den Datenträger geschrieben. Der Zugriff auf die Trainingsdaten über den Datenträger ist aufwendig, da es sich um eine Zufallszugriffsmethode handelt, die zu übermäßiger Datenträgeraktivität führen kann. Sie können die Verwendung des Datenträgerspeichers für diese Algorithmen inaktivieren, sodass alle Daten im Arbeitsspeicher gespeichert werden müssen. Wählen Sie dazu im Dialogfeld des Knotens auf der Registerkarte "Modell" die Option **Auf Geschwindigkeit optimieren** aus. Wenn der zum Speichern der Daten erforderliche Arbeitsspeicher größer ist als die Arbeitsmenge des Serverprozesses, wird ein Teil davon auf den Datenträger ausgelagert und die Leistung wird entsprechend beeinträchtigt.

Bei Auswahl von **Speicher optimiert** wird dem Algorithmus ein Prozentsatz des physischen RAM zugeordnet. Dies geschieht gemäß dem Wert in der Konfigurationsoption *Obergrenze für Modellierungsarbeitsspeicher (in Prozent)* von IBM SPSS Modeler Server. Um mehr Arbeitsspeicher für das Trainieren neuronaler Netze zu verwenden, stellen Sie entweder mehr RAM zur Verfügung oder erhöhen Sie den Wert für diese Option. Beachten Sie jedoch, dass ein zu hoher Wert zu Paging führt.

Die Laufzeit des Algorithmus für neuronale Netze hängt von der gewünschten Genauigkeitsstufe ab. Sie können die Laufzeit durch Festlegung einer Stoppbedingung im Dialogfeld des Knotens steuern.

K-Means. Der K-Means-Clusteralgorithmus weist dieselben Optionen für die Steuerung der Arbeitsspeicherverwendung auf wie die Algorithmen für neuronale Netze. Die Leistung für auf Datenträger gespeicherte Daten ist jedoch besser, weil der Zugriff auf die Daten sequenziell erfolgt.

Leistung: CLEM-Ausdrücke

CLEM-Sequenzfunktionen ("@-Funktionen"), die einen Rückblick auf den Datenstream ermöglichen, müssen so viele Daten speichern, wie für den längsten Rückblick erforderlich sind. Bei Operationen mit unbegrenztem Rückblicksgrad müssen alle Werte im Feld gespeichert werden. Eine unbegrenzte Operation ist eine Operation, bei der es sich beim Offset-Wert nicht um ein ganzzahliges Literal handelt; beispielsweise @OFFSET(Verkäufe, Monat). Der Offset-Wert ist der Feldname *Monat*, dessen Wert bis zur Ausführung unbekannt ist. Der Server muss alle Werte des Felds *Verkäufe* speichern, um genaue Ergebnisse zu gewährleisten. Wenn eine Obergrenze bekannt ist, sollten Sie sie als zusätzliches Argument angeben; Beispiel: @OFFSET(Verkäufe, Monat, 12). Diese Operation weist den Server an, nicht mehr als die 12 aktuellsten Werte von *Verkäufe* zu speichern. Sequenzfunktionen, ob mit Grenzen oder ohne, behindern fast immer die SQL-Generierung.

Kapitel 16. Eingabehilfen in IBM SPSS Modeler

Übersicht über die Eingabehilfen in IBM SPSS Modeler

IBM SPSS Modeler bietet Eingabehilfen für alle Benutzer sowie besondere Unterstützung für Benutzer mit beeinträchtigtem Sehvermögen oder anderen funktionalen Einschränkungen. In diesem Abschnitt werden die Funktionen und Methoden für die Arbeit mit erweiterten Eingabehilfen, wie Sprachausgabeprogramm und Tastenkombinationen, beschrieben.

Arten der Unterstützung von Eingabehilfen

Ob Sie nun unter einer Sehbehinderung leiden oder aus anderen Gründen auf die Tastatur angewiesen sind, es gibt eine Vielzahl alternativer Methoden zur Verwendung dieses Data-Mining-Toolkits. Sie können z. B. Streams erstellen, Optionen angeben und die Ausgabe lesen, ohne die Maus zu verwenden. Die verfügbaren Tastenkombinationen werden in den folgenden Themen beschrieben. Außerdem bietet IBM SPSS Modeler umfangreiche Unterstützung für Sprachausgabeprogramme, wie JAWS für Windows. Sie können auch das Farbschema ändern, um für zusätzlichen Kontrast zu sorgen. Die verschiedenen Möglichkeiten zur Unterstützung werden in den folgenden Themen beschrieben.

Zugriffsmöglichkeiten für Sehbehinderte

In IBM SPSS Modeler gibt es viele Eigenschaften, die Sie individuell festlegen können, um besser mit der Software arbeiten zu können.

Anzeigeoptionen

Sie können zum Anzeigen von Diagrammen verschiedene Farben auswählen. Sie können für die Software auch Ihre individuellen Windows-Einstellungen übernehmen. So wird der visuelle Kontrast unter Umständen erhöht.

1. Um die Anzeigeoptionen einzustellen, klicken Sie im Menü "Tools" auf die Option **Benutzeroptionen**.
2. Klicken Sie auf die Registerkarte **Anzeige**. Die Optionen auf dieser Registerkarte beziehen sich auf das Farbschema der Software, die Farben für Diagramme sowie die Schriftgrößen für Knoten.

Anmerkung: Da das Sprachausgabeprogramm keine Grafiken lesen kann, sind diese für sehbehinderte Benutzer nicht zugänglich.

Verwenden von Tönen zur Benachrichtigung

Durch das Aktivieren bzw. Inaktivieren von Tönen können Sie steuern, wie Ihnen bestimmte Operationen in der Software gemeldet werden. Zum Beispiel können Sie Töne für Ereignisse aktivieren, wie die Erstellung von Knoten und die Löschung bzw. Generierung neuer Ausgabe bzw. Modelle.

1. Um die Benachrichtigungsoptionen einzustellen, klicken Sie im Menü "Tools" auf die Option **Benutzeroptionen**.
2. Klicken Sie auf die Registerkarte **Benachrichtigungen**.

Steuern des automatischen Aufrufs neuer Fenster

Die Registerkarte "Benachrichtigungen" im Dialogfeld "Benutzeroptionen" wird auch dazu verwendet, zu steuern, ob die neu generierte Ausgabe, wie Tabellen und Diagramme, in einem separaten Fenster aufgerufen wird. Unter Umständen ist es für Sie einfacher, diese Option zu inaktivieren und nur im Bedarfsfall ein Ausgabefenster zu öffnen.

1. Um diese Optionen einzustellen, klicken Sie im Menü "Tools" auf die Option **Benutzeroptionen**.

2. Klicken Sie auf die Registerkarte **Benachrichtigungen**.
3. Wählen Sie im Dialogfeld in der Liste in der Gruppe **Visuelle Benachrichtigungen** die Option **Neue Ausgabe** aus.
4. Wählen Sie unter **Fenster öffnen** die Option **Nie** aus.

Knotengröße

Knoten können entweder in der Standardgröße oder in einer kleineren Größe angezeigt werden. Sie können die verschiedenen Größeneinstellungen an Ihre Anforderungen anpassen.

1. Um die Größenoptionen für die Knoten festzulegen, klicken Sie im Menü "Datei" auf die Option **Streameigenschaften**.
2. Klicken Sie auf die Registerkarte **Layout**.
3. Wählen Sie in der Liste **Symbolgröße** die Option **Standard** aus.

Zugriffsmöglichkeiten für Blinde

Die Unterstützung für blinde Benutzer ist nur möglich, wenn ein Sprachausgabeprogramm verwendet wird, wie z. B. JAWS für Windows. Um die Verwendung eines Sprachausgabeprogramms in IBM SPSS Modeler zu optimieren, können Sie verschiedene Einstellungen festlegen.

Anzeigeoptionen

Sprachausgabeprogramme funktionieren in der Regel besser, wenn der visuelle Bildschirmkontrast höher ist. Wenn bereits eine Windows-Einstellung mit hohem Kontrast gewählt ist, können Sie diese Windows-Einstellungen für die Software verwenden.

1. Um die Anzeigeoptionen einzustellen, klicken Sie im Menü "Tools" auf die Option **Benutzeroptionen**.
2. Klicken Sie auf die Registerkarte **Anzeige**.

Anmerkung: Da das Sprachausgabeprogramm keine Grafiken lesen kann, sind diese für blinde Benutzer nicht zugänglich.

Verwenden von Tönen zur Benachrichtigung

Durch das Aktivieren bzw. Inaktivieren von Tönen können Sie steuern, wie Ihnen bestimmte Operationen in der Software gemeldet werden. Zum Beispiel können Sie Töne für Ereignisse aktivieren, wie die Erstellung von Knoten und die Löschung bzw. Generierung neuer Ausgabe bzw. Modelle.

1. Um die Benachrichtigungsoptionen einzustellen, klicken Sie im Menü "Tools" auf die Option **Benutzeroptionen**.
2. Klicken Sie auf die Registerkarte **Benachrichtigungen**.

Steuern des automatischen Aufrufs neuer Fenster

Die Registerkarte "Benachrichtigungen" im Dialogfeld "Benutzeroptionen" wird auch dazu verwendet, zu steuern, ob die neu generierte Ausgabe in einem separaten Fenster aufgerufen wird. Unter Umständen ist es für Sie einfacher, diese Option zu inaktivieren und nur im Bedarfsfall ein Ausgabefenster zu öffnen.

1. Um diese Optionen einzustellen, klicken Sie im Menü "Tools" auf die Option **Benutzeroptionen**.
2. Klicken Sie auf die Registerkarte **Benachrichtigungen**.
3. Wählen Sie im Dialogfeld in der Liste in der Gruppe **Visuelle Benachrichtigungen** die Option **Neue Ausgabe** aus.
4. Wählen Sie unter **Fenster öffnen** die Option **Nie** aus.

Tastatureingabehilfen

Die Produktfunktionen können über die Tastatur aufgerufen werden. So können Sie ALT und die entsprechende Taste drücken, um Fenstermenüs zu aktivieren (z. B. Alt+D, um das Menü "Datei" aufzurufen), oder die Tabulatortaste drücken, um durch die Dialogfeldsteuerelemente zu blättern. Darüber hinaus gibt es spezielle Zugriffsmöglichkeiten, die nur für die einzelnen Hauptfenster des Produkts gelten, und hilfreiche Hinweise zum Navigieren in Dialogfeldern.

In diesem Abschnitt werden die wichtigsten Tastatureingabehilfen beschrieben, vom Öffnen eines Streams über die Verwendung von Knotendialogfeldern bis hin zur Arbeit mit Ausgaben. Für eine noch wirksamere Navigation stehen Listen mit Tastenkombinationen zur Verfügung.

Tastenkombinationen für die Navigation im Hauptfenster

Die meiste Data-Mining-Arbeit findet im Hauptfenster von IBM SPSS Modeler statt. Der Hauptbereich ist der **Streamerstellungsbereich**, der zum Erstellen und Ausführen von Datenstreams verwendet wird. Der untere Teil des Fensters enthält die **Knotenpaletten** mit allen verfügbaren Knoten. Die Paletten werden je nach Typ der Data-Mining-Operation für jede Knotengruppe auf Registerkarten angeordnet. Knoten, die zum Importieren von Daten in IBM SPSS Modeler verwendet werden, sind auf der Registerkarte "Datenquellen" gruppiert und Knoten, die zum Ableiten, Filtern und Eingeben von Feldern verwendet werden, sind auf der Registerkarte "Feldoperationen" gruppiert.

Die rechte Seite des Fensters enthält mehrere Tools zur Verwaltung von Streams, Ausgaben und Projekten. Die obere Hälfte auf der rechten Seite enthält die **Manager** mit drei Registerkarten, die zum Verwalten von Streams, Ausgaben und generierten Modellen verwendet werden. Sie können diese Objekte auswählen, indem Sie die Registerkarte und ein Objekt aus der Liste auswählen. Die untere Hälfte auf der rechten Seite enthält den **Projektbereich**, mit dem Sie Ihre Arbeit in Projekte aufteilen können. In diesem Bereich gibt es zwei Registerkarten, mit denen zwei unterschiedliche Ansichten eines Projekts möglich sind. Die Ansicht **Klassen** sortiert Projektobjekte nach Typ, während die Ansicht **CRISP-DM** Objekte nach der entsprechenden Data-Mining-Phase sortiert, z. B. Datenvorbereitung oder Modellierung. Diese verschiedenen Aspekte des IBM SPSS Modeler-Fensters werden im gesamten Hilfesystem sowie im Benutzerhandbuch erläutert.

Die folgende Tabelle enthält Tastenkombinationen, mit denen Sie im IBM SPSS Modeler-Hauptfenster navigieren und Streams erstellen können. Die Tastenkombinationen für Dialogfelder und die Ausgabe werden in den folgenden Themen aufgelistet. Beachten Sie, dass diese Tastenkombinationen nur im Hauptfenster zur Verfügung stehen.

