

# IBM SPSS Custom Tables 20



*Hinweis:* Lesen Sie zunächst die allgemeinen Informationen unter Hinweise auf S. 212, bevor Sie dieses Informationsmaterial sowie das zugehörige Produkt verwenden.

Diese Ausgabe bezieht sich auf IBM® SPSS® Statistics 20 und alle nachfolgenden Versionen sowie Anpassungen, sofern dies in neuen Ausgaben nicht anders angegeben ist.

Screenshots von Adobe-Produkten werden mit Genehmigung von Adobe Systems Incorporated abgedruckt.

Screenshots von Microsoft-Produkten werden mit Genehmigung der Microsoft Corporation abgedruckt.

Lizenziertes Material - Eigentum von IBM

© **Copyright IBM Corporation 1989, 2011.**

Eingeschränkte Rechte für Benutzer der US-Regierung: Verwendung, Vervielfältigung und Veröffentlichung eingeschränkt durch GSA ADP Schedule Contract mit der IBM Corp.

---

# Vorwort

IBM® SPSS® Statistics ist ein umfassendes System zum Analysieren von Daten. Das optionale Zusatzmodul Custom Tables bietet die zusätzlichen Analyseverfahren, die in diesem Handbuch beschrieben sind. Die Prozeduren im Zusatzmodul Custom Tables müssen zusammen mit SPSS Statistics Core verwendet werden. Sie sind vollständig in dieses System integriert.

## **Informationen zu IBM Business Analytics**

Die Software IBM Business Analytics liefert umfassende, einheitliche und korrekte Informationen, mit denen Entscheidungsträger die Unternehmensleistung verbessern können. Ein umfassendes Portfolio aus [Business Intelligence](#), [Vorhersageanalyse](#), [Finanz- und Strategiemangement](#) sowie [Analyseanwendungen](#) bietet Ihnen sofort klare und umsetzbare Einblicke in die aktuelle Leistung und gibt Ihnen die Möglichkeit, zukünftige Ergebnisse vorherzusagen. Durch umfassende Branchenlösungen, bewährte Vorgehensweisen und professionellen Service können Unternehmen jeder Größe die Produktivität maximieren, Entscheidungen automatisieren und bessere Ergebnisse erzielen.

Als Teil dieses Portfolios unterstützt IBM SPSS Predictive Analytics-Software Unternehmen dabei, zukünftige Ereignisse vorherzusagen und proaktiv Maßnahmen zu ergreifen, um bessere Geschäftsergebnisse zu erzielen. Kunden aus Wirtschaft, öffentlichem Dienst und dem Bildungsbereich weltweit nutzen IBM SPSS-Technologie als Wettbewerbsvorteil für Kundengewinnung, Kundenbindung und Erhöhung der Kundenumsätze bei gleichzeitiger Eindämmung der Betrugsmöglichkeiten und Minderung von Risiken. Durch die Einbindung von IBM SPSS-Software in ihre täglichen Operationen wandeln sich Organisationen zu “Predictive Enterprises” – die Entscheidungen auf Geschäftsziele ausrichten und automatisieren und einen messbaren Wettbewerbsvorteil erzielen können. Wenn Sie weitere Informationen wünschen oder Kontakt zu einem Mitarbeiter aufnehmen möchten, besuchen Sie die Seite <http://www.ibm.com/spss>.

## **Technischer Support**

Kunden mit Wartungsvertrag können den technischen Support in Anspruch nehmen. Kunden können sich an den Technischen Support wenden, wenn sie Hilfe bei der Arbeit mit den Produkten von IBM Corp. oder bei der Installation in einer der unterstützten Hardware-Umgebungen benötigen. Zur Kontaktaufnahme mit dem technischen Support besuchen Sie die Website von IBM Corp. unter <http://www.ibm.com/support>. Wenn Sie Hilfe anfordern, halten Sie bitte Informationen bereit, um sich, Ihre Organisation und Ihren Supportvertrag zu identifizieren.

## **Technischer Support für Studenten**

Wenn Sie in der Ausbildung eine Studenten-, Bildungs- oder Grad Pack-Version eines IBM SPSS-Softwareprodukts verwenden, informieren Sie sich auf unseren speziellen Online-Seiten für Studenten zu [Lösungen für den Bildungsbereich](http://www.ibm.com/spss/rd/students/) (<http://www.ibm.com/spss/rd/students/>). Wenn Sie in der Ausbildung eine von der Bildungsstätte gestellte Version der IBM SPSS-Software verwenden, wenden Sie sich an den IBM SPSS-Produktkoordinator an Ihrer Bildungsstätte.

## ***Kundendienst***

Bei Fragen bezüglich der Lieferung oder Ihres Kundenkontos wenden Sie sich bitte an Ihre lokale Niederlassung. Halten Sie bitte stets Ihre Seriennummer bereit.

## ***Ausbildungsseminare***

IBM Corp. bietet öffentliche und unternehmensinterne Seminare an. Alle Seminare beinhalten auch praktische Übungen. Seminare finden in größeren Städten regelmäßig statt. Weitere Informationen zu diesen Seminaren finden Sie unter <http://www.ibm.com/software/analytics/spss/training>.

## ***Weitere Veröffentlichungen***

Die Handbücher *SPSS Statistics: Guide to Data Analysis*, *SPSS Statistics: Statistical Procedures Companion* und *SPSS Statistics: Advanced Statistical Procedures Companion*, die von Marija Norušis geschrieben und von Prentice Hall veröffentlicht wurden, werden als Quelle für Zusatzinformationen empfohlen. Diese Veröffentlichungen enthalten statistische Verfahren in den Modulen “Statistics Base”, “Advanced Statistics” und “Regression” von SPSS. Diese Bücher werden Sie dabei unterstützen, die Funktionen und Möglichkeiten von IBM® SPSS® Statistics optimal zu nutzen. Dabei ist es unerheblich, ob Sie ein Neuling im Bereich der Datenanalyse sind oder bereits über umfangreiche Vorkenntnisse verfügen und damit in der Lage sind, auch die erweiterten Anwendungen zu nutzen. Weitere Informationen zu den Inhalten der Veröffentlichungen sowie Auszüge aus den Kapiteln finden Sie auf der folgenden Autoren-Website: <http://www.norusis.com>

## **1 Erste Schritte mit Custom Tables** **1**

Tabellenstruktur und Terminologie .....	1
Pivot-Tabellen .....	1
Variablen und Messniveau .....	2
Zeilen, Spalten und Zellen .....	3
Stapeln .....	3
Kreuztabelle .....	3
Verschachteln .....	4
Schichten .....	4
Tabellen für Variablen mit gemeinsamen Kategorien .....	5
Mehrfachantworten-Sets .....	5
Gesamtsummen und Zwischenergebnisse .....	6
Benutzerdefinierte Auswertungsstatistiken für Gesamtsummen .....	7
Beispiel-Datendatei .....	7
Konstruieren einer Tabelle .....	8
Öffnen des Dialogfelds zum Konstruieren von benutzerdefinierten Tabellen .....	9
Auswählen von Zeilen- und Spaltenvariablen .....	10
Einfügen von Gesamtsummen und Zwischenergebnissen .....	15
Auswerten von metrischen Variablen .....	19

## **2 Benutzeroberfläche zum Konstruieren von Tabellen** **27**

Konstruieren von Tabellen .....	27
So konstruieren Sie eine Tabelle: .....	30
Stapeln von Variablen .....	31
Verschachteln von Variablen .....	32
Schichten .....	33
Einblenden und Ausblenden von Variablennamen bzw. -labels .....	34
Auswertungsstatistik .....	35
Kategorien und Gesamtsummen .....	42
Berechnete Kategorien .....	46
Tabellen für Variablen mit gemeinsamen Kategorien (Comperimeter-Tabellen) .....	48
Anpassen des Dialogfelds zum Konstruieren von Tabellen .....	49
Benutzerdefinierte Tabellen: Registerkarte "Optionen" .....	49
Benutzerdefinierte Tabellen: Registerkarte "Titel" .....	51
Benutzerdefinierte Tabellen: Registerkarte "Teststatistiken" .....	52

### **3 Einfache Tabellen für kategoriale Variablen** **56**

Eine einzelne kategoriale Variable	57
Prozente	58
Gesamt	60
Kreuztabelle	61
Prozentwerte in Kreuztabellen	62
Steuern des Anzeigeformats	63
Randsummen	64
Sortieren und Ausschließen von Kategorien	65

### **4 Stapeln, Verschachteln und Schichten von kategorialen Variablen** **70**

Stapeln von kategorialen Variablen	70
Stapeln bei Kreuztabellen	71
Verschachteln von kategorialen Variablen	73
Unterdrücken von Variablenlabels	76
Verschachtelte Kreuztabellen	78
Schichten	81
Zwei gestapelte kategoriale Schichtvariablen	83
Zwei verschachtelte kategoriale Schichtvariablen	85

### **5 Gesamtsummen und Zwischenergebnisse für kategoriale Variablen** **87**

Einfache Gesamtsumme für eine Variable	87
Die Gesamtsumme entspricht der Summe aller angezeigten Werte	88
Anzeigeposition der Gesamtsummen	90
Gesamtsummen für verschachtelte Tabellen	90
Gesamtsummen für Schichtvariablen	92
Zwischenergebnisse	94
Das Zwischenergebnis entspricht der Summe der angezeigten Werte	95
Ausblenden der Kategorien für das Zwischenergebnis	96
Zwischenergebnisse für Schichtvariablen	99

## **6 Berechnete Kategorien für kategoriale Variablen 100**

Einfache berechnete Kategorie . . . . .	100
Ausblenden von Kategorien in einer berechneten Kategorie . . . . .	103
Verweis auf Zwischenergebnisse in einer berechneten Kategorie . . . . .	104
Einsatz von berechneten Kategorien zur Anzeige nicht abschließender Zwischenergebnisse . . . . .	108

## **7 Tabellen für Variablen mit gemeinsamen Kategorien 111**

Häufigkeitstabelle. . . . .	111
Tabelle mit Prozentwerten . . . . .	113
Gesamtsummen und Kategorien . . . . .	116
Verschachteln in Tabellen mit gemeinsamen Kategorien. . . . .	118

## **8 Auswertungsstatistik 121**

Quellvariable für Auswertungsstatistik . . . . .	122
Quelle von Auswertungsstatistiken bei kategorialen Variablen. . . . .	122
Quelle von Auswertungsstatistiken für metrische Variablen. . . . .	125
Gestapelte Variablen . . . . .	127
Angepasste Auswertungsstatistik für Gesamtsummen bei kategorialen Variablen. . . . .	130
Anzeigen von Kategoriewerten . . . . .	134

## **9 Auswerten von metrischen Variablen 137**

Gestapelte metrische Variablen . . . . .	137
Mehrere Auswertungsstatistiken . . . . .	138
Anzahl, gültige N und fehlende Werte . . . . .	139
Unterschiedliche Auswertungen für unterschiedliche Variablen . . . . .	141
Gruppierte Auswertungen in Kategorien . . . . .	143
Mehrere Gruppenvariablen . . . . .	144
Verschachteln von kategorialen in metrischen Variablen . . . . .	147

## **10 Teststatistiken** **148**

Tests auf Unabhängigkeit (Chi-Quadrat) . . . . .	148
Auswirkungen von verschachtelten und gestapelten Variablen auf Tests auf Unabhängigkeit	152
Vergleichen der Spaltenmittel. . . . .	153
Auswirkungen von verschachtelten und gestapelten Variablen auf Spaltenmitteltests . . . . .	158
Vergleichen der Spaltenanteile. . . . .	160
Auswirkungen von verschachtelten und gestapelten Variablen auf Spaltenanteiltests . . . . .	166
Anmerkung zu Gewichtungen und Mehrfachantworten-Sets . . . . .	168

## **11 Mehrfachantworten-Sets** **169**

Häufigkeiten, Antworten, Prozentwerte und Gesamtwerte . . . . .	169
Verwenden von Mehrfachantworten-Sets mit anderen Variablen . . . . .	173
Statistikquellvariablen und verfügbare Auswertungsstatistiken . . . . .	175
Sets aus kategorialen Variablen und doppelte Antworten . . . . .	176
Signifikanztests bei Mehrfachantworten-Sets. . . . .	178
Test auf Unabhängigkeit bei Mehrfachantworten-Sets. . . . .	179
Vergleichen von Spaltenmitteln mit Mehrfachantworten-Sets . . . . .	181

## **12 Fehlende Werte (Define Variable)** **184**

Tabellen ohne fehlende Werte . . . . .	184
Aufnehmen von fehlenden Werten in Tabellen . . . . .	186

## **13 Formatieren und Anpassen von Tabellen** **190**

Anzeigeformat für Auswertungsstatistiken . . . . .	190
Anzeigen von Beschriftungen für Auswertungsstatistiken. . . . .	194
Spaltenbreite . . . . .	196
Anzeigewert für leere Zellen. . . . .	198
Anzeigewert für fehlende Statistiken . . . . .	198

## ***Anhänge***

***A Beispieldateien*** **201**

***B Hinweise*** **212**

***Index*** **215**



# ***Erste Schritte mit Custom Tables***

Viele Prozeduren produzieren Ausgaben in Form von Tabellen. Das Erweiterungsmodul “Custom Tables” bietet jedoch besondere Funktionen, die eine große Vielzahl von Funktionen für benutzerdefinierte Berichte unterstützen. Viele dieser Funktionen eignen sich besonders für die Analyse von Umfragen und für die Marktforschung.

In diesem Handbuch wird angenommen, dass Sie mit den Grundlagen von IBM® SPSS® Statistics vertraut sind. Wenn Sie mit den Grundlagen der Funktionsweise noch nicht vertraut sind, sollten Sie zunächst das einleitende Lernprogramm durcharbeiten, das mit dem Programm ausgeliefert wird. Wählen Sie aus der Menüleiste eines beliebigen SPSS Statistics-Fensters die folgenden Optionen aus:

Hilfe > Lernprogramm

## ***Tabellenstruktur und Terminologie***

Mit Custom Tables kann eine breite Auswahl an benutzerdefinierten Tabellen erstellt werden. Sie können viele der Möglichkeiten entdecken, indem Sie einfach mit der Benutzeroberfläche für die Tabellenkonstruktion experimentieren. Möglicherweise hilft es Ihnen jedoch, wenn Sie sich mit der grundlegenden Tabellenstruktur und den Begriffen vertraut machen, mit denen die verschiedenen strukturellen Elemente bezeichnet werden, die in Tabellen verwendet werden können.

## ***Pivot-Tabellen***

Tabellen, die aus dem Modul “Custom Tables” stammen, werden im Fenster des Viewers als **Pivot-Tabellen** angezeigt. Pivot-Tabellen geben Ihnen große Flexibilität für die Formatierung und Darstellung von Tabellen.

Detaillierte Informationen zur Arbeit mit Pivot-Tabellen finden Sie im Hilfesystem.

- ▶ Wählen Sie in einem beliebigen Fenster die folgenden Befehle aus den Menüs aus:  
Hilfe > Themen
- ▶ Doppelklicken Sie im Fenster “Inhalt” auf Core-System.
- ▶ Doppelklicken Sie anschließend im erweiterten Inhaltsverzeichnis auf Pivot-Tabellen.