Tabelle 39. Tastenkombinationen im Hauptfenster

Tastenkombination	Funktion
Strg+F5	Verschiebt den Fokus zu den Knotenpaletten.
Strg+F6	Verschiebt den Fokus zum Streamerstellungsbereich.
Strg+F7	Verschiebt den Fokus zum Managerbereich.
Strg+F8	Verschiebt den Fokus zum Projektbereich.

Tabelle 40. Knoten- und Streamtastenkombinationen

Tastenkombination	Funktion
Strg+N	Erstellt einen neuen, leeren Streamerstellungsbereich.
Strg+O	Zeigt das Dialogfeld "Öffnen" an, über das Sie bereits vorhandene Streams auswählen und öffnen können.
Strg+Zifferntasten	Verschiebt den Fokus in einem Fenster zur entsprechenden Registerkarte. In einem Fensterbereich oder in einem Fenster mit Registerkarten springt Strg+1 zur ersten Registerkarte von links, Strg+2 zur zweiten usw.

Tabelle 40. Knoten- und Streamtastenkombinationen (Forts.)

Tastenkombination	Funktion
Strg+Abwärtspfeil	Wird in der Knotenpalette zum Verschieben des Fokus von einer Palettenregisterkarte zum ersten Knoten auf dieser Registerkarte verwendet.
Strg+Aufwärtspfeil	Wird in der Knotenpalette zum Verschieben des Fokus von einem Knoten zu dessen Palettenregisterkarte verwendet.
Eingabetaste	Wenn in der Knotenpalette ein Knoten ausgewählt ist (einschließlich verfeinerter Modelle in der Palette der generierten Modelle), wird der Knoten durch Drücken dieser Taste in den Streamerstellungsbereich eingefügt. Durch Drücken der Eingabetaste wird bei bereits ausgewähltem Knoten im Erstellungsbereich das Dialogfeld für diesen Knoten geöffnet.
Strg+Eingabetaste	Wenn in der Palette ein Knoten ausgewählt ist, wird dieser Knoten durch diese Tastenkombination dem Streamerstellungsbereich hinzugefügt, ohne dass dieser ausgewählt wird. Gleichzeitig wird der Fokus auf den ersten Knoten in der Palette verschoben.
ALT+Eingabetaste	Wenn in der Palette ein Knoten ausgewählt ist, wird dieser Knoten durch diese Tastenkombination dem Streamerstellungsbereich hinzugefügt und ausgewählt. Gleichzeitig wird der Fokus auf den ersten Knoten in der Palette verschoben.
Umschalt+Leertaste	Wenn ein Knoten in der Palette einen Fokus aufweist, wird hiermit der Knoten oder Kommentar ausgewählt bzw. seine Auswahl aufgehoben. Wenn auch andere Knoten oder Kommentare ausgewählt sind, wird deren Auswahl ebenfalls aufgehoben.
Strg+Umschalt+Leertaste	Wenn ein Knoten oder Kommentar im Stream oder ein Knoten oder Kommentar in der Palette einen Fokus aufweist, wird hiermit der Knoten oder Kommentar ausgewählt bzw. seine Auswahl aufgehoben. Dies beeinflusst keine anderen ausgewählten Knoten oder Kommentare.
Linkspfeil/Rechtspfeil	Verschiebt den gesamten Stream horizontal, wenn der Fokus auf dem Streamerstellungsbereich ist. Wenn der Fokus auf einer Palettenregisterkarte liegt, wird zwischen den Registerkarten gewechselt. Wenn der Fokus auf einem Palettenknoten liegt, wird zwischen den Knoten in der Palette gewechselt.
Aufwärtspfeil/Abwärtspfeil	Verschiebt den gesamten Stream vertikal, wenn der Fokus auf dem Streamerstellungsbereich ist. Wenn der Fokus auf einem Palettenknoten liegt, wird zwischen den Knoten in der Palette gewechselt. Wenn der Fokus auf einer Unterpalette liegt, wird zwischen anderen Unterpaletten dieser Palettenregisterkarte gewechselt.
Alt+Linkspfeil/Rechtspfeil	Verschiebt ausgewählte Knoten und Kommentare im Streamerstellungsbereich horizontal in Richtung der Pfeiltaste.
Alt+Aufwärtspfeil/Abwärtspfeil	Verschiebt ausgewählte Knoten und Kommentare im Streamerstellungsbereich vertikal in Richtung der Pfeiltaste.
Strg+A	Wählt alle Knoten in einem Stream aus.
Strg+Q	Mit dieser Tastenkombination werden der Knoten, auf dem der Fokus ist, sowie alle nachgeordneten Knoten ausgewählt. Die Auswahl aller vorgeordneten Knoten wird aufgehoben.
Strg+W	Mit dieser Tastenkombination wird bei einem ausgewählten Knoten, auf dem der Fokus ist, sowie allen nachgeordneten Knoten die Auswahl aufgehoben.
Strg+Alt+D	Mit dieser Tastenkombination wird ein ausgewählter Knoten dupliziert.

Tabelle 40. Knoten- und Streamtastenkombinationen (Forts.)

Tastenkombination	Funktion
Strg+Alt+L	Wenn im Stream ein Modellnugget ausgewählt ist, wird ein Dialogfeld "Einfügen" geöffnet, über das Sie ein gespeichertes Modell aus einer .nod-Datei in den Stream laden können.
Strg+Alt+R	Zeigt für einen ausgewählten Knoten die Registerkarte "Anmerkungen" an, über die Sie den Knoten umbenennen können.
Strg+Alt+U	Erstellt einen Benutzereingabe-Quellenknoten.
Strg+Alt+C	Mit dieser Tastenkombination wird der Cache für einen Knoten aktiviert/inaktiviert.
Strg+Alt+F	Leert den Cache für einen Knoten.
Tabulator	Auf dem Streamerstellungsbereich wird zwischen allen Quellenknoten und Kommentaren im aktuellen Stream gewechselt. In einer Knotenpalette wird zwischen den Knoten in der Palette gewechselt. In einer ausgewählten Unterpalette wird zum ersten Knoten in der Unterpalette gewechselt.
Umschalt+Tabulatortaste	Führt denselben Befehl wie die Tabulatortaste aus, jedoch in umgekehrter Reihenfolge.
Strg+Tabulatortaste	Wenn sich der Fokus im Managerbereich oder im Projektbereich befindet, wird der Fokus in den Streamerstellungsbereich verschoben. Wenn sich der Fokus in einer Knotenpalette befindet, wird der Fokus zwischen einem Knoten und seiner Palettenregisterkarte gewechselt.
Beliebige Buchstabetaste	Wenn sich der Fokus auf einem Knoten im aktuellen Stream befindet, wird der Fokus auf den nächsten Knoten gesetzt, dessen Name mit der gedrückten Taste beginnt, und springt dorthin.
F1	Startet das Hilfesystem bei einem für den Fokus relevanten Thema.
F2	Startet den Verbindungsprozess für einen im Erstellungsbereich ausgewählten Knoten. Verwenden Sie die Tabulatortaste, um zum gewünschten Knoten im Erstellungsbereich zu gelangen, und drücken Sie Umschalt+Leertaste, um die Verbindung zu beenden.
F3	Löscht alle Verbindungen für den ausgewählten Knoten im Erstellungsbereich.
F6	Verschiebt den Fokus zwischen dem Managerbereich, dem Projektbereich und der Knotenpalette.
F10	Öffnet das Menü "Datei".
Umschalt+F10	Öffnet das Popup-Menü für den Knoten oder Stream.
Entf	Löscht einen ausgewählten Knoten aus dem Erstellungsbereich.
Esc	Schließt ein Popup-Menü oder Dialogfeld.
Strg+Alt+X	Mit dieser Tastenkombination wird der Superknoten erweitert.
Strg+Alt+Z	Vergrößert einen Superknoten.
Strg+Alt+Umschalt+Z	Verkleinert einen Superknoten.
Strg+E	Wenn sich der Fokus im Streamerstellungsbereich befindet, wird mit dieser Tastenkombination der aktuelle Stream ausgeführt.

In IBM SPSS Modeler werden auch viele der Standardtastenkombinationen verwendet, wie Strg+C zum Kopieren. Weitere Informationen finden Sie in „Verwenden von Tastenkombinationen“ auf Seite 22.

Tastenkombinationen für Dialogfelder und Tabellen

Einige der Tastenkombinationen und das Sprachausgabeprogramm sind hilfreich, wenn Sie mit Dialogfeldern, Tabellen und Tabellen in Dialogfeldern arbeiten. Nachfolgend erhalten Sie eine vollständige Liste mit speziellen Tastenkombinationen und das Sprachausgabeprogramm.

Table 41. Dialogfeld- und Expression Builder-Tastenkombinationen

Tastenkombination	Funktion
Alt+4	Wird zum Verwerfen aller geöffneten Dialogfelder oder Ausgabefenster verwendet. Die Ausgabe kann auf der Registerkarte "Ausgaben" im Managerbereich abgerufen werden.
Strg+Ende	Wenn sich der Fokus auf einem beliebigen Steuerelement im Expression Builder befindet, wird die Einfügemarke mit dieser Tastenkombination an das Ende des Ausdrucks verschoben.
Strg+1	Im Expression Builder wird der Fokus mit dieser Tastenkombination zum Steuerelement für die Bearbeitung des Ausdrucks verschoben.
Strg+2	Im Expression Builder wird der Fokus mit dieser Tastenkombination zur Funktionsliste verschoben.
Strg+3	Mit dieser Tastenkombination wird der Fokus im Expression Builder zur Feldliste verschoben.

Tastenkombinationen für Tabellen

Tabellentastenkombinationen werden für Ausgabebibliotheken sowie für Tabellensteuerelemente in Dialogfeldern für Knoten, wie Typ, Filter und Zusammenführen, verwendet. In der Regel verwenden Sie die Tabulatortaste, um zwischen Tabellenzellen zu navigieren, und Strg+Tab, um die Tabellensteuerung zu verlassen. *Hinweis:* Es kann vorkommen, dass ein Sprachausgabeprogramm nicht sofort mit dem Lesen des Inhalts einer Zelle beginnt. Wenn Sie die Pfeiltasten ein- oder zweimal drücken, wird die Software zurückgesetzt und die Sprachausgabe beginnt.

Table 42. Tastenkombinationen für Tabellen

Tastenkombination	Funktion
Strg+W	Mit dieser Tastenkombination wird die Kurzbeschreibung der ausgewählten Zeile in Tabellen gelesen. Zum Beispiel: "Die ausgewählten Werte aus Zeile 2 sind Geschlecht, Flag, m/w usw."
Strg+Alt+W	Mit dieser Tastenkombination wird die lange Beschreibung der ausgewählten Zeile in Tabellen gelesen. Zum Beispiel: "Die ausgewählten Werte aus Zeile 2 sind Feld = Geschlecht, Typ = Flag, Geschlecht = m/w usw."
Strg+D	Mit dieser Tastenkombination wird die Kurzbeschreibung des ausgewählten Bereichs in Tabellen gelesen. Zum Beispiel: "Auswahl ist eine Zeile mal sechs Spalten."
Strg+Alt+D	Mit dieser Tastenkombination wird die lange Beschreibung des ausgewählten Bereichs in Tabellen gelesen. Zum Beispiel: "Auswahl ist eine Zeile mal sechs Spalten. Die ausgewählten Spalten sind Feld, Typ, Fehlend. Die ausgewählte Zeile ist 1."
Strg+T	Diese Tastenkombination zeigt in Tabellen eine Kurzbeschreibung der ausgewählten Spalten an. Zum Beispiel: "Felder, Typ, Fehlend."
Strg+Alt+T	Diese Tastenkombination zeigt in Tabellen eine lange Beschreibung der ausgewählten Spalten an. Zum Beispiel: "Ausgewählte Spalten sind Felder, Typ, Fehlend."
Strg+R	Mit dieser Tastenkombination wird in Tabellen die Anzahl der in der Tabelle vorhandenen Datensätze angezeigt.

Tabelle 42. Tastenkombinationen für Tabellen (Forts.)

Tastenkombination	Funktion
Strg+Alt+R	Mit dieser Tastenkombination wird in Tabellen die Anzahl der in der Tabelle vorhandenen Datensätze sowie Spaltennamen angezeigt.
Strg+I	Mit dieser Tastenkombination werden in Tabellen die Zelleninformation bzw. der Inhalt für die Zelle, auf der der Fokus ist, gelesen.
Strg+Alt+I	Mit dieser Tastenkombination wird in Tabellen die lange Beschreibung der Zelleninformation (Spaltenname und Inhalt der Zelle) für die Zelle, auf der der Fokus ist, gelesen.
Strg+G	Diese Tastenkombination zeigt für Tabellen kurze allgemeine Auswahlinformationen an.
Strg+Alt+G	Diese Tastenkombination zeigt für Tabellen lange allgemeine Auswahlinformationen an.
Strg+Q	Ermöglicht in Tabellen ein schnelles Wechseln zwischen den Tabellenzellen. Mit Strg+Q werden lange Beschreibungen, wie "Geschlecht=Weiblich", gelesen, wenn Sie mit den Pfeiltasten durch die Tabelle blättern. Wenn Sie Strg+Q erneut auswählen, werden die kurzen Beschreibungen (Zelleninhalte) angezeigt.
F8	Versetzt in Tabellen den Fokus von der Tabelle auf die Spaltenüberschrift.
Leertaste	Ermöglicht in Tabellen die Spaltensortierung, wenn der Fokus sich in der Spaltenüberschrift befindet.

Tastenkombinationen für Kommentare

Beim Arbeiten mit Kommentaren am Bildschirm können Sie die folgenden Tastenkombinationen verwenden.

Tabelle 43. Tastenkombinationen für Kommentare

Tastenkombination	Funktion
Alt+C	Aktiviert/inaktiviert die Funktion "Kommentare ein-/ausblenden".
Alt+M	Fügt einen neuen Kommentar ein, falls Kommentare angezeigt werden, und zeigt Kommentare an, wenn Sie gerade ausgeblendet sind.
Tabulator	Auf dem Streamerstellungsbereich wird zwischen allen Quellenknoten und Kommentaren im aktuellen Stream gewechselt.
Eingabetaste	Wenn ein Kommentar den Fokus aufweist, gibt dies den Beginn der Bearbeitung an.
Alt+Eingabetaste oder Strg+Tabulatortaste	Beendet die Bearbeitung und speichert vorgenommene Änderungen.
Esc	Bricht die Bearbeitung ab. Vorgenommene Änderungen gehen verloren.
Alt+Umschalt+Aufwärtspfeil	Verringert die Höhe des Textbereichs um eine Rasterzelle (oder ein Pixel), wenn die Funktion zum Ausrichten am Raster aktiviert (bzw. inaktiviert) ist.
Alt+Umschalt+Abwärtspfeil	Vergrößert die Höhe des Textbereichs um eine Rasterzelle (oder ein Pixel), wenn die Funktion zum Ausrichten am Raster aktiviert (bzw. inaktiviert) ist.
Alt+Umschalt+Linkspfeil	Verringert die Breite des Textbereichs um eine Rasterzelle (oder ein Pixel), wenn die Funktion zum Ausrichten am Raster aktiviert (bzw. inaktiviert) ist.

Tabelle 43. Tastenkombinationen für Kommentare (Forts.)

Tastenkombination	Funktion
Alt+Umschalt+Rechtspfeil	Vergrößert die Breite des Textbereichs um eine Rasterzelle (oder ein Pixel), wenn die Funktion zum Ausrichten am Raster aktiviert (bzw. inaktiviert) ist.

Tastenkombinationen für Cluster-Viewer und Modellviewer

Für die Navigation in den Fenstern des Cluster-Viewers und des Modellviewers stehen Tastenkombinationen zur Verfügung.

Tabelle 44. Allgemeine Tastenkombinationen - Cluster-Viewer und Modellviewer

Tastenkombination	Funktion
Tabulator	Verschiebt den Fokus zum nächsten Bildschirmsteuerelement.
Umschalt+Tabulatortaste	Verschiebt den Fokus zum vorherigen Bildschirmsteuerelement.
Abwärtspfeil	Wenn eine Dropdown-Liste den Fokus enthält, wird die Liste geöffnet bzw. erhält das nächste Element in der Liste den Fokus. Wenn ein Menü den Fokus enthält, wird der Fokus zum nächsten Element im Menü verschoben. Wenn sich der Fokus auf einem Piktogramm befindet, wird der Fokus zum nächsten Element im Set (oder zum ersten, wenn der Fokus auf dem letzten Piktogramm liegt) verschoben.
Aufwärtspfeil	Wenn eine Dropdown-Liste den Fokus enthält, wird der Fokus zum vorherigen Element in der Liste verschoben. Wenn ein Menü den Fokus enthält, wird der Fokus zum vorherigen Element im Menü verschoben. Wenn sich der Fokus auf einem Piktogramm befindet, wird der Fokus zum vorherigen Element im Set (oder zum letzten, wenn der Fokus auf dem ersten Piktogramm liegt) verschoben.
Eingabetaste	Schließt eine geöffnete Dropdown-Liste oder nimmt eine Auswahl in einem geöffneten Menü vor.
F6	Schaltet den Fokus zwischen dem linken und rechten Bereich des Fensters um.
Linkspfeil und Rechtspfeil	Wenn eine Registerkarte den Fokus enthält, wird der Fokus zur vorherigen bzw. nächsten Registerkarte verschoben. Wenn ein Menü den Fokus enthält, wird der Fokus zum vorherigen bzw. nächsten Menü verschoben.
Alt+Buchstabe	Wählt die Schaltfläche bzw. das Menü aus, in dessen Namen dieser Buchstabe unterstrichen ist.
Esc	Schließt ein geöffnetes Menü oder Dropdown-Listenfeld.

Nur Cluster-Viewer

Im Cluster-Viewer kann ein Cluster-nach-Funktionen-Raster angezeigt werden.

So wählen Sie die Clusteransicht anstelle der Ansicht "Modellübersicht":

1. Drücken Sie wiederholt die Tabulatortaste, bis die Schaltfläche **Ansicht** ausgewählt ist.
2. Drücken Sie zweimal die Abwärtspfeiltaste, um **Cluster** auszuwählen.

Von hier aus können Sie eine einzelne Zelle im Raster auswählen:

3. Drücken Sie wiederholt die Tabulatortaste, bis Sie das letzte Symbol in der Visualisierungssymbolleiste erreichen.



Abbildung 19. Symbol "Visualisierungsbaum anzeigen"

4. Drücken Sie noch einmal die Tabulatortaste, anschließend die Leertaste und dann eine Pfeiltaste.

Nun stehen die folgenden Tastenkombinationen zur Verfügung:

Tabelle 45. Cluster-Viewer - Tastenkombinationen

Tastenkombination	Funktion
Pfeiltaste	Verschiebt den Fokus zwischen einzelnen Zellen im Raster. Die Anzeigegellenverteilung im rechten Bereich ändert sich mit dem Verschieben des Fokus.
Strg+, (Komma)	Wählt die ganze Spalte im Raster aus, in der eine Zelle den Fokus enthält, bzw. hebt die Auswahl der ganzen Spalte auf. Sie fügen der Auswahl eine Spalte hinzu, indem Sie mithilfe der Pfeiltasten zu einer Zelle in der entsprechenden Spalte wechseln und erneut Strg+, drücken.
Tabulator	Verschiebt den Fokus aus dem Raster und zum nächsten Bildschirmsteuerelement.
Umschalt+Tabulatortaste	Verschiebt den Fokus aus dem Raster und zum vorherigen Bildschirmsteuerelement.
F2	Wechselt in den Bearbeitungsmodus (nur Beschriftungen und Beschreibungszellen).
Eingabetaste	Speichert Bearbeitungsänderungen und beendet den Bearbeitungsmodus (nur Beschriftungen und Beschreibungszellen).
Esc	Beendet den Bearbeitungsmodus, ohne Änderungen zu speichern (nur Beschriftungen und Beschreibungszellen).

Beispiel für Tastenkombinationen: Erstellen von Streams

Um den Streamerstellungsprozess für Benutzer zu verdeutlichen, die auf die Tastatur oder ein Sprachausgabeprogramm angewiesen sind, folgt nun ein Beispiel für eine Streamerstellung ohne Verwendung der Maus. In diesem Beispiel erstellen Sie einen Stream, der einen Knoten vom Typ "Datei (var.)", einen Knoten vom Typ "Ableitung" und einen Knoten vom Typ "Histogramm" enthält. Gehen Sie dazu wie folgt vor:

1. **Starten Sie IBM SPSS Modeler.** Beim Start von IBM SPSS Modeler befindet sich der Fokus auf der Registerkarte "Favoriten" der Knotenpalette.
2. **Strg+Abwärtspfeil.** Verschiebt den Fokus von der Registerkarte selbst zum Inhalt der Registerkarte.
3. **Rechtspfeil.** Verschiebt den Fokus zum Knoten "Variable Datei".
4. **Leertaste.** Wählt den Knoten "Variable Datei" aus.
5. **Strg+Eingabetaste.** Nimmt den neuen Variablendateiknoten in den Streamerstellungsbereich auf. Mit dieser Tastenkombination bleibt die Auswahl am Knoten "Variable Datei", sodass der nächste hinzugefügte Knoten damit verbunden wird.
6. **Tabulatortaste.** Verschiebt den Fokus zurück in die Knotenpalette.
7. **Viermal Rechtspfeil.** So gelangen Sie zum Ableitungsknoten.
8. **Leertaste.** Mit dieser Taste wird der Ableitungsknoten ausgewählt.

9. **Alt+Eingabetaste.** Fügt den Ableitungsknoten im Erstellungsbereich ein und verschiebt die Auswahl zum Ableitungsknoten. Dieser Knoten kann nun mit dem nächsten hinzugefügten Knoten verbunden werden.
10. **Tabulatortaste.** Verschiebt den Fokus zurück in die Knotenpalette.
11. **Fünfmal Rechtspfeil.** Verschiebt den Fokus zum Knoten "Histogramm" in der Palette.
12. **Leertaste.** Wählt den Knoten "Histogramm" aus.
13. **Eingabetaste.** Fügt den Knoten dem Stream hinzu und verschiebt den Fokus in den Streamerstellungsbereich.

Fahren Sie mit dem nächsten Beispiel fort oder speichern Sie den Stream, falls Sie das nächste Beispiel zu einem späteren Zeitpunkt ausprobieren möchten.

Beispiel für Tastenkombinationen: Bearbeiten von Knoten

In diesem Beispiel verwenden Sie den im vorherigen Beispiel erstellten Stream. Der Stream besteht aus einem Knoten vom Typ "Datei (var.)", einem Ableitungsknoten und einem Knoten vom Typ "Histogramm". Die Anweisungen beginnen mit dem Fokus auf dem dritten Knoten im Stream, dem Knoten vom Typ "Histogramm".

1. **Strg+zweimal Linkspfeil.** Verschiebt den Fokus zurück zum Knoten "Datei (.var)".
2. **Eingabetaste.** Öffnet das Dialogfeld "Variable Datei". Wechseln Sie mithilfe der Tabulatortaste in das Feld "Datei" und geben Sie einen Dateipfad und einen Dateinamen ein, um die entsprechende Datei auszuwählen. Wechseln Sie dann mithilfe der Tastenkombination Strg+Tab in den unteren Bereich des Dialogfelds, drücken Sie die Tabulatortaste, bis die Schaltfläche "OK" ausgewählt ist, und drücken Sie die Eingabetaste, um das Dialogfeld zu schließen.
3. **Strg+Rechtspfeil.** Verschiebt den Fokus zum zweiten Knoten, einem Ableitungsknoten.
4. **Eingabetaste.** Öffnet das Dialogfeld "Ableitungsknoten". Blättern Sie mit der Tabulatortaste durch dieses Dialogfeld, um die gewünschten Felder auszuwählen und die Ableitungsbedingungen festzulegen. Wechseln Sie dann mithilfe der Tastenkombination Strg+Tab zur Schaltfläche "OK" und drücken Sie die Eingabetaste, um das Dialogfeld zu schließen.
5. **Strg+Rechtspfeil.** Verschiebt den Fokus zum dritten Knoten, einem Knoten vom Typ "Histogramm".
6. **Eingabetaste.** Öffnet das Dialogfeld "Histogrammknoten". Blättern Sie mit der Tabulatortaste durch dieses Dialogfeld, um die gewünschten Felder auszuwählen und Diagrammoptionen festzulegen. Für Dropdown-Listen drücken Sie die Abwärtspfeiltaste, um die Liste zu öffnen und ein Element zu markieren, und drücken Sie dann die Eingabetaste, um das Listenelement auszuwählen. Wechseln Sie dann mithilfe der Tabulatortaste zur Schaltfläche "OK" und drücken Sie die Eingabetaste, um das Dialogfeld zu schließen.

Zu diesem Zeitpunkt können Sie weitere Knoten hinzufügen oder den aktuellen Stream ausführen. Beachten Sie beim Erstellen von Streams folgende Hinweise:

- Verwenden Sie beim manuellen Verbinden von Knoten die Taste F2, um den Startpunkt einer Verbindung zu erstellen. Navigieren Sie dann mithilfe der Tabulatortaste zum Endpunkt und drücken Sie Umschalt+Leertaste, um die Verbindung fertigzustellen.
- Verwenden Sie die Taste F3, um alle Verbindungen für einen ausgewählten Knoten im Erstellungsbereich zu zerstören.
- Nachdem Sie einen Stream erstellt haben, drücken Sie Strg+E, um den aktuellen Stream auszuführen.

Eine vollständige Liste mit Tastenkombinationen ist vorhanden. Weitere Informationen finden Sie in „Tastenkombinationen für die Navigation im Hauptfenster“ auf Seite 245.

Verwenden eines Sprachausgabeprogramms

Auf dem Markt sind einige Sprachausgabeprogramme erhältlich. IBM SPSS Modeler ist für die Unterstützung von JAWS für Windows mithilfe der Java Access Bridge konfiguriert, die zusammen mit IBM SPSS Modeler installiert wird. Wenn JAWS installiert ist, starten Sie JAWS einfach vor dem Start von IBM SPSS Modeler, um das Produkt zu verwenden.