## **Variablen und Messniveau**

Zu einem gewissen Grad schränkt das für eine Variable festgelegte Messniveau ein, wie Sie diese in einer Tabelle einsetzen können. In der Prozedur “Benutzerdefinierte Tabellen” wird anhand der Messniveaus zwischen zwei grundlegenden Variablentypen unterschieden:

**Kategorial.** Daten mit einer begrenzten Anzahl von eindeutigen Werten bzw. Kategorien (beispielsweise Geschlecht oder Religion). Auch als qualitative Daten bezeichnet. Bei kategorialen Variablen kann es sich um String-Variablen (alphanumerisch) oder um numerische Variablen handeln, bei denen numerische Codes zum Darstellen der Kategorien verwendet werden (beispielsweise 0 = *Weiblich* und 1 = *Männlich*). Kategoriale Variablen werden in die folgenden Untergruppen unterteilt:

- **Nominal.** Eine Variable kann als nominal behandelt werden, wenn ihre Kategorien sich nicht in eine natürliche Reihenfolge bringen lassen, z. B. die Firmenabteilung, in der eine Person arbeitet. Beispiele für nominale Variablen sind Region, Postleitzahl oder Religionszugehörigkeit.
- **Ordinal.** Eine Variable kann als ordinal behandelt werden, wenn ihre Werte für Kategorien stehen, die eine natürliche Reihenfolge aufweisen (z. B. Grad der Zufriedenheit mit Kategorien von sehr unzufrieden bis sehr zufrieden). Ordinale Variablen treten beispielsweise bei Einstellungsmessungen (Zufriedenheit oder Vertrauen) und bei Präferenzbeurteilungen auf.

Variablen, die im Daten-Editor als nominal oder ordinal festgelegt wurden, werden in der Prozedur “Benutzerdefinierte Tabellen” als kategoriale Variablen betrachtet.

**Metrisch.** Eine Variable kann als metrisch (stetig) behandelt werden, wenn ihre Werte geordnete Kategorien mit einer sinnvollen Metrik darstellen, sodass man sinnvolle Aussagen über die Abstände zwischen den Werten machen kann. Metrische Variablen sind beispielsweise Alter (in Jahren) oder Einkommen (in Geldeinheiten). Sie werden auch als quantitative Daten oder stetige Daten bezeichnet. Variablen, die im Daten-Editor als metrisch festgelegt wurden, werden in der Prozedur “Benutzerdefinierte Tabellen” als metrische Variablen betrachtet.

## **Wertelabels**

Bei kategorialen Variablen stützt sich die auf der Zeichenfläche angezeigte Vorschau der zu erstellenden Tabelle auf die definierten **Wertelabels**. Die in der Tabelle dargestellten Kategorien sind nämlich die für die entsprechende Variable definierten Wertelabels. Wenn für die Variable keine Wertelabels definiert sind, werden zwei generische Kategorien angezeigt. Die Anzahl an unterschiedlichen Werten in den Daten bestimmt, wie viele Kategorien letztlich in der Tabelle dargestellt werden. In der Vorschau wird lediglich angenommen, dass mindestens zwei Kategorien vorliegen.

Außerdem sind bei kategorialen Variablen, für die keine Wertelabels definiert wurden, einige benutzerdefinierte Funktionen zum Konstruieren von Tabellen nicht verfügbar.

## Zeilen, Spalten und Zellen

Jede Dimension einer Tabelle wird durch eine einzelne Variable oder eine Kombination von Variablen definiert. Die Variablen, die untereinander an der linken Seite der Tabelle dargestellt werden, werden als **Zeilenvariablen** bezeichnet. Durch diese Variablen werden die Zeilen einer Tabelle definiert. Die Variablen, die nebeneinander an der Oberseite der Tabelle dargestellt werden, werden als **Spaltenvariablen** bezeichnet. Durch diese Variablen werden die Spalten einer Tabelle definiert. Der Tabellenkörper besteht aus **Zellen**. Diese enthalten die Informationen, die durch die Tabelle vermittelt werden sollen, also die Häufigkeiten, Summen, Mittelwerte, Prozentwerte usw. Eine Zelle wird durch den Schnittpunkt einer Zeile und einer Spalte der Tabelle gebildet.

## Stapeln

Beim Stapeln werden separate Tabellen in dieselbe Anzeige eingefügt. So können beispielsweise Informationen über das *Geschlecht* und die *Alterskategorie* in unterschiedlichen Abschnitten derselben Tabelle angezeigt werden.

Abbildung 1-1  
Gestapelte Variablen

		Anzahl
Gender	Male	1232
	Female	1600
Age category	Less than 25	242
	25 to 34	627
	35 to 44	679
	45 to 54	481
	55 to 64	320
	65 or older	479

Obwohl durch “Stapeln” in der Regel eine vertikale Anzeige bezeichnet wird, können Variablen auch horizontal gestapelt werden.

Abbildung 1-2  
Horizontales Stapeln

Gender		Age category					
Male	Female	Less than 25	25 to 34	35 to 44	45 to 54	55 to 64	65 or older
1232	1600	242	627	679	481	320	479

## Kreuztabelle

Eine Kreuztabelle stellt ein grundlegendes Verfahren dar, mit dem die Beziehung zwischen zwei kategorialen Variablen untersucht werden kann. Wenn Sie beispielsweise *Alterskategorie* als Zeilenvariable und *Geschlecht* als Spaltenvariable verwenden, können Sie eine zweidimensionale

Kreuztabelle erstellen, mit der die Anzahl an Männern und Frauen in jeder Alterskategorie dargestellt wird.

Abbildung 1-3  
Einfache zweidimensionale Kreuztabelle

		Gender	
		Male	Female
Age category	Less than 25	108	134
	25 to 34	276	351
	35 to 44	309	370
	45 to 54	221	260
	55 to 64	136	184
	65 or older	178	301

## Verschachteln

Ähnlich wie bei Kreuztabellen kann mit Verschachtelung die Beziehung zwischen zwei kategorialen Variablen aufgezeigt werden. Bei der Verschachtelung wird eine der Variablen jedoch in derselben Dimension innerhalb der anderen Variablen verschachtelt. So können Sie beispielsweise *Geschlecht* in *Alterskategorie* in der Zeilendimension verschachteln, um die Anzahl der Männer und Frauen in jeder Alterskategorie darzustellen.

In diesem Beispiel zeigt die verschachtelte Tabelle im wesentlichen dieselben Informationen wie eine Kreuztabelle derselben zwei Variablen.

Abbildung 1-4  
Verschachtelte Variablen

		Anzahl	
Age category	Less than 25	Male	108
		Female	134
	25 to 34	Male	276
		Female	351
	35 to 44	Male	309
		Female	370
	45 to 54	Male	221
		Female	260
	55 to 64	Male	136
		Female	184
	65 or older	Male	178
		Female	301

## Schichten

Mit Schichten können Tabellen um eine Tiefendimension erweitert werden, um dreidimensionale "Würfel" zu erstellen. Schichten ähneln stark der Verschachtelung. Der wichtigste Unterschied besteht darin, dass stets nur eine Schichtenkategorie sichtbar ist. Wenn Sie beispielsweise *Alterskategorie* als Zeilenvariable und *Geschlecht* als Schichtvariable verwenden, erhalten Sie

eine Tabelle, in der die Informationen für Männer und Frauen in unterschiedlichen Schichten dargestellt werden.

Abbildung 1-5  
Variablen in Schichten

Geschlecht Weiblich		
Alterskategorie		Anzahl
Weniger als 25		134
25 bis 34		351
35 bis 44		370

Geschlecht Männlich		
Alterskategorie		Anzahl
Weniger als 25		108
25 bis 34		276
35 bis 44		309
45 bis 54		221
55 bis 64		136
65 oder älter		178

## Tabellen für Variablen mit gemeinsamen Kategorien

Oft enthalten Umfragen viele Fragen mit übereinstimmenden möglichen Antworten. Die Beispielumfrage enthält beispielsweise eine Reihe von Variablen, die sich auf das Vertrauen in verschiedene öffentliche und private Institutionen und Dienstleistungen beziehen. Für diese Variablen ist jeweils dieselbe Gruppe von Antwortkategorien verfügbar: 1 = *Sehr viel*, 2 = *Mäßig* und 3 = *Fast kein*. Sie können diese zusammenhängenden Variablen stapeln, damit sie in einer Tabelle dargestellt werden, und die gemeinsamen Antwortkategorien können als Spalten der Tabelle erscheinen.

Abbildung 1-6  
Gestapelte Variablen mit gemeinsamen Antwortkategorien in Spalten

	Sehr viel	Mäßig	Fast kein
Vertrauen in die Medizin	844	864	167
Vertrauen in Banken & Geldinstitute	490	1068	306
Vertrauen in das Fernsehen	196	936	744
Vertrauen in die Bildung	511	1055	315
Vertrauen in die Presse	176	878	808
Vertrauen in Großkonzerne	500	1078	243

## Mehrfachantworten-Sets

Mehrfachantworten-Sets verwenden mehrere Variablen, um Antworten auf Fragen aufzuzeichnen, auf welche der Befragte mehr als eine Antwort geben kann. In der Beispielumfrage wird beispielsweise die folgende Frage gestellt: "Aus welchen der folgenden Quellen beziehen Sie Nachrichten?" Die Befragten können eine beliebige Kombination der fünf möglichen Antworten auswählen: *Internet*, *Fernsehen*, *Radio*, *Zeitungen* und *Zeitschriften*. Jede dieser

Auswahlmöglichkeiten wird als separate Variable in der Datendatei gespeichert. Diese bilden zusammen ein Mehrfachantworten-Set. Mit dem Modul "Custom Tables" können Sie ein Mehrfachantworten-Set anhand dieser Variablen definieren und dieses Mehrfachantworten-Set in den von Ihnen erstellten Tabellen verwenden.

Abbildung 1-7  
Mehrfachantworten-Sets in einer Tabelle

		Anzahl	Anzahl als ...
News sources	Get news from internet	867	41.7%
	Get news from radio	551	26.5%
	Get news from television	1077	51.8%
	Get news from news magazines	294	14.1%
	Get news from newspapers	805	38.7%

In diesem Beispiel wird Ihnen auffallen, dass die Summe der Prozentwerte größer als 100 % ist. Da jeder Befragte mehr als eine Antwort auswählen darf, kann die Gesamtzahl der Antworten die Gesamtzahl der Befragten übersteigen.

## ***Gesamtsummen und Zwischenergebnisse***

Es stehen Ihnen zahlreiche Einstellungsmöglichkeiten für die Anzeige von Gesamtsummen und Zwischenergebnissen zur Verfügung. Dazu gehören:

- Gesamtsummen für Zeilen und Spalten
- Gesamtsummen für Gruppen in Tabellen mit verschachtelten und gestapelten Variablen sowie Tabellen mit Schichten
- Gesamtsummen für Untergruppen

Abbildung 1-8  
Zwischenergebnisse, Gesamtsummen für Gruppen und Gesamtsummen für Tabellen

		Anzahl	Prozent
Male	Less than 25	108	8.8%
	25 to 34	276	22.5%
	35 to 44	309	25.2%
	Subtotal < 45	693	56.4%
	45 to 54	221	18.0%
	55 to 64	136	11.1%
	65 or older	178	14.5%
	Subtotal 45+	535	43.6%
	Total	1228	100.0%
Female	Less than 25	134	8.4%
	25 to 34	351	21.9%
	35 to 44	370	23.1%
	Subtotal < 45	855	53.4%
	45 to 54	260	16.3%
	55 to 64	184	11.5%
	65 or older	301	18.8%
	Subtotal 45+	745	46.6%
	Total	1600	100.0%
Total	Less than 25	242	8.6%
	25 to 34	627	22.2%
	35 to 44	679	24.0%
	Subtotal < 45	1548	54.7%
	45 to 54	481	17.0%
	55 to 64	320	11.3%
	65 or older	479	16.9%
	Subtotal 45+	1280	45.3%
	Total	2828	100.0%

### Benutzerdefinierte Auswertungsstatistiken für Gesamtsummen

Tabellen, die Gesamtsummen oder Zwischenergebnisse enthalten, können andere Auswertungsstatistiken als die Zusammenfassungen enthalten, die für jede Kategorie angezeigt werden. So können beispielsweise Häufigkeiten für eine ordinale kategoriale Zeilenvariable und der Mittelwert für die Gesamtsummen angezeigt werden.

Abbildung 1-9  
Kategoriale Variable und Auswertungsstatistiken in derselben Dimension

Confidence in television	1 A great deal	Anzahl	196
	2 Only some	Anzahl	936
	3 Hardly any	Anzahl	744
	Gesamt	Anzahl	1876
		Median	2.29

### Beispiel-Datendatei

Für die meisten hier vorgestellten Beispiele wird die Datendatei *survey\_sample.sav* verwendet. Für weitere Informationen siehe Thema [Beispieldateien in Anhang A auf S. 201](#). Diese Datendatei stellt eine fiktive Umfrage unter mehreren Tausend Personen dar. Sie enthält grundlegende

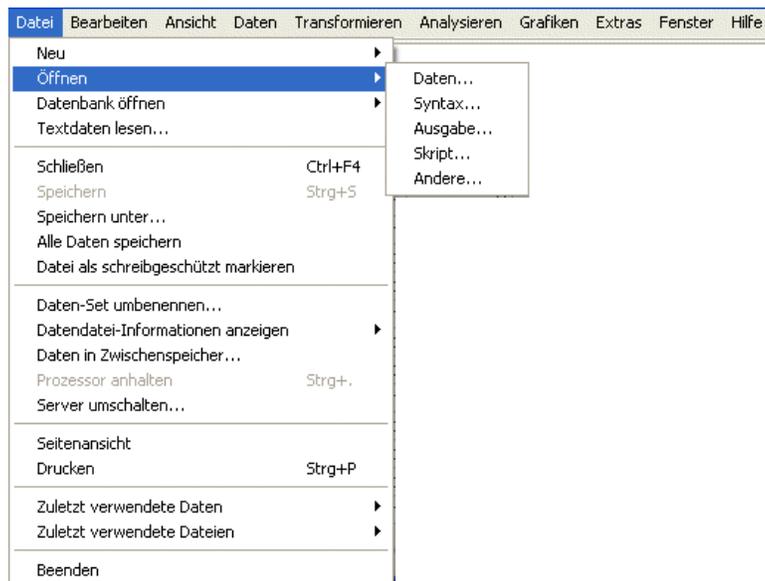
demografische Informationen und Antworten auf eine Reihe von Fragen, die diverse Themen von politischen Ansichten bis hin zu Fernsehgewohnheiten abdecken.

## Konstruieren einer Tabelle

Bevor Sie eine Tabelle konstruieren können, müssen Sie über Daten verfügen, die in der Tabelle dargestellt werden sollen.

- Wählen Sie die folgenden Befehle aus den Menüs aus:  
Datei > Öffnen > Daten...

Abbildung 1-10  
Menü "Datei"; "Öffnen"



Wahlweise können Sie auch auf der Symbolleiste auf die Schaltfläche "Datei öffnen" klicken.