Anmerkung: Zum Ausführen von JAWS mit SPSS Modeler sollten mindestens 6 GB Speicherbereich verfügbar sein.

Aufgrund der Natur der einmaligen grafischen Darstellung des Data-Mining-Prozesses von IBM SPSS Modeler werden die Diagramme und Grafiken optimal visuell verwendet. Es ist jedoch möglich, dass Sie Entscheidungen basierend auf der Ausgabe und der Modelle treffen, die unter Verwendung eines Sprachausgabeprogramms textuell angezeigt werden.

Anmerkung: Auf 64-Bit-Clientmaschinen funktionieren bestimmte Technologien für behindertengerechte Bedienung nicht. Das liegt daran, dass die Java Access Bridge nicht für den 64-Bit-Betrieb konfiguriert ist.

Verwenden der Wörterbuchdatei von IBM SPSS Modeler

Eine Wörterbuchdatei von IBM SPSS Modeler (*Awt.JDF*) steht für die Verwendung mit JAWS zur Verfügung. So verwenden Sie diese Datei:

1. Navigieren Sie zum Unterverzeichnis */accessibility* Ihrer IBM SPSS Modeler-Installation und kopieren Sie die Wörterbuchdatei (*Awt.JDF*).
2. Kopieren Sie die Datei in das Verzeichnis, in dem sich Ihre JAWS-Scripts befinden.

Unter Umständen befindet sich bereits eine Datei namens *Awt.JDF* auf Ihrem System, falls andere JAVA-Anwendungen ausgeführt werden. In diesem Fall können Sie diese Wörterbuchdatei eventuell nicht verwenden, ohne die Wörterbuchdatei manuell zu bearbeiten.

Verwenden eines Sprachausgabeprogramms mit HTML-Ausgabe

Wenn Sie unter Verwendung eines Sprachausgabeprogramms eine Ausgabe anzeigen, die in IBM SPSS Modeler als HTML angezeigt wird, können eventuell einige Schwierigkeiten auftreten. Es sind mehrere Ausgabetypen betroffen, darunter:

- Ausgabe, die auf der Registerkarte "Erweitert" für Knoten vom Typ "Regression", "Logistische Regression" und "Faktor/PCA" angezeigt wird
- Ausgabe des Berichtsknotens

In jedem dieser Fenster oder Dialogfelder gibt es ein Tool in der Symbolleiste, das zum Aufrufen der Ausgabe in Ihrem Standardbrowser verwendet wird, der Unterstützung für ein Standardsprachausgabeprogramm bietet. Sie können das Sprachausgabeprogramm dann verwenden, um die Ausgabeinformationen zu übermitteln.

Eingabehilfen im Fenster "Interaktiver Baum"

Die Standardanzeige eines Entscheidungsbaummodells im Fenster "Interaktiver Baum" kann zu Problemen mit Sprachausgabeprogrammen führen. Um eine Version mit Eingabehilfen aufzurufen, wählen Sie in den Menüs "Interaktiver Baum":

Ansicht > Fenster mit Eingabehilfen

Die Anzeige ist ähnlich wie die Standardbaumübersicht, kann jedoch von JAWS ordnungsgemäß gelesen werden. Mit den Standardpfeiltasten können Sie sich nach oben, nach unten, nach rechts oder nach links bewegen. Bei der Navigation im Fenster mit Eingabehilfen ändert sich der Fokus im Fenster "Interaktiver Baum" entsprechend. Verwenden Sie die Leertaste, um die Auswahl zu ändern, oder Strg+Leertaste, um die aktuelle Auswahl zu erweitern.

Tipps zur Verwendung

Es gibt mehrere Tipps, wie die IBM SPSS Modeler-Umgebung für Sie benutzerfreundlicher gestaltet werden kann. Im Folgenden erhalten Sie allgemeine Tipps für die Arbeit in IBM SPSS Modeler.

- **Verlassen von erweiterten Textfeldern.** Verwenden Sie Strg+Tab, um die erweiterten Textfelder zu verlassen. Beachten Sie, dass Strg+Tab auch zum Beenden von Tabellensteuerelementen verwendet wird.
- **Verwenden der Tabulatortaste anstelle der Pfeiltasten.** Verwenden Sie bei der Auswahl von Optionen für ein Dialogfeld die Tabulatortasten, um zwischen den Optionsfeldern hin- und herzuwechseln. Die Pfeiltasten funktionieren in diesem Zusammenhang nicht.
- **Dropdown-Listen.** In einer Dropdown-Liste für Dialogfelder können Sie entweder die Abbruchtaste oder die Leertaste verwenden, um ein Element auszuwählen, und dann die Liste schließen. Sie können die Abbruchtaste auch verwenden, um die Dropdown-Listen zu schließen, die nicht geschlossen werden können, wenn Sie mit der Tabulatortaste zu einem anderen Steuerelement wechseln.
- **Ausführungsstatus.** Wenn Sie einen Stream für eine große Datenbank ausführen, kann JAWS mit dem Lesen des Streamstatus hinterherhinken. Drücken Sie die Steuertaste in regelmäßigen Abständen, um den Statusbericht zu aktualisieren.
- **Verwenden der Knotenpaletten.** Wenn Sie zum ersten Mal eine Registerkarte der Knotenpaletten aufrufen, liest JAWS manchmal "Gruppenfeld" anstatt den Namen des Knotens. In diesem Fall können Sie Strg+Rechtspfeil und dann Strg+Linkspfeil verwenden, um das Sprachausgabeprogramm zurückzusetzen und den Knotennamen zu hören.
- **Lesen von Menüs.** Beim erstmaligen Öffnen eines Menüs kann es gelegentlich vorkommen, dass JAWS den ersten Menüeintrag nicht lesen kann. Wenn Sie vermuten, dass dies der Fall gewesen sein könnte, verwenden Sie die Abwärtspfeiltaste und dann die Aufwärtspfeiltaste, um das erste Element im Menü zu hören.
- **Untermenüfenster.** JAWS liest die erste Ebene eines überlappenden Menüs nicht. Wenn Sie beim Sprechen eine Unterbrechung hören, während Sie durch das Menü blättern, drücken Sie die Rechtspfeiltaste, um die untergeordneten Menüelemente zu hören.

Wenn auf Ihrem System IBM SPSS Modeler Text Analytics installiert ist, können Ihnen außerdem die folgenden Tipps die Arbeit mit der interaktiven Workbench-Schnittstelle erleichtern.

- **Aufrufen von Dialogfeldern.** Möglicherweise müssen Sie die Tabulatortaste drücken, um beim Aufrufen eines Dialogfelds den Fokus auf das erste Steuerelement zu legen.
- **Verlassen von erweiterten Textfeldern.** Verwenden Sie Strg+Tab, um die erweiterten Textfelder zu verlassen und zum nächsten Steuerelement überzugehen. Beachten Sie, dass Strg+Tab auch zum Beenden von Tabellensteuerelementen verwendet wird.
- **Eingabe des ersten Buchstabens zum Auffinden von Elementen in der Baumliste.** Beim Suchen nach einem Element im Fensterbereich für Kategorien, im Fensterbereich für die extrahierten Ergebnisse oder im Bibliotheksbaum können Sie den ersten Buchstaben des Elements eingeben, wenn sich der Fokus im entsprechenden Fensterbereich befindet. Dadurch wird das nächste Vorkommen eines Elements ausgewählt, das mit dem eingegebenen Buchstaben beginnt.
- **Dropdown-Listen.** In einer Dropdown-Liste für Dialogfelder können Sie die Leertaste verwenden, um ein Element auszuwählen, und dann die Liste schließen.

Kompatibilität mit anderer Software

Beim Testen von IBM SPSS Modeler mit Sprachausgabeprogrammen, wie JAWS, hat unser Entwicklungsteam festgestellt, dass die Verwendung eines Systems Management Servers (SMS) in Ihrem Unternehmen eventuell mit der Funktion von JAWS kollidiert, Java-basierte Anwendungen, wie IBM SPSS Modeler, zu lesen. Die Inaktivierung von SMS behebt diese Situation. Weitere Informationen zu SMS finden Sie auf der Microsoft-Website.

JAWS und Java

Die verschiedenen Versionen von JAWS bieten unterschiedliche Unterstützung für Java-basierte Software-Anwendungen. IBM SPSS Modeler funktioniert zwar mit allen neueren Versionen von JAWS, bei einigen Versionen können jedoch kleinere Probleme auftreten, wenn Sie zusammen mit Java-basierten Systemen eingesetzt werden. Besuchen Sie die Website von JAWS für Windows unter <http://www.FreedomScientific.com>.

Verwenden von Diagrammen in IBM SPSS Modeler

Die visuelle Anzeige von Informationen, wie Histogramme, Evaluierungsdiagramme, Multiplots und Streudiagramme, sind mit einem Sprachausgabeprogramm schwer zu interpretieren. Beachten Sie jedoch, dass die Webdiagramme und Verteilungen mithilfe der Textübersicht im Ausgabefenster angezeigt werden können.

Kapitel 17. Unicode-Unterstützung

Unicode-Unterstützung in IBM SPSS Modeler

IBM SPSS Modeler ist für IBM SPSS Modeler und IBM SPSS Modeler Server vollständig Unicode-kompatibel. Dadurch können Daten mit anderen Anwendungen ausgetauscht werden, die Unicode unterstützen, einschließlich mehrsprachiger Datenbanken, und das ohne jeden Informationsverlust, der durch die Konvertierung in ein oder aus einem Codierungsschema verursacht werden könnte, das für eine bestimmte Ländereinstellung spezifisch ist.

- IBM SPSS Modeler speichert Unicode-Daten intern und kann mehrsprachige Daten, die als Unicode in Datenbanken gespeichert sind, verlustfrei lesen und schreiben.
- IBM SPSS Modeler kann UTF-8-codierte Textdateien lesen und schreiben. Für den Import und Export von Textdateien wird standardmäßig die für die Ländereinstellung spezifische Codierung verwendet, doch UTF-8 wird als Alternative unterstützt. Diese Einstellung kann in den Import- bzw. Exportknoten der Datei festgelegt werden. Alternativ kann die Standardcodierung im Dialogfeld "Stromeigenschaften" geändert werden. Weitere Informationen finden Sie in „Festlegen allgemeiner Optionen für Streams“ auf Seite 44.
- Statistics-, SAS- und Textdateien, die in der für die Ländereinstellung spezifischen Codierung gespeichert sind, werden beim Import in UTF-8 konvertiert und beim Export wieder in das ursprüngliche Format zurückkonvertiert. Wenn beim Schreiben in eine Datei Unicode-Zeichen vorhanden sind, die im Zeichensatz der Ländereinstellung nicht vorhanden sind, werden diese Zeichen ersetzt und eine Warnung wird angezeigt. Dies sollte nur passieren, wenn die Daten aus einer Datenquelle importiert wurden, die Unicode unterstützt (Datenbank oder UTF-8-Textdatei) und Zeichen aus einer anderen Ländereinstellung oder aus mehreren Ländereinstellungen oder Zeichensätzen enthält.
- Bilder aus IBM SPSS Modeler Solution Publisher sind UTF-8-codiert und vollständig zwischen Plattformen und Ländereinstellungen übertragbar.

Informationen zu Unicode

Das Ziel des Unicode-Standards besteht darin, ein konsistentes Verfahren zur Codierung von mehrsprachigem Text bereitzustellen, sodass dieser problemlos über Grenzen, Ländereinstellungen und Anwendungen hinweg verwendet werden kann. Der Unicode-Standard, der mittlerweile in Version 4.0.1 vorliegt, definiert einen Zeichensatz, der eine Obermenge aller derzeit häufig verwendeten Zeichensätze weltweit darstellt, und weist jedem Zeichen einen eindeutigen Namen und Codepunkt zu. Die Zeichen und ihre Codepunkte stimmen mit denen des in ISO-10646 definierten Universal Character Set ("universeller Zeichensatz", Abk. UCS) überein. Weitere Informationen finden Sie auf der Unicode-Homepage.

Bemerkungen

Die vorliegenden Informationen wurden für Produkte und Services entwickelt, die auf dem deutschen Markt angeboten werden. IBM stellt dieses Material möglicherweise auch in anderen Sprachen zur Verfügung. Für den Zugriff auf das Material in einer anderen Sprache kann eine Kopie des Produkts oder der Produktversion in der jeweiligen Sprache erforderlich sein.

Möglicherweise bietet IBM die in dieser Dokumentation beschriebenen Produkte, Services oder Funktionen in anderen Ländern nicht an. Informationen über die gegenwärtig im jeweiligen Land verfügbaren Produkte und Services sind beim zuständigen IBM Ansprechpartner erhältlich. Hinweise auf IBM Lizenzprogramme oder andere IBM Produkte bedeuten nicht, dass nur Programme, Produkte oder Services von IBM verwendet werden können. Anstelle der IBM Produkte, Programme oder Services können auch andere, ihnen äquivalente Produkte, Programme oder Services verwendet werden, solange diese keine gewerblichen oder anderen Schutzrechte von IBM verletzen. Die Verantwortung für den Betrieb von Produkten, Programmen und Services anderer Anbieter liegt beim Kunden.