Abbildung 1-11  
Schaltfläche "Datei öffnen" auf der Symbolleiste



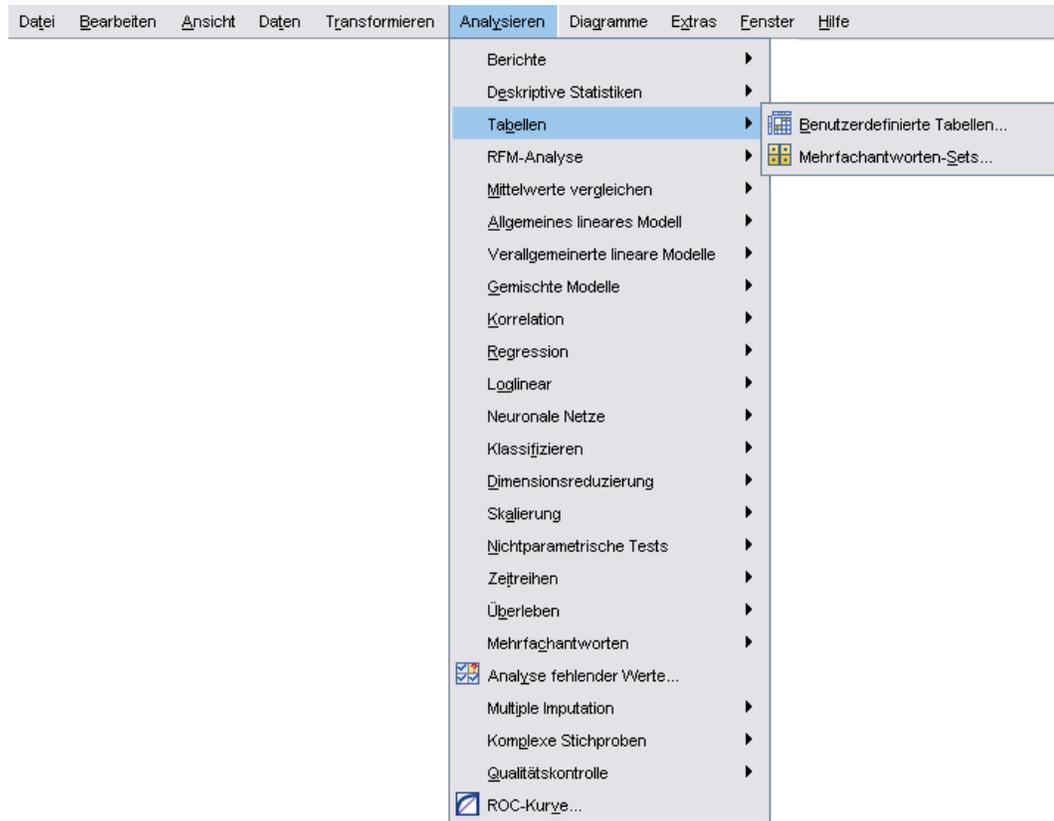
- Wenn Sie die in diesem Beispiel verwendete Datendatei öffnen möchten, lesen Sie unter [Beispieldateien](#) auf S. 201 nach, um weitere Informationen zu Speicherorten von Datendateien zu erhalten.
- Öffnen Sie die Datei *survey\_sample.sav*.

## Öffnen des Dialogfelds zum Konstruieren von benutzerdefinierten Tabellen

- Um das Dialogfeld zum Konstruieren von benutzerdefinierten Tabellen zu öffnen, wählen Sie die folgenden Befehle aus den Menüs aus:

Analysieren > Tabellen > Benutzerdefinierte Tabellen...

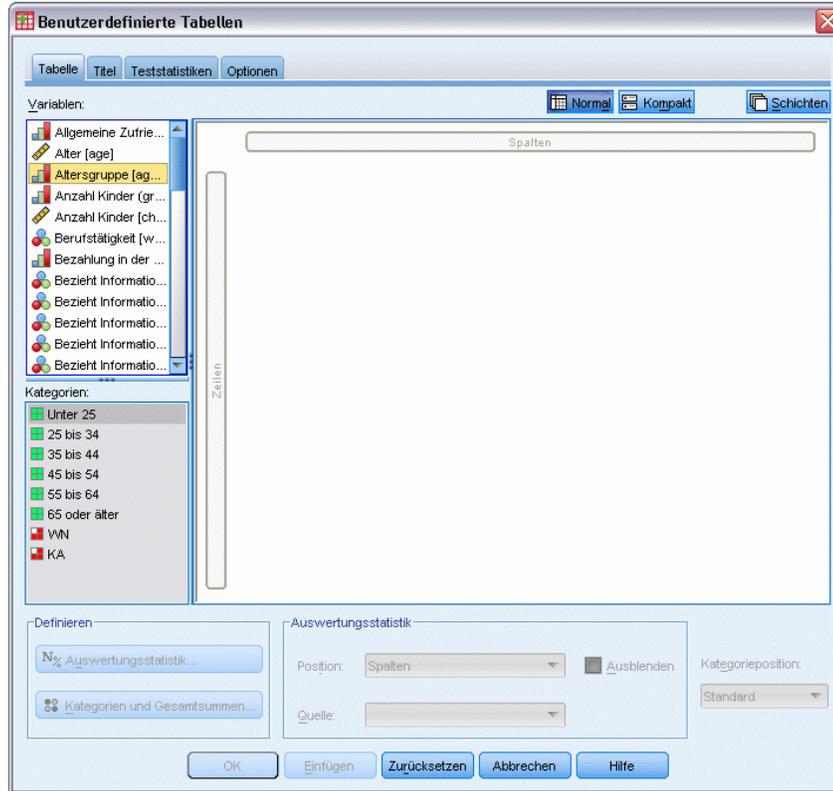
Abbildung 1-12  
Menü "Analysieren," Option "Tabellen"



Hiermit öffnen Sie das Dialogfeld zum Konstruieren von benutzerdefinierten Tabellen.

Abbildung 1-13

Dialogfeld zum Konstruieren von benutzerdefinierten Tabellen

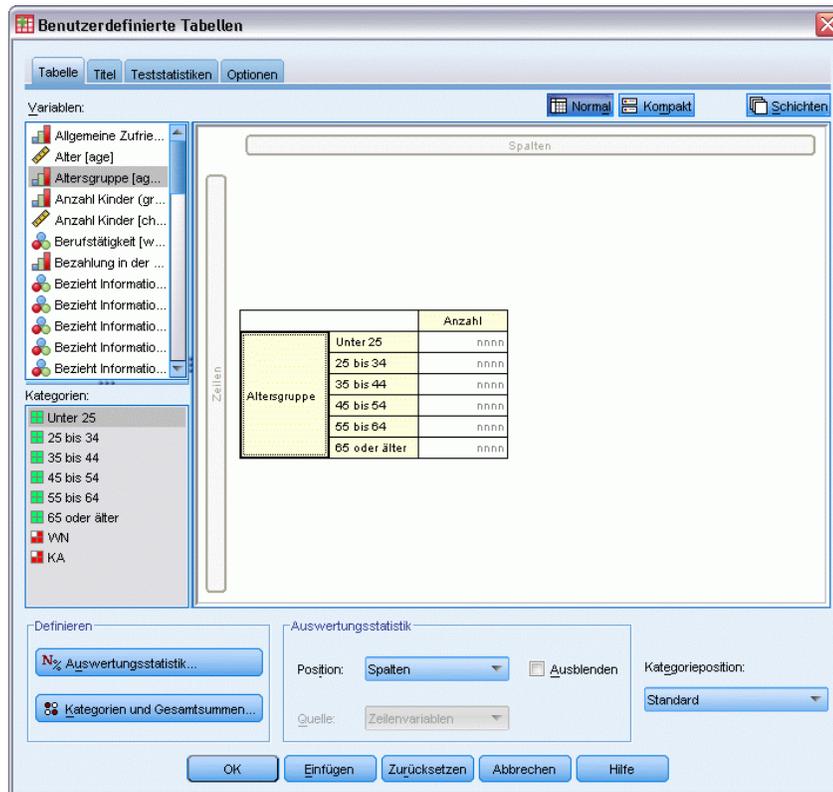


### **Auswählen von Zeilen- und Spaltenvariablen**

Um eine Tabelle zu konstruieren, ziehen Sie einfach Variablen und legen Sie diese an den Stellen ab, an denen sie in der Tabelle angezeigt werden sollen.

- ▶ Wählen Sie in der Variablenliste die Variable *Altersgruppe* aus, indem Sie darauf klicken, ziehen Sie die Variable auf die Zeichenfläche und legen Sie sie im Bereich für die Zeilen ab.