Für in diesem Handbuch beschriebene Erzeugnisse und Verfahren kann es IBM Patente oder Patentanmeldungen geben. Mit der Auslieferung dieses Handbuchs ist keine Lizenzierung dieser Patente verbunden. Lizenzanforderungen sind schriftlich an folgende Adresse zu richten (Anfragen an diese Adresse müssen auf Englisch formuliert werden):

*IBM Director of Licensing
IBM Europe, Middle East & Africa
Tour Descartes
2, avenue Gambetta
92066 Paris La Defense
France*

Diese Informationen können technische Ungenauigkeiten oder typografische Fehler enthalten. Die hier enthaltenen Informationen werden in regelmäßigen Zeitabständen aktualisiert und als Neuausgabe veröffentlicht. IBM kann ohne weitere Mitteilung jederzeit Verbesserungen und/oder Änderungen an den in dieser Veröffentlichung beschriebenen Produkten und/oder Programmen vornehmen.

Verweise in diesen Informationen auf Websites anderer Anbieter werden lediglich als Service für den Kunden bereitgestellt und stellen keinerlei Billigung des Inhalts dieser Websites dar. Das über diese Websites verfügbare Material ist nicht Bestandteil des Materials für dieses IBM Produkt. Die Verwendung dieser Websites geschieht auf eigene Verantwortung.

Werden an IBM Informationen eingesandt, können diese beliebig verwendet werden, ohne dass eine Verpflichtung gegenüber dem Einsender entsteht.

Lizenznehmer des Programms, die Informationen zu diesem Produkt wünschen mit der Zielsetzung: (i) den Austausch von Informationen zwischen unabhängig voneinander erstellten Programmen und anderen Programmen (einschließlich des vorliegenden Programms) sowie (ii) die gemeinsame Nutzung der ausgetauschten Informationen zu ermöglichen, wenden sich an folgende Adresse:

*IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
USA*

Die Bereitstellung dieser Informationen kann unter Umständen von bestimmten Bedingungen - in einigen Fällen auch von der Zahlung einer Gebühr - abhängig sein.

Die Lieferung des in diesem Dokument beschriebenen Lizenzprogramms sowie des zugehörigen Lizenzmaterials erfolgt auf der Basis der IBM Rahmenvereinbarung bzw. der Allgemeinen Geschäftsbedingungen von IBM, der IBM Internationalen Nutzungsbedingungen für Programmpakete oder einer äquivalenten Vereinbarung.

Die angeführten Leistungsdaten und Kundenbeispiele dienen nur zur Illustration. Die tatsächlichen Ergebnisse beim Leistungsverhalten sind abhängig von der jeweiligen Konfiguration und den Betriebsbedingungen.

Alle Informationen zu Produkten anderer Anbieter stammen von den Anbietern der aufgeführten Produkte, deren veröffentlichten Ankündigungen oder anderen allgemein verfügbaren Quellen. IBM hat diese Produkte nicht getestet und kann daher keine Aussagen zu Leistung, Kompatibilität oder anderen Merkmalen machen. Fragen zu den Leistungsmerkmalen von Produkten anderer Anbieter sind an den jeweiligen Anbieter zu richten.

Aussagen über Pläne und Absichten von IBM unterliegen Änderungen oder können zurückgenommen werden und repräsentieren nur die Ziele von IBM.

Diese Veröffentlichung enthält Beispiele für Daten und Berichte des alltäglichen Geschäftsablaufs. Sie sollen nur die Funktionen des Lizenzprogramms illustrieren und können Namen von Personen, Firmen, Marken oder Produkten enthalten. Alle diese Namen sind frei erfunden; Ähnlichkeiten mit tatsächlichen Namen und Adressen sind rein zufällig.

Marken

IBM, das IBM Logo und ibm.com sind Marken oder eingetragene Marken der IBM Corporation in den USA und/oder anderen Ländern. Weitere Produkt- und Servicennamen können Marken von IBM oder anderen Unternehmen sein. Eine aktuelle Liste der IBM Marken finden Sie auf der Webseite "Copyright and trademark information" unter www.ibm.com/legal/copytrade.shtml.

Adobe, das Adobe-Logo, PostScript und das PostScript-Logo sind Marken oder eingetragene Marken der Adobe Systems Incorporated in den USA und/oder anderen Ländern.

Intel, das Intel-Logo, Intel Inside, das Intel Inside-Logo, Intel Centrino, das Intel Centrino-Logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium und Pentium sind Marken oder eingetragene Marken der Intel Corporation oder ihrer Tochtergesellschaften in den USA oder anderen Ländern.

Linux ist eine eingetragene Marke von Linus Torvalds in den USA und/oder anderen Ländern.

Microsoft, Windows, Windows NT und das Windows-Logo sind Marken von Microsoft Corporation in den USA und/oder anderen Ländern.

UNIX ist eine eingetragene Marke von The Open Group in den USA und anderen Ländern.

Java und alle auf Java basierenden Marken und Logos sind Marken oder eingetragene Marken der Oracle Corporation und/oder ihrer verbundenen Unternehmen.

Bedingungen für Produktdokumentation

Die Berechtigungen zur Nutzung dieser Veröffentlichungen werden Ihnen auf der Basis der folgenden Bedingungen gewährt.

Anwendbarkeit

Diese Bedingungen sind eine Ergänzung der Nutzungsbedingungen auf der IBM Website.

Persönliche Nutzung

Sie dürfen diese Veröffentlichungen für Ihre persönliche, nicht kommerzielle Nutzung unter der Voraussetzung vervielfältigen, dass alle Eigentumsvermerke erhalten bleiben. Sie dürfen diese Veröffentlichungen oder Teile der Veröffentlichungen ohne ausdrückliche Genehmigung von IBM weder weitergeben oder anzeigen noch abgeleitete Werke davon erstellen.

Kommerzielle Nutzung

Sie dürfen diese Veröffentlichungen nur innerhalb Ihres Unternehmens und unter der Voraussetzung, dass alle Eigentumsvermerke erhalten bleiben, vervielfältigen, weitergeben und anzeigen. Sie dürfen diese Veröffentlichungen oder Teile der Veröffentlichungen ohne ausdrückliche Genehmigung von IBM außerhalb Ihres Unternehmens weder vervielfältigen, weitergeben oder anzeigen noch abgeleitete Werke davon erstellen.

Berechtigungen

Abgesehen von den hier gewährten Berechtigungen werden keine weiteren Berechtigungen, Lizenzen oder Rechte (veröffentlicht oder stillschweigend) in Bezug auf die Veröffentlichungen oder darin enthaltene Informationen, Daten, Software oder geistiges Eigentum gewährt.

IBM behält sich das Recht vor, die hierin gewährten Berechtigungen nach eigenem Ermessen zurückzuziehen, wenn sich die Nutzung der Veröffentlichungen für IBM als nachteilig erweist oder wenn die obigen Nutzungsbestimmungen nicht genau befolgt werden.

Sie dürfen diese Informationen nur in Übereinstimmung mit allen anwendbaren Gesetzen und Vorschriften, einschließlich aller US-amerikanischen Exportgesetze und Verordnungen, herunterladen und exportieren.

IBM übernimmt keine Gewährleistung für den Inhalt dieser Veröffentlichungen. Diese Veröffentlichungen werden auf der Grundlage des gegenwärtigen Zustands (auf "as-is"-Basis) und ohne eine ausdrückliche oder stillschweigende Gewährleistung für die Handelsüblichkeit, die Verwendungsfähigkeit für einen bestimmten Zweck oder die Freiheit von Rechten Dritter zur Verfügung gestellt.

Index

Sonderzeichen

@BLANK (Funktion) 138, 162, 188
@DIFF (Funktion) 182
@FIELD (Funktion) 138, 189
@FIELDS_BETWEEN (Funktion) 138, 145, 189
@FIELDS_MATCHING (Funktion) 138, 145, 189
@INDEX (Funktion) 182
@LAST_NON_BLANK (Funktion) 182, 188
@MAX (Funktion) 182
@MEAN (Funktion) 182
@MIN (Funktion) 182
@MULTI_RESPONSE_SET (Funktion) 146, 189
@NULL (Funktion) 138, 162, 188
@OFFSET (Funktion) 182
 Erläuterungen zur Leistung 241
@PARTITION_FIELD (Funktion) 189
@PREDICTED (Funktion) 189
@SDEV (Funktion) 182
@SINCE (Funktion) 182
@SUM (Funktion) 182
@TARGET (Funktion) 189
@TESTING_PARTITION (Funktion) 189
@THIS (Funktion) 182
@TODAY (Funktion) 177
@TRAINING_PARTITION (Funktion) 189
@VALIDATION_PARTITION (Funktion) 189

Numerische Stichwörter

3-D-Balkendiagramme 81
3-D-Diagramme 92
3-D-Streudiagramm 76
508-Kompatibilität 243

A

abs (Funktion) 167
Abwechselnde Zeilenfarben
 Pivot-Tabellen 124
Aggregatknotten
 Leistung 239
Aktualisieren von Modellen 208
Aktualisierung
 Quellenknoten 44
allbutfirst (Funktion) 172
allbutlast (Funktion) 172
alphabefore (Funktion) 172
and (Operator) 166
Anmeldung bei IBM SPSS Modeler Server 10
Anmerkungen
 in Kommentare umwandeln 64

Anmerkungen (*Forts.*)
 Knoten 60, 64
 Ordner 219
 Projekt 218
 Streams 60, 64
Anwendungen 25
Anwendungsbeispiele 3
Anwendungsgebiete des Data-Mining 26
Anzeigeformate
 Dezimalstellen 47
 georäumliche Koordinaten 51
 Symbol für Zifferngruppierung 47
 Währung 47
 Wissenschaftlich 47
 Zahlen 47
Anzeigereihenfolge 118
arccos (Funktion) 168
arccosh (Funktion) 168
arcsin (Funktion) 168
arcsinh (Funktion) 168
arctan (Funktion) 168
arctan2 (Funktion) 168
arctanh (Funktion) 168
area (Funktion) 169
Attribut 25
Auflisten aller Kommentare für einen Stream 63
Aufteilen von Tabellen 129
 Festlegen von Tabellenumbrüchen 129
Aus früheren Versionen übernommene Tabellen 130
Ausblenden 101, 102, 121, 122
 Dimensionsbeschriftungen 122
 Fußnoten 126
 Prozedurergebnisse 102
 Titel 122
 Titelzeilen 126
 Zeilen und Spalten 121
Ausdrücke 155
Ausführen von Streams 59
Ausführung anhalten 18
Ausführungszeiten anzeigen 52
Ausgabe 16, 101, 102, 105, 107, 116, 131
 Ändern der Ausgabesprache 120
 ausblenden 101
 Ausrichtung 102, 131
 einblenden 101
 Einfügen in andere Anwendungen 105
 Export 107
 interaktiv 106
 kopieren 102
 Löschung 102
 speichern 116
 verschieben 102
 verschlüsseln 116
 Viewer 101
 zentrieren 102, 131

Ausgabedateien
 speichern 66
Ausgabeknoten 38
Ausgabeobjekte
 in IBM SPSS Collaboration and Deployment Services Repository speichern 197
Ausrichtung 102, 131
 Ausgabe 102, 131
Ausschneiden 18
Auswahl georäumlicher Koordinatensysteme 51
Auswahlmethoden 128
 Auswählen von Zeilen und Spalten in Pivot-Tabellen 128
Automatisierung 139

B

Balkendiagramme 81
Baumbasierte Analyse
 typische Anwendungen 25
Bedingungen 142
Befehlszeile
 IBM SPSS Modeler starten 9
Begrüßungsdialogfeld 226
Beispiele
 Anwendungshandbuch 3
 Übersicht 4
Benachrichtigungen
 Festlegen von Optionen 225
Benennen von Knoten und Streams 64
Benutzer-ID
 IBM SPSS Modeler Server 10
Benutzerdefinierte Diagramme 97
Benutzerdefinierte Funktionen (UDFs) 148
Benutzerdefinierte Palettenerstellung 233
 Unterpalettenerstellung 235
Benutzeroptionen 224
Berechtigungsnachweise
 für IBM SPSS Collaboration and Deployment Services Repository 193
Bereitstellung 192
Bereitstellungsoptionen 205
Bereitstellungstyp 205
Berichte
 Eigenschaften festlegen 220
 Generierung 220
 Speichern der Ausgabe 66
 zu Projekten hinzufügen 216
Beschriftungen 119
 anzeigen 44
 Einfügen von Gruppenbeschriftungen 119
 Löschung 119
 Variablen 212
 Wert 212