Abbildung 1-14  
Auswählen einer Zeilenvariablen

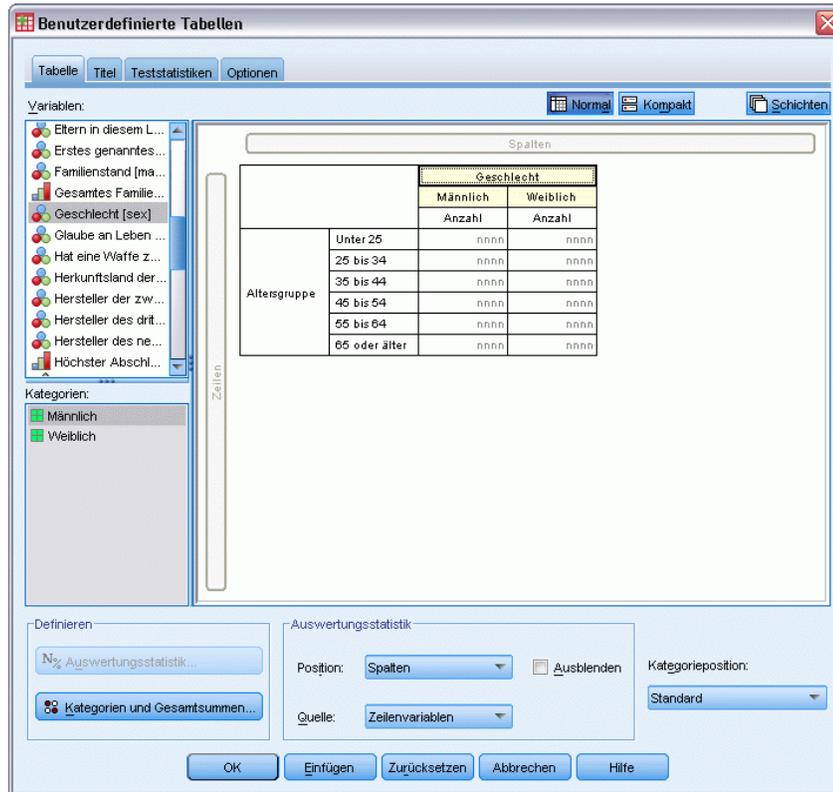


Auf der Zeichenfläche wird die Tabelle angezeigt, die auf der Grundlage dieser einzelnen Zeilenvariablen erstellt würde.

Die Vorschau enthält nicht die tatsächlichen Werte, die in der Tabelle letztlich angezeigt werden, sondern nur das grundlegende Layout der Tabelle.

- Wählen Sie in der Variablenliste die Variable *Geschlecht* aus, ziehen Sie die Variable auf die Zeichenfläche und legen Sie sie im Bereich für die Spalten ab. (Möglicherweise müssen Sie in der Variablenliste einen Bildlauf abwärts durchführen, um diese Variable zu finden.)

Abbildung 1-15  
Auswählen einer Spaltenvariablen



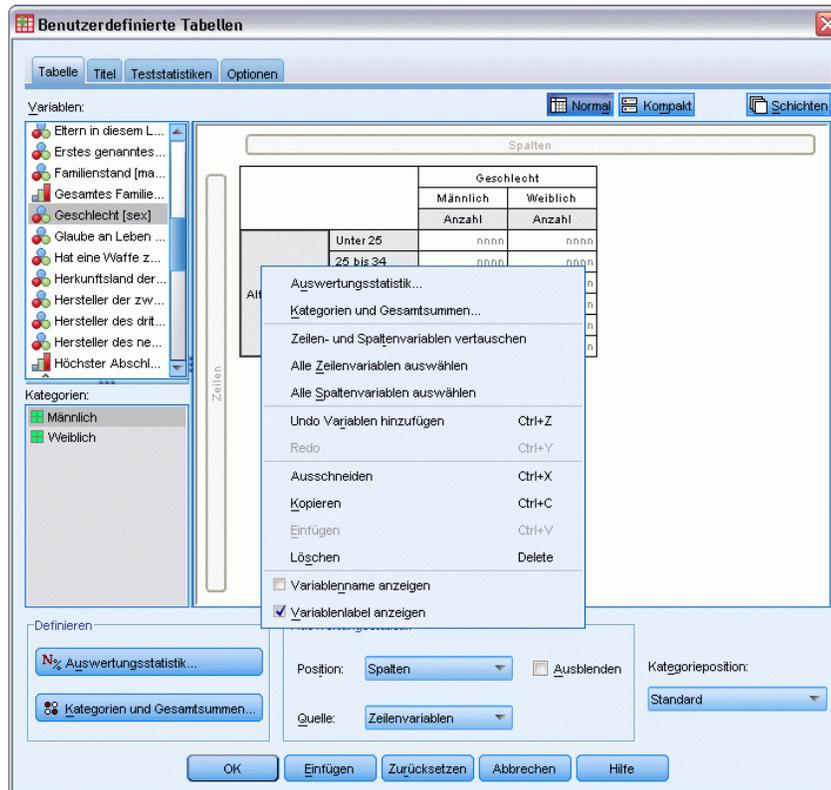
Auf der Zeichenfläche wird jetzt eine Zwei-Weg-Kreuztabelle der *Alterskategorie* nach *Geschlecht* angezeigt.

In der Standardeinstellung werden in den Zellen für kategoriale Variablen Häufigkeiten angezeigt. Es können jedoch auch Zeilen-, Spalten- oder Gesamtprozentwerte angezeigt werden.

- Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option *Auswertungsstatistik* aus.

Abbildung 1-16

Kontextmenü für kategoriale Variablen auf der Zeichenfläche



- Wählen Sie im Dialogfeld "Auswertungsstatistik" in der Liste "Statistik" die Option *Anzahl der Zeilen (%)* aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste "Anzeigen" zu übertragen.

Jetzt werden sowohl die Häufigkeiten als auch die Zeilenprozentwerte in der Tabelle angezeigt.

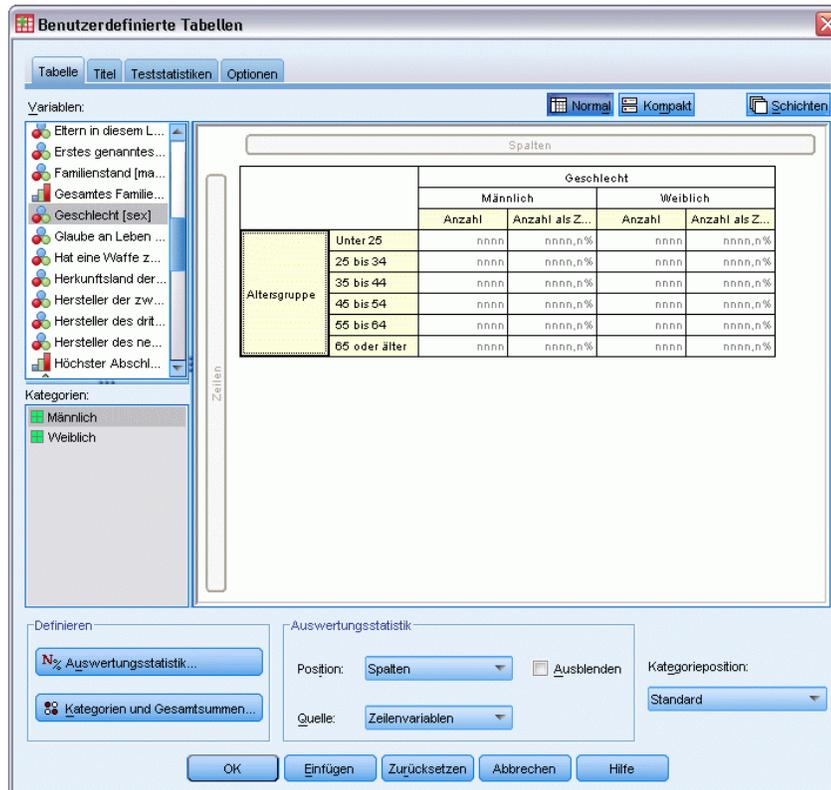
Abbildung 1-17  
Dialogfeld "Auswertungsstatistik" für kategoriale Variablen



- Klicken Sie auf **Der Auswahl zuweisen**, um diese Einstellungen zu speichern und zur Tabellenkonstruktion zurückzukehren.

Auf der Zeichenfläche werden die vorgenommenen Änderungen wiedergegeben. Es werden jetzt Spalten für die Häufigkeiten und die Zeilenprozentwerte angezeigt.

Abbildung 1-18  
Häufigkeiten und Zeilenprozentwerte auf der Zeichenfläche



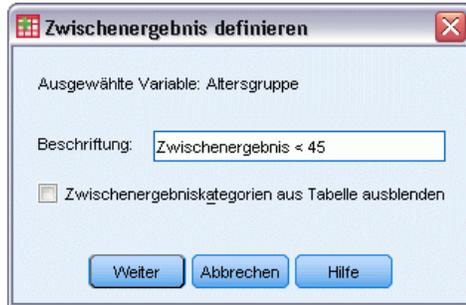
## Einfügen von Gesamtsummen und Zwischenergebnissen

In der Standardeinstellung werden in benutzerdefinierten Tabellen keine Gesamtsummen angezeigt. Gesamtsummen und Zwischenergebnisse können einer Tabelle jedoch problemlos hinzugefügt werden.

- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option *Kategorien und Gesamtsummen* aus.
- ▶ Wählen Sie im Dialogfeld "Kategorien und Gesamtsummen" in der Liste "Wert(e)" den Eintrag 3,00 aus, indem Sie darauf klicken.
- ▶ Klicken Sie auf *Zwischenergebnis* hinzufügen.

- ▶ Geben Sie im Dialogfeld “Zwischenergebnis definieren” Zwischenergebnis <45 ein und klicken Sie dann auf Weiter.

Abbildung 1-19  
Dialogfeld “Zwischenergebnis definieren”



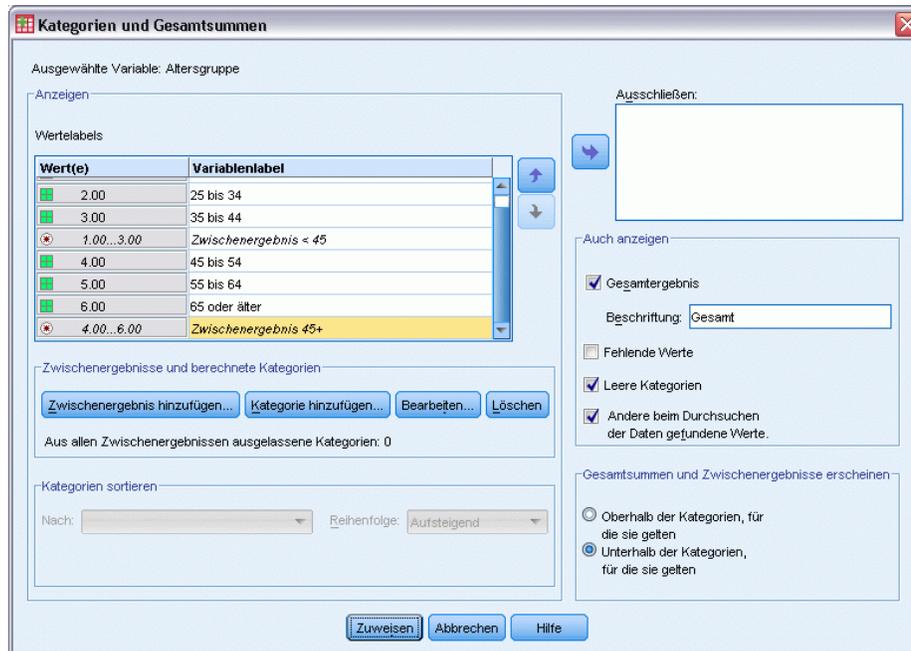
Hiermit fügen Sie eine Zeile mit dem Zwischenergebnis für die ersten drei Alterskategorien ein.

- ▶ Wählen Sie in der Liste “Wert(e)” den Eintrag 6,00 aus, indem Sie darauf klicken.
- ▶ Klicken Sie auf Zwischenergebnis hinzufügen.
- ▶ Geben Sie im Dialogfeld “Zwischenergebnis definieren” Zwischenergebnis 45+ ein und klicken Sie dann auf Weiter.

Hiermit fügen Sie eine Zeile mit dem Zwischenergebnis für die letzten drei Alterskategorien ein.

- ▶ Um eine Gesamtsumme in die Tabelle aufzunehmen, aktivieren Sie im Gruppenfeld “Auch anzeigen” das Kontrollkästchen Gesamtergebnis.

Abbildung 1-20  
Einfügen von Gesamtsummen und Zwischenergebnissen

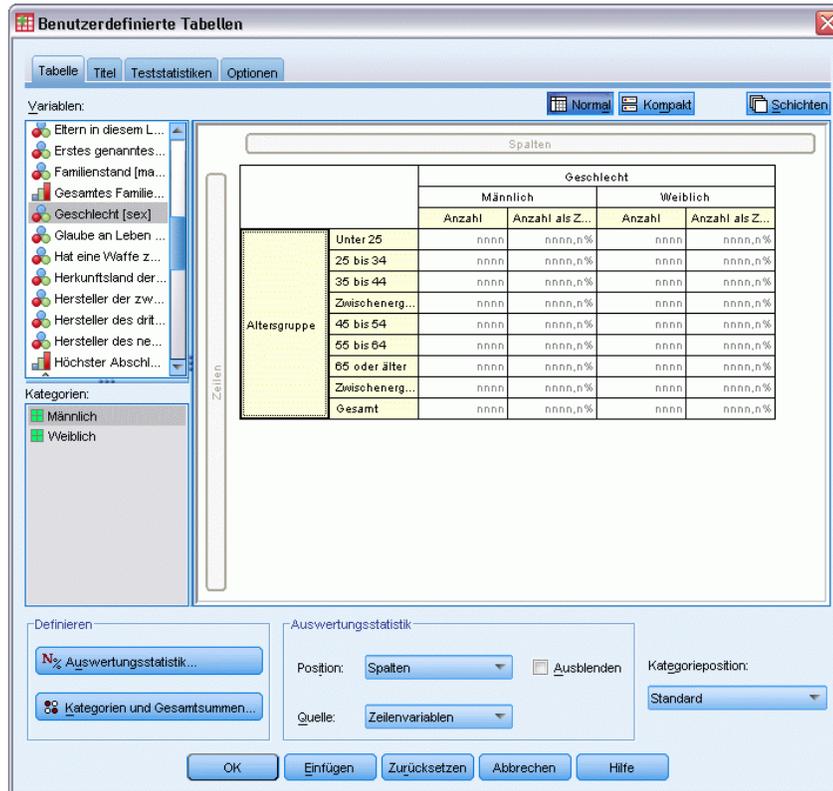


- Klicken Sie dann auf Zuweisen.

Die Vorschau auf der Zeichenfläche enthält jetzt Zeilen für die beiden Zwischenergebnisse und die Gesamtsumme.

Abbildung 1-21

*Gesamtsumme und Zwischenergebnisse auf der Zeichenfläche*



- Klicken Sie auf OK, um diese Tabelle zu erstellen.

Die Tabelle wird im Viewer angezeigt.