Beschriftungen, Objekt von IBM SPSS
 Collaboration and Deployment Services
 Repository 204
 Bevölkerungspyramiden 79
 Beziehungsdiagramme 83
 Bitweise Funktionen 170
 Blasendiagramme 85
 Blättern
 Festlegen von Optionen 49
 BMP-Dateien 107, 114
 Exportieren von Diagrammen 107,
 114
 Boxplots 84
 Breite Tabellen
 Einfügen in Microsoft Word 105

C

Cache
 Aktivierung 224
 Cache einrichten 40
 leeren 41, 44
 speichern 41
 CACHEDATEIKNOTEN
 laden 67
 Candlestick-Diagramme 94
 cdf_chisq (Funktion) 169
 cdf_f (Funktion) 169
 cdf_normal (Funktion) 169
 cdf_t (Funktion) 169
 Champion-Challenger-Analyse 192, 205
 Chi-Quadrat-Verteilung
 Wahrscheinlichkeitsfunktionen 169
 CLEM 147
 Ausdrücke 142, 155
 Ausdrücke erstellen 147
 Ausdrücke überprüfen 151
 Beispiele 139
 Datentypen 155, 156, 157
 Einführung 24, 139
 Funktionen 148
 Sprache 155
 CLEM-Ausdrücke
 Leistung 241
 CLEM-Funktionen
 bitweise 170
 datetime 177
 fehlende Werte 138
 Globalwert 187
 Information 162
 Leerstellen und Nullen 188
 Liste der verfügbaren Funktionen 160
 logisch 166
 numerisch 167
 räumlich 169
 Sequenz 182
 Sonderfunktionen 189
 trigonometrisch 168
 Umwandlung 163
 Vergleich 164
 Wahrscheinlichkeit 169
 Zeichenfolge 172
 Zufall 171
 Client
 Standardverzeichnis 224

close_to
 räumliche Funktionen 169
 close_to (Funktion) 169
 Codierung 44, 257
 Cognos Active Report 109
 Coordinator of Processes 11
 COP 11
 cos (Funktion) 168
 cosh (Funktion) 168
 count_equal (Funktion) 145, 164
 count_greater_than (Funktion) 145, 164
 count_less_than (Funktion) 145, 164
 count_non_nulls (Funktion) 164
 count_not_equal (Funktion) 145, 164
 count_nulls (Funktion) 138, 145, 164
 count_substring (Funktion) 172
 CRISP-DM 18, 215
 Projektansicht 215
 CRISP-DM, Prozessmodell 27, 28
 crosses
 räumliche Funktionen 169
 crosses (Funktion) 169

D

Dashboard 97
 Diagramme 97
 Data Audit-Knoten
 bei Datenexploration verwenden 25
 beim Data-Mining verwenden 26
 Data-Mining 25
 Anwendungsbeispiele 35
 Strategie 27
 date_before (Funktion) 164
 Dateien 104
 Hinzufügen von Textdateien im Viewer 104
 Daten
 Vorschau 43
 Datenbank
 Funktionen 148
 Datenbankfunktionen
 benutzerdefinierte Funktionen
 (UDFs) 148
 in CLEM-Ausdrücken 148
 Datendateien
 mehrere geöffnete Datendateien 131
 Dateneditor
 deskriptive Statistiken, Ausgaben 132
 mehrere geöffnete Datendateien 131
 Datensätze 25
 fehlende Werte 134
 systemdefiniert fehlende Werte 135
 Datenstreams
 Erstellung 37
 Datentypen 141
 in Parametern 54
 Datenzuordnung 69
 Datenzuordnungstool 68
 datetime_date (Funktion) 163
 Datums-/Uhrzeitfunktionen
 datetime_date 177
 datetime_day 177
 datetime_day_name 177
 datetime_day_short_name 177
 datetime_hour 177
 datetime_in_seconds 177

Datums-/Uhrzeitfunktionen (Forts.)
 datetime_minute 177
 datetime_month 177
 datetime_month_name 177
 datetime_month_short_name 177
 datetime_now datetime_second 177
 datetime_time 177
 datetime_timestamp 177
 datetime_weekday 177
 datetime_year 177
 Datums-/Uhrzeitwerte 144
 Datumsangaben
 bearbeiten 181
 Konvertierung 181
 Datumsformate 46, 157, 158
 Datumsfunktionen 157, 158
 @TODAY (Funktion) 177
 date_before 164, 177
 date_days_difference 177
 date_in_days 177
 date_in_months 177
 date_in_weeks 177
 date_in_years 177
 date_months_difference 177
 date_weeks_difference 177
 date_years_difference 177
 Dezimalstellen
 Anzeigeformate 47
 Dezimalzeichen
 Zahlenanzeigeformate 44
 Diagramme 75, 101, 107, 130
 3D 92
 ausblenden 101
 Balken 81
 benutzerdefiniert 97
 Bevölkerungspyramide 79
 Beziehung 83
 Blase 85
 Boxplot 84
 Candlestick 94
 Dashboard 97
 Erstellen aus Pivot-Tabellen 130
 Export 107
 Fehlerbalken 81, 91
 Häufigkeitspolygon 79
 Heat-Map 88
 Histogramm 79
 Karte 86, 87
 Kreis 80
 Kreispackung 85
 Linie 77
 mehrere Zeitreihen 78
 parallel 82
 Punktdiagramm 76
 Q-Q-Diagramm 75
 Spalte 81
 Speichern der Ausgabe 66
 Streudiagramm 76
 t-SNE 90
 ThemeRiver 95
 Übersichtspunktdiagramm 76
 Verbundlinie 76
 Vorlagen 98
 Wortwolke 90
 Zeitdiagramme 96
 zu Projekten hinzufügen 216
 Diagramme mit mehreren Zeitreihen 78

Diagrammerstellung
 Begriffe 73
 Galerie 74
 Layout 73
 Dialogfelder 131
 Anzeigen von Variablenbeschriftungen 131
 Anzeigen von Variablenamen 131
 Anzeigereihenfolge für Variablen 131
 DIFF (Funktion) 182
 distance
 räumliche Funktionen 169
 distance (Funktion) 169
 div (Funktion) 167
 Dokumentation 3
 Domänenname (Windows)
 IBM SPSS Modeler Server 10
 Drehen von Beschriftungen 119
 Druck 23, 114, 115, 116, 123, 125, 129
 Abstand zwischen Ausgabeelementen 116
 Diagramme 114
 Diagrammgröße 116
 Festlegen von Tabellennummern 129
 Kopf- und Fußzeilen 115
 Pivot-Tabellen 114
 Schichten 114, 123, 125
 Seitenansicht 115
 Seitennummern 116
 Skalieren von Tabellen 123, 125
 Streams 21, 40
 Textausgabe 114
 DTD 212
 Duplikatknotten
 Leistung 239

E

Eigenschaften 123
 Berichtphasen 220
 für Datenstreams 44
 Pivot-Tabellen 123
 Projektordner 219
 Tabellen 123
 Einblenden 101, 121, 122
 Dimensionsbeschriftungen 122
 Ergebnisse 101
 Fußnoten 126
 Titel 122
 Titelzeilen 126
 Zeilen oder Spalten 121
 Einfügen 18
 Einfügen von Ausgaben in andere Anwendungen 105
 Einfügen von Gruppenbeschriftungen 119
 Einführung 155
 IBM SPSS Modeler 9, 223
 Eingabehilfen 243, 255
 Beispiel 251, 252
 Funktionen in IBM SPSS Modeler 243
 Tipps in IBM SPSS Modeler 254
 Endknotten 38
 endstring (Funktion) 172

Entfernen von Gruppenbeschriftungen 119
 Entscheidungsbäume
 Eingabehilfen 253
 EPS-Dateien 107, 114
 Exportieren von Diagrammen 107, 114
 Erforderliche Felder 70
 Ersetzen von Modellen 225
 Erstellungsbereich 14
 Evaluierungsknoten
 Leistung 239
 Excel-Format
 Exportieren von Ausgaben 107, 110
 exponential (Funktion) 167
 Export
 PMML 212, 213
 Streambeschreibungen 59
 Exportieren von Ausgaben 107, 111, 112
 Excel-Format 107, 110
 HTML 108
 HTML-Format 107
 PDF-Format 107, 112
 PowerPoint-Format 107
 Webbericht 109
 Word-Format 107, 109
 Exportieren von Diagrammen 107, 113, 114
 Expression Builder 248
 Übersicht 147
 verwenden 147
 Zugriff 147

F

f-Verteilung
 Wahrscheinlichkeitsfunktionen 169
 Faktor 253
 Fall 25
 Farben
 festlegen 226
 Farben in Pivot-Tabellen 125
 Rahmen 125
 Fehlende Werte 134, 144
 CLEM-Ausdrücke 138
 in Datensätzen 134
 System 135
 Fehler beim Ansichtsrendering
 nicht genug Speicher 223
 Fehlerbalkendiagramme 81, 91
 Fehlernachrichten 52
 Felder 25, 155, 157
 in CLEM-Ausdrücken 151
 Werte anzeigen 151
 Feldzuordnungen aufheben 68
 Festlegen der anzuzeigenden Zeilenanzahl 123
 first_index (Funktion) 146, 164
 first_non_null (Funktion) 146, 164
 first_non_null_index (Funktion) 146, 164
 Fläche
 räumliche Funktionen 169
 Formate georäumlicher Koordinaten 51
 Fortsetzungstext 125
 für Pivot-Tabellen 125
 fracof (Funktion) 167

Füllerknotten
 fehlende Werte 137
 Funktionen 157, 158, 161, 162, 182
 @BLANK 137
 @FIELD 147, 189
 @GLOBAL_MAX 187
 @GLOBAL_MEAN 187
 @GLOBAL_MIN 187
 @GLOBAL_SDEV 187
 @GLOBAL_SUM 187
 @PARTITION 189
 @PREDICTED 147, 189
 @TARGET 147, 189
 Beispiele 139
 benutzerdefinierte Funktionen (UDFs) 148
 Datenbank 148
 in CLEM-Ausdrücken 148
 Umgang mit fehlenden Werten 138
 Fußnoten 124, 126, 127, 128
 Markierungen 124
 neu nummerieren 127
 Fußzeilen 115

G

Ganze Zahlen 155
 Generierte Modelle (Palette) 16
 Georäumliche Koordinaten
 Anzeigeformat 51
 Systeme auswählen 51
 Gestapelte Balkendiagramme 81
 gleich (Operator) 164
 Gliederung 102, 103
 Ändern von Ebenen 103
 erweitern 103
 im Viewer 102
 reduzieren 103
 Globale Funktionen 187
 Globale Werte
 in CLEM-Ausdrücken 151
 Grad
 Maßeinheiten 47
 Größe ändern 20
 Größen 103
 in der Gliederung 103
 größer als (Operator) 164
 Gruppenbeschriftungen 119
 Gruppieren von Zeilen oder Spalten 119
 Gruppierte Liniendiagramme 77
 Gruppierte Streudiagramme 76
 Gruppiertes Balkendiagramm 81

H

hasendstring (Funktion) 172
 hasmidstring (Funktion) 172
 hasstartstring (Funktion) 172
 hassubstring (Funktion) 172
 Häufigkeitspolygone 79
 Hauptfenster 14
 Heat-Map-Diagramme 88
 Hintergrundfarbe 126
 Hinweise
 allgemeine Verwendung 70

- Hinzufügen
 - zu einem Projekt 216
- Hinzufügen von Gruppenbeschriftungen 119
- Histogramme 79
- Hostname
 - IBM SPSS Modeler Server 10, 11
- Hotkeys 22
- HTML 107, 108
 - Exportieren von Ausgaben 107, 108
- HTML-Ausgabe
 - Sprachausgabeprogramm 253

I

- IBM SPSS Analytic Server
 - mehrere Verbindungen 12
 - Verbindung 12
- IBM SPSS Collaboration and Deployment Services 192
- IBM SPSS Collaboration and Deployment Services Repository 191, 192
 - Berechtigungsanzeige 193
 - durchsuchen 193
 - Objekte abrufen 198
 - Objekte speichern 193
 - Objekte sperren und entsperren 201
 - Objekte und Versionen löschen 201
 - Objekteigenschaften 203
 - Ordner 201, 202
 - Projekte übertragen 218
 - Single Sign-on 192
 - Suche in 199
 - Verbindung herstellen zu 192, 193
- IBM SPSS Modeler 1, 14
 - Dokumentation 3
 - Eingabehilfen 243
 - erste Schritte 9
 - Optionen 223
 - Tipps und Tastenkombinationen 70
 - über Befehlszeile ausführen 9
 - Übersicht 9, 223
- IBM SPSS Modeler Advantage 192
- IBM SPSS Modeler Server 1
 - Benutzer-ID 10
 - Domänenname (Windows) 10
 - Hostname 10, 11
 - Kennwort 10
 - Portnummer 10, 11
- IBM SPSS Modeler Server-Verbindungen hinzufügen 11
- if, then, else (Funktionen) 166
- Import
 - PMML 212, 213
- INDEX (Funktion) 182
- Informationsfunktionen 162
- Inhalte kopieren 105
- integer_bitcount (Funktion) 170
- integer_leastbit (Funktion) 170
- integer_length (Funktion) 170
- Interaktive Ausgabe 106
- Interaktiver Baum (Fenster)
 - Eingabehilfen 253
- intof (Funktion) 167
- is_date (Funktion) 162
- is_datetime (Funktion) 162
- is_integer (Funktion) 162

- is_number (Funktion) 162
- is_real (Funktion) 162
- is_string (Funktion) 162
- is_time (Funktion) 162
- is_timestamp (Funktion) 162
- isalphacode (Funktion) 172
- isendstring (Funktion) 172
- islowercode (Funktion) 172
- ismidstring (Funktion) 172
- isnumbercode (Funktion) 172
- isstartstring (Funktion) 172
- issubstring (Funktion) 172
- issubstring_count (Funktion) 172
- issubstring_lim (Funktion) 172
- isuppercode (Funktion) 172

J

- Java 255
- JAWS 243, 253, 254, 255
- JPEG-Dateien 107, 113
 - Exportieren von Diagrammen 107, 113

K

- K-Means-Knoten
 - große Sets 44
 - Leistung 240
- Kartenservice 87
- Kennwort
 - IBM SPSS Analytic Server 12
 - IBM SPSS Modeler Server 10
- Klassen 18, 215, 216
- Klassierungsknoten
 - Leistung 239
- kleiner als (Operator) 164
- Knoten 9
 - Aktivierung 40
 - Anmerkungen 60, 64
 - aus Palette entfernen 234
 - Ausführungszeiten 52
 - bearbeiten 40
 - benutzerdefinierte Palettenerstellung 233
 - benutzerdefinierte Unterpalettenerstellung 235
 - Datenvorschau 43
 - duplizieren 40
 - Einführung 38
 - Festlegen von Optionen 40
 - hinzufügen 38, 40
 - in einem Stream inaktivieren 40
 - in einem Stream umgehen 39
 - in einem Stream verbinden 38
 - in IBM SPSS Collaboration and Deployment Services Repository speichern 197
 - inaktivieren 40
 - Kommentare hinzufügen zu 60
 - laden 67
 - Leistung 239, 240
 - Löschung 38
 - Palette anzeigen 234
 - Palettenregisterkarte, Anpassung 235
 - Reihenfolge 237

- Knoten (*Forts.*)
 - speichern 66
 - sperren 43
 - suchen 56
 - Verbindungen löschen 40
 - Vorschau von Daten 43
 - zu Projekten hinzufügen 216, 217
- Knoten aktivieren 40
- Knoten inaktivieren 40
- Knotenennamen 64
- Knotenpalettenauswahl 234
- Knowledge Discovery 25
- Kohonen-Knoten
 - große Sets 44
 - Leistung 240
- Komma 44
- Kommentare
 - alle in einem Stream auflisten 63
 - Tastenkombinationen 249, 250
 - zu Knoten und Streams 60
- Konventionen 161
- Konvertierungsfunktionen 163
- Kopfzeilen 115
- Kopieren 18
- Kopieren und Einfügen von Ausgaben in andere Anwendungen 105
- Kreisdiagramme 80
- Kreispackungsdiagramme 85

L

- Laden
 - Knoten 67
 - Statusangaben 67
- Ländereinstellung
 - Optionen 223
- Landkartendiagramme 86
- last_index (Funktion) 146, 164
- LAST_NON_BLANK (Funktion) 182
- last_non_null (Funktion) 146, 164
- last_non_null_index (Funktion) 146, 164
- Laufzeit, Eingabeaufforderungen während 54
- Leere Bereiche
 - aus Zeichenfolgen entfernen 143, 172
- Leerstellenbehandlung
 - CLEM-Funktionen 188
- Leerzeichen 134, 144
 - aus Zeichenfolgen entfernen 143, 172
- Leistung
 - CLEM-Ausdrücke 241
 - Modellierungsknoten 240
 - von Prozessknoten 239
- length (Funktion) 172
- Lineare Regression
 - als PMML exportieren 228
- Liniendiagramme 77
 - Verbundlinie 76, 93
 - Zeitdiagramme 96
- Listen 155, 156
- locchar (Funktion) 172
- locchar_back (Funktion) 172
- log (Funktion) 167
- log10 (Funktion) 167
- Logische Funktionen 166
- Logistische Regression 253
 - als PMML exportieren 228

Löschen von Ausgaben 102
lowertoupper (Funktion) 172

M

Manager 16
Maschinenlernen 25
Maßeinheit 131
matches (Funktion) 172
Matrixstreudiagramme 76, 93
Maus
 in IBM SPSS Modeler verwenden 22, 38
max (Funktion) 164
MAX (Funktion) 182
max_index (Funktion) 146, 164
max_n (Funktion) 145, 164
MEAN (Funktion) 182
mean_n (Funktion) 145, 167
Mehrere geöffnete Datendateien 131
Mehrere IBM SPSS Modeler-Sitzungen 13
Mehrfachantwortsets
 in CLEM-Ausdrücken 146, 151
Mehrfachliniendiagramme 77
member (Funktion) 164
Merkmalauswahlknoten
 fehlende Werte 134
Metadateien 107
 Exportieren von Diagrammen 107
min (Funktion) 164
MIN (Funktion) 182
min_index (Funktion) 146, 164
min_n (Funktion) 145, 164
Minimierung 20
Mittlere Maustaste
 Simulierung 22, 38
mod (Funktion) 167
Modellaktualisierung 205
Modelle 59
 Aktualisierung 208
 Ersetzung 225
 Export 228
 in IBM SPSS Collaboration and Deployment Services Repository speichern 197
 zu Projekten hinzufügen 216
Modellierung
 Verzweigung 59
Modellierungsknoten 38
 Leistung 240
 Modellierungs-Palettenregisterkarte, Anpassung 235
Modellnuggets 59

N

Nachrichten
 generiertes SQL anzeigen 49
Navigieren
 Tastenkombinationen 245
negate (Funktion) 167
Neue Funktionen 7
Neuronaler Netzknoten
 Leistung 240

Neuronales Netz, Knoten
 große Sets 44
Nicht genug Speicher 223
Normalverteilung
 Wahrscheinlichkeitsfunktionen 169
not (Operator) 166
Nuggets 59
 definiert 16
Nullen 144
num_points
 räumliche Funktionen 169
num_points (Funktion) 169
Numerische Funktionen 167
Nutzer
 IBM SPSS Analytic Server 12

O

Objekte
 Eigenschaften 219
Objekte aus IBM SPSS Collaboration and Deployment Services Repository abrufen 198
Objekte in IBM SPSS Collaboration and Deployment Services Repository speichern 193
Objekte von IBM SPSS Collaboration and Deployment Services Repository entsperren 201
Objekte von IBM SPSS Collaboration and Deployment Services Repository sperren 201
Objekteigenschaften, IBM SPSS Collaboration and Deployment Services Repository 203
Öffnen
 Ausgabe 67
 Knoten 67
 Modelle 67
 Projekte 216
 Statusangaben 67
 Streams 67
OFFSET (Funktion) 182
oneof (Funktion) 171
Operatoren
 in CLEM-Ausdrücken 148
 Zeichenfolgen verbinden 163
Operatorenrangfolge 158
Optionen 131, 132, 223
 allgemein 131
 anzeigen 226
 Benutzer 224
 Beschriftung der Ausgabe 132
 deskriptive Statistiken im Dateneditor 132
 für IBM SPSS Modeler 223
 PMML 228
 Streameigenschaften 44, 46, 47, 49, 51, 52
 Syntax 227
 Viewer 131
 Vorlage für Pivot-Tabellen 131
or (Operator) 166
Ordner, IBM SPSS Collaboration and Deployment Services Repository 201, 202
overlap
 räumliche Funktionen 169

overlap (Funktion) 169

P

Palette der Modelle 197
Paletten 14
 anpassen 233
Palettenregisterkarte, Anpassung 235
Paralleldiagramme 82
Parameter
 Eingabeaufforderungen während der Laufzeit 54
 in CLEM-Ausdrücken 151
 Modellerstellung 206
 Scoring 206
 Sitzung 54
 Stream 54
 Typ 54
PDF
 Exportieren von Ausgaben 107, 112
pi (Funktion) 168
Pivot-Tabellen 101, 105, 107, 114, 118, 119, 121, 122, 123, 124, 125, 128, 129, 130, 131
 abwechselnde Zeilenfarben 124
 allgemeine Eigenschaften 123
 Ändern der Anzeigereihenfolge 118
 Ändern der Vorlage 122
 Anpassen an die Seitengröße 123, 125
 Anpassen der Standardspaltenbreite 131
 Anzeigen ausgeblendeter Rahmen 128
 Aufheben der Gruppierung von Zeilen oder Spalten 119
 aus früheren Versionen übernommene Tabellen 130
 ausblenden 101
 Ausrichtung 126
 Auswählen von Zeilen und Spalten 128
 bearbeiten 118
 Drehen von Beschriftungen 119
 Drucken umfangreicher Tabellen 129
 Drucken von Schichten 114
 Eigenschaften 123
 Ein- und Ausblenden von Zellen 121
 Einfügen als Tabellen 105
 Einfügen in andere Anwendungen 105
 Einfügen von Gruppenbeschriftungen 119
 Einfügen von Zeilen und Spalten 120
 Erstellen aus Pivot-Tabellen 130
 Exportieren als HTML 107
 Festlegen der anzuzeigenden Zeilenanzahl 123
 Festlegen von Tabellenumbrüchen 129
 Fortsetzungstext 125
 Fußnoten 126, 127, 128
 Fußnoteneigenschaften 124
 Gruppieren von Zeilen oder Spalten 119
 Hintergrundfarbe 126

- Pivot-Tabellen (*Forts.*)
 - Löschen von Gruppenbeschriftungen 119
 - pivotieren 118
 - Rahmen 125
 - Ränder 126
 - Rasterlinien 128
 - Rendern von Tabellen mit höherer Geschwindigkeit 131
 - Rückgängigmachen von Änderungen 121
 - Schichten 121
 - schnelle Pivot-Tabellen 131
 - Schriftarten 126
 - Sortieren von Zeilen 119
 - Sprache 120
 - Standardvorlage für neue Tabellen 131
 - Titelzeilen 126
 - Transponieren von Zeilen und Spalten 119
 - Variablenbeschriftungen 120
 - Verschieben von Zeilen und Spalten 118
 - Verwenden von Symbolen 118
 - Wertbeschriftungen 120
 - Zellenbreiten 128
 - Zelleneigenschaften 126
 - Zellenformate 124
 - PMML
 - Exportoptionen 228
 - Modelle exportieren 212, 213
 - Modelle importieren 212, 213
 - PMML-Modelle
 - lineare Regression 228
 - logistische Regression 228
 - PNG-Dateien 107, 114
 - Exportieren von Diagrammen 107, 114
 - Portnummer
 - IBM SPSS Modeler Server 10, 11
 - PostScript-Dateien (Encapsulated) 107, 114
 - Exportieren von Diagrammen 107, 114
 - power (exponentielle Funktion) 167
 - PowerPoint 111
 - Exportieren von Ausgaben als PowerPoint 111
 - PowerPoint-Dateien 216
 - PowerPoint-Format
 - Exportieren von Ausgaben 107
 - Produktionsmodus 131
 - Verwenden von Befehlsyntax aus der Journaldatei 131
 - Projekte 18, 215
 - Anmerkungen 218
 - Berichte generieren 220
 - CRISP-DM, Ansicht 215
 - Eigenschaften festlegen 218
 - Erstellung 216
 - in IBM SPSS Collaboration and Deployment Services Repository 218
 - in IBM SPSS Collaboration and Deployment Services Repository speichern 196
 - Klassenansicht 216
 - Projekte (*Forts.*)
 - Neuerstellung 216
 - Objekte hinzufügen 217
 - Objekteigenschaften 219
 - Ordneigenschaften 219
 - schließen 220
 - Standardordner einrichten 216
 - Protokolldateien
 - generiertes SQL anzeigen 49
 - Prozessknoten 38
 - Leistung 239
 - Punkt 44
- Q**
- Q-Q-Diagramm 75
 - Qualitätsknoten
 - fehlende Werte 134
 - Quellenknoten 38
 - Aktualisierung 44
 - Datenzuordnung 68
 - QuickInfos
 - Anmerkungen zu Knoten 64
- R**
- Radiant
 - Maßeinheiten 47
 - Rahmen 125, 128
 - Anzeigen ausgeblendeter Rahmen 128
 - random (Funktion) 171
 - random0 (Funktion) 171
 - Rangfolge 158
 - Rasterlinien 128
 - Pivot-Tabellen 128
 - Räumliche Funktionen 169
 - Reelle Zahlen 155, 156
 - Regelerstellungsknoten
 - laden 67
 - Regelsets
 - Auswertung 44
 - Regression 253
 - rem (Funktion) 167
 - replace (Funktion) 172
 - replicate (Funktion) 172
 - round (Funktion) 167
 - Rückgängig 18
- S**
- SAS-Dateien
 - Codierung 257
 - Schichten 114, 121, 123, 125
 - anzeigen 121
 - Druck 114, 123, 125
 - Erstellung 121
 - in Pivot-Tabellen 121
 - Schlüsselwörter
 - Anmerkungen zu Knoten 64
 - Schnelle Pivot-Tabellen 131
 - Schriftarten 103, 126, 226, 227
 - im Gliederungsfenster 103
 - Scoring
 - Verzweigung 59, 207, 209
 - Scriptfarben
 - festlegen 227
 - Scripts 24, 139
 - SDEV (Funktion) 182
 - sdev_n (Funktion) 145, 167
 - Seiteneinrichtung 115, 116
 - Diagrammgröße 116
 - Kopf- und Fußzeilen 115
 - Seitennummerierung 116
 - Sequenzfunktionen 182
 - Server
 - Anmeldung 10
 - Standardverzeichnis 224
 - Suche nach COP für Server 11
 - Verbindungen hinzufügen 11
 - Set aus dichotomen Variablen
 - in CLEM-Ausdrücken 146
 - Sets 44
 - Sets aus kategorialen Variablen
 - in CLEM-Ausdrücken 146
 - sign (Funktion) 167
 - sin (Funktion) 168
 - SINCE (Funktion) 182
 - Single Sign-on 10
 - Single Sign-on, IBM SPSS Collaboration and Deployment Services Repository 191, 192
 - sinh (Funktion) 168
 - Sitzungsparameter 54
 - Skalierung
 - Pivot-Tabellen 123, 125
 - Skalierung der Streams für Ansicht 21
 - skipchar (Funktion) 172
 - skipchar_back (Funktion) 172
 - Sonderfunktionen 189
 - Sonderzeichen
 - aus Zeichenfolgen entfernen 143
 - Sortierknoten
 - Leistung 239
 - Sortierung
 - Pivot-Tabellenzeilen 119
 - soundex (Funktion) 177
 - soundex_difference (Funktion) 177
 - Spalten 128
 - Ändern der Breite in Pivot-Tabellen 128
 - Auswählen in Pivot-Tabellen 128
 - Spaltenbreite 123, 128, 131
 - Beeinflussen der Breite bei Texten mit Zeilenumbrüchen 123
 - Einstellen der Standardbreite 131
 - Festlegen der maximalen Breite 123
 - Pivot-Tabellen 128
 - Spaltendiagramme 81
 - Speicher 131
 - Stapelüberlauffehler 223
 - verwalten 223, 224
 - Speichern
 - Ausgabeobjekte 66
 - Knoten 66
 - mehrere Objekte 66
 - Statusangaben 66
 - Streams 66
 - Speichern der Ausgabe 107, 111, 112
 - Excel-Format 107, 110
 - HTML 107, 108
 - HTML-Format 107

- Speichern der Ausgabe (*Forts.*)
 - PDF-Format 107, 112
 - PowerPoint-Format 107, 111
 - Textformat 107, 112
 - Webbericht 109
 - Word-Format 107, 109
- Speichern von Diagrammen 107, 113, 114
 - BMP-Dateien 107, 114
 - EMF-Dateien 107
 - EPS-Dateien 107, 114
 - JPEG-Dateien 107, 113
 - Metadateien 107
 - PICT-Dateien 107
 - PNG-Dateien 114
 - PostScript-Dateien 114
 - TIFF-Dateien 114
- Sperren von Knoten 43
- SPLOMs 76
- Sprachausgabeprogramme 245, 248, 253
 - Beispiel 251, 252
- Sprache
 - Ändern der Ausgabesprache 120
 - Optionen 223
- SQL-Generierung
 - Protokollierung 49
 - Vorschau 49
- sqrt (Funktion) 167
- Standard
 - Projektphase 216
- Stapelüberlauffehler 223
- Startdialogfeld 226
- startstring (Funktion) 172
- Statistikdateien
 - Codierung 257
- Statusangaben
 - laden 67
 - speichern 66
- Stream 14
- Streambeschreibungen 57, 59
- Streameigenschaften
 - Analytic Server 50
- Streamerstellungsbereich
 - Einstellungen 49
- Streamnamen 64
- Streamparameter 54
- Streams 9, 204
 - Anmerkungen 60, 64
 - ausführen 59
 - Ausführungszeiten anzeigen 52
 - Bereitstellungsoptionen 205
 - Erstellung 37
 - georäumliches Koordinatensystem 51
 - in IBM SPSS Collaboration and Deployment Services Repository speichern 196
 - Knoten hinzufügen 38, 40
 - Knoten inaktivieren 40
 - Knoten umgehen 39
 - Knoten verbinden 38
 - Kommentare hinzufügen 60
 - laden 67
 - Optionen 44, 46, 47, 49, 51
 - Sicherungsdateien 66
 - Skalierung für Ansicht 21
 - speichern 66
 - Umbenennung 57, 64

- Streams (*Forts.*)
 - zu Projekten hinzufügen 216, 217
- Streamsicherungsdateien
 - wiederherstellen 66
- Streamstandardcodierung 44
- Streudiagramme 76, 93
 - 1-D 76
 - 3-D 76
 - einfache 76
 - gruppiert 76
 - Matrix 76, 93
 - Punktprogramme 76
 - Überlagerung 76
- stripchar (Funktion) 172
- strmember (Funktion) 172
- subscrs (Funktion) 172
- substring (Funktion) 172
- substring_between (Funktion) 172
- Suche
 - Knoten in einem Stream 56
- Suche nach COP für Verbindungen 11
- Suche nach Objekten in IBM SPSS Collaboration and Deployment Services Repository 199
- Suchen und Ersetzen
 - Viewer-Dokumente 104
- SUM (Funktion) 182
- sum_n (Funktion) 145, 167
- Symbol für Zifferngruppierung
 - Zahlenanzeigeformate 44
- Symbole
 - Festlegen von Optionen 21, 49
- Symbolleiste 18
- System
 - Optionen 223
- Systemdefiniert fehlende Werte
 - in Datensätzen 135

T

- t-SNE-Diagramme 90
- t-Verteilung
 - Wahrscheinlichkeitsfunktionen 169
- Tabellen 129, 248
 - Ausrichtung 126
 - Festlegen von Tabellenumbrüchen 129
 - Hintergrundfarbe 126
 - Ränder 126
 - Schriftarten 126
 - Speichern der Ausgabe 66
 - Zelleneigenschaften 126
 - zu Projekten hinzufügen 216
- Tabellendiagramm 130
- Tabellenumbrüche 129
- Tabellenvorlagen 122
 - Erstellung 122
 - zuweisen 122
- Tage übertragen 46
- tan (Funktion) 168
- tanh (Funktion) 168
- Tastenkombinationen 245, 248, 249, 250
 - allgemeine Verwendung 70
 - Tastatur 22, 245, 248, 249, 250
- Temporäres Verzeichnis 13
- testbit (Funktion) 170

- Textdateien
 - Codierung 257
- Text 103, 104, 107, 112
 - Einfügen im Viewer 103
 - Exportieren von Ausgaben als Text 107, 112
 - Hinzufügen von Textdateien im Viewer 104
- Textcodierung 44
- ThemeRiver-Diagramme 95
- THIS (Funktion) 182
- TIFF-Dateien 114
 - Exportieren von Diagrammen 107, 114
- time_before (Funktion) 164
- Tipps
 - allgemeine Verwendung 70
 - für Eingabehilfen 254
- Titel 103
 - Einfügen im Viewer 103
- Titelzeilen 126
- to_date (Funktion) 163, 177
- to_dateline (Funktion) 177
- to_datetime (Funktion) 163
- to_integer (Funktion) 163
- to_number (Funktion) 163
- to_real (Funktion) 163
- to_string (Funktion) 163
- to_time (Funktion) 163, 177
- to_timestamp (Funktion) 163, 177
- Transponieren von Zeilen und Spalten 119
- Trigonometrische Funktionen 168
- trim (Funktion) 172
- trim_start (Funktion) 172
- trimend (Funktion) 172
- Typische Anwendungen 25
- Typknoten
 - fehlende Werte 137
 - Leistung 239

U

- Überlagerte Streudiagramme 76
- Überprüfen von CLEM-Ausdrücken 151
- Übersichtspunktdiagramm 76
- Umbenennung
 - Knoten 64
 - Streams 57
- Umgekehrter Schrägstrich in CLEM-Ausdrücken 156
- Umordnen von Zeilen und Spalten 118
- undef (Funktion) 188
- ungleich (Operator) 164
- unicode_char (Funktion) 172
- Unicode-Unterstützung 257
- unicode_value (Funktion) 172
- Unterpalette
 - auf Palettenregisterkarte anzeigen 234
 - aus Palettenregisterkarte entfernen 234
 - Erstellung 235
- uppertolower (Funktion) 172
- URL
 - IBM SPSS Analytic Server 12
 - UTF-8-Codierung 44, 257

V

- value_at (Funktion) 146, 164
- Variablen 25, 131
 - Anzeigereihenfolge in Dialogfeldern 131
- Variablenbeschriftungen 120, 131, 132
 - im Gliederungsfenster 132
 - in Dialogfeldern 131
 - in Pivot-Tabellen 132
- Variablennamen 131
 - in Dialogfeldern 131
- Verbindungen
 - Server-Cluster 11
 - zu IBM SPSS Analytic Server 12
 - zu IBM SPSS Collaboration and Deployment Services Repository 192, 193
 - zu IBM SPSS Modeler Server 10, 11
- Verbundliniendiagramm 76, 93
- Vergleichsfunktionen 164
- Verketten von Zeichenfolgen 163
- Verrauschte Daten 26
- Verschieben von Zeilen und Spalten 118
- Versionsbeschriftungen, Objekt von IBM SPSS Collaboration and Deployment Services Repository 204
- Verteilungsfunktionen 169
- Vertikaler Beschriftungstext 119
- Verzeichnis
 - Standard 224
- Verzweigungen, Scoring und Modellierung 59, 207, 209
- Viewer 101, 102, 103, 116, 131, 132
 - Abstand zwischen Ausgabeelementen 116
 - Ändern der Größe von Gliederungselementen 103
 - Ändern der Schriftart in der Gliederung 103
 - Ändern von Gliederungsebenen 103
 - Anzeigen von Datenwerten 132
 - Anzeigen von Variablenbeschriftungen 132
 - Anzeigen von Variablennamen 132
 - Anzeigen von Wertbeschriftungen 132
 - Anzeigeoptionen 131
 - Ausblenden von Ergebnissen 101
 - Ergebnisfenster 101
 - Erweitern der Gliederung 103
 - Gliederung 102
 - Gliederungsfenster 101
 - Löschen von Ausgaben 102
 - Reduzieren der Gliederung 103
 - Speichern von Dokumenten 116
 - Suchen und Ersetzen von Informationen 104
 - Verschieben der Ausgabe 102
- Visuelle Programmierung 14
- Vorlagen 68, 98
 - Diagramme 98
- Vorlagenfelder 70
- Vorschau
 - Knotendaten 43

W

- Wahrscheinlichkeitsfunktionen 169
- Währungsanzeigeformat 47
- Warnungen 52
 - Festlegen von Optionen 225
- Webbericht 109
 - Exportieren von Ausgaben 109
- Wertbeschriftungen 120, 132
 - im Gliederungsfenster 132
 - in Pivot-Tabellen 132
- Werte 141
 - aus einem Data Audit anzeigen 151
 - zu CLEM-Ausdrücken hinzufügen 151
- Wichtige Felder 68, 69
- Wissenschaftliche Notation 131
 - Anzeigeformat 47
 - Unterdrücken in der Ausgabe 131
- within
 - räumliche Funktionen 169
- within (Funktion) 169
- Word-Format
 - breite Tabellen 107
 - Exportieren von Ausgaben 107, 109
- Wörterbuchdatei 253
- Wortwolkenendiagramme 90

Z

- Zahlen 144, 155, 156
- Zahlenanzeigeformate 47
- Zeichen 155, 156
- Zeichenfolgefunktionen 172
- Zeichenfolgen 155, 156
 - abgleichen 143
 - Ersetzung 143
 - in CLEM-Ausdrücken ändern 143
- Zeilen 128
 - Auswählen in Pivot-Tabellen 128
- Zeilenumbrüche 123
 - Beeinflussen der Spaltenbreite bei Texten mit Zeilenumbrüchen 123
- Zeit- und Datumsfunktionen 157, 158
- Zeitdiagramme 96
- Zeitfelder
 - Konvertierung 181
- Zeitformate 46, 157, 158
- Zeitfunktionen 157, 158
 - time_before 164, 177
 - time_hours_difference 177
 - time_in_hours 177
 - time_in_mins 177
 - time_in_secs 177
 - time_mins_difference 177
 - time_secs_difference 177
- Zellen in Pivot-Tabellen 121, 124, 128
 - ausblenden 121
 - Auswahl 128
 - Breiten 128
 - einblenden 121
 - Formate 124
- Zelleneigenschaften 126
- Zentrieren der Ausgabe 102, 131
- Zoomen 18
- Zuordnen von Feldern 68

Zusammenführungsknoten
Leistung 239



Gedruckt in Deutschland