Abbildung 1-22

*Kreuztabelle mit Gesamtsummen und Zwischenergebnissen*

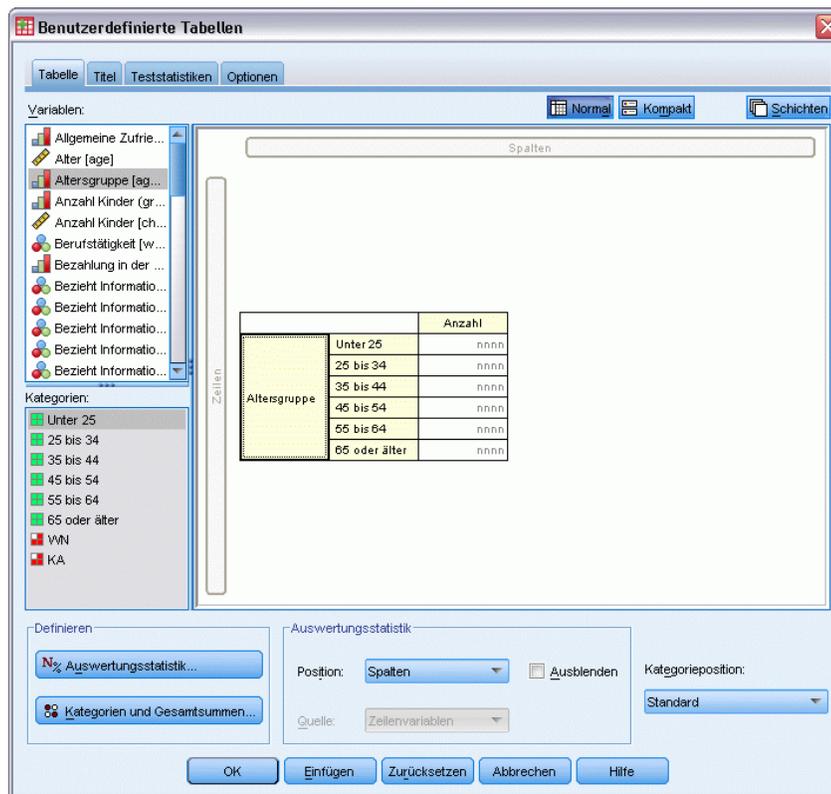
		Gender			
		Male		Female	
		Anzahl	Anzahl als ...	Anzahl	Anzahl als ...
Age category	Less than 25	108	44.6%	134	55.4%
	25 to 34	276	44.0%	351	56.0%
	35 to 44	309	45.5%	370	54.5%
	Subtotal < 45	693	44.8%	855	55.2%
	45 to 54	221	45.9%	260	54.1%
	55 to 64	136	42.5%	184	57.5%
	65 or older	178	37.2%	301	62.8%
	Subtotal 45+	535	41.8%	745	58.2%
	Total	1228	43.4%	1600	56.6%

## Auswerten von metrischen Variablen

In einer einfachen Kreuztabelle von zwei kategorialen Variablen werden Häufigkeiten oder Prozentwerte in den Tabellenzellen angezeigt. Bei metrischen Variablen können Sie jedoch auch Zusammenfassungen in die Zellen aufnehmen.

- ▶ Um das Dialogfeld zum Konstruieren von benutzerdefinierten Tabellen erneut zu öffnen, wählen Sie die folgenden Befehle aus den Menüs aus:  
Analysieren > Tabellen > Benutzerdefinierte Tabellen...
- ▶ Klicken Sie auf Zurücksetzen, um frühere Einstellungen zu löschen.
- ▶ Wählen Sie in der Variablenliste die Variable *Altersgruppe* aus, indem Sie darauf klicken, ziehen Sie die Variable auf die Zeichenfläche und legen Sie sie im Bereich für die Zeilen ab.

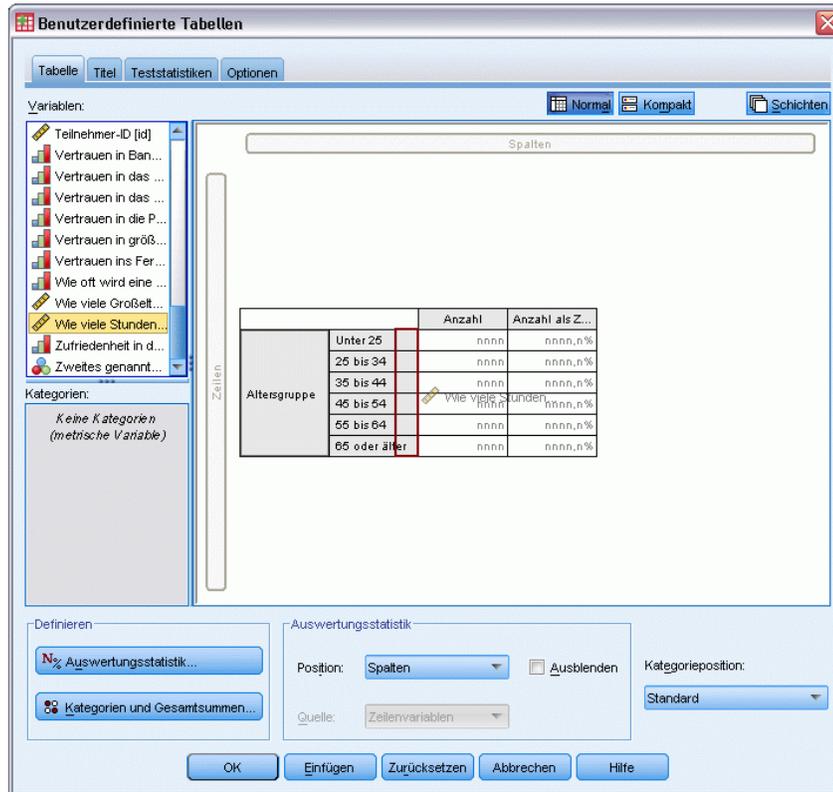
Abbildung 1-23  
Auswählen einer Zeilenvariablen



- Wählen Sie in der Variablenliste die Variable *Täglicher Fernsehkonsum in Stunden* aus, ziehen Sie sie auf die Zeichenfläche und legen Sie sie rechts neben *Alterskategorie* in der Zeilendimension der Tabelle ab.

Abbildung 1-24

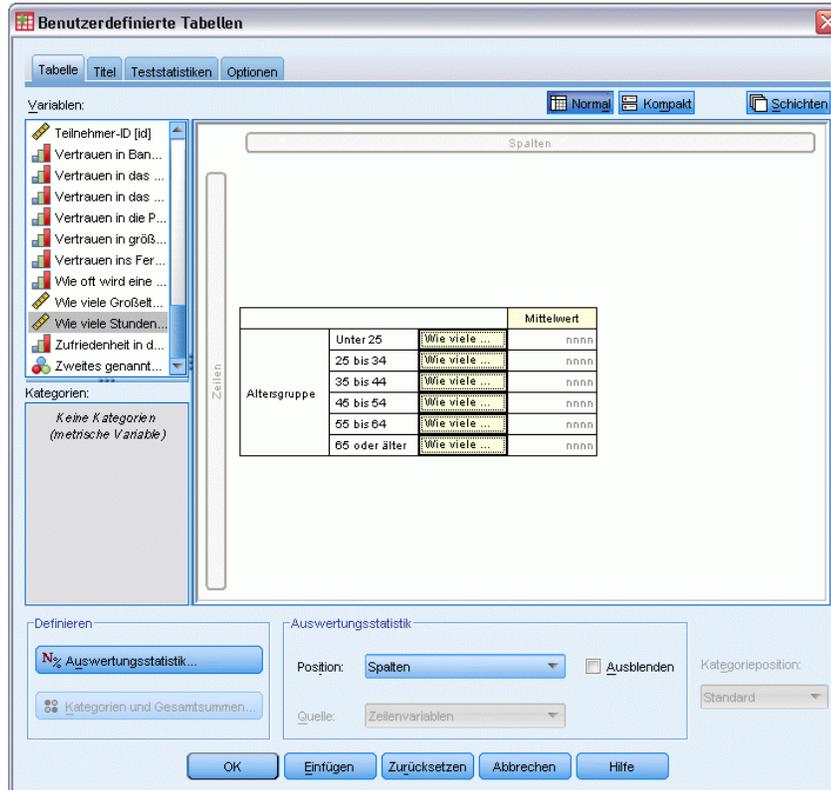
Ziehen und Ablegen einer metrischen Variablen in der Zeilendimension



Statt Kategoriehäufigkeiten wird in der Tabelle jetzt die durchschnittliche Anzahl an Stunden Fernsehkonsum pro Alterskategorie angezeigt.

Abbildung 1-25

Zusammenfassung einer metrischen Variablen in den Tabellenzellen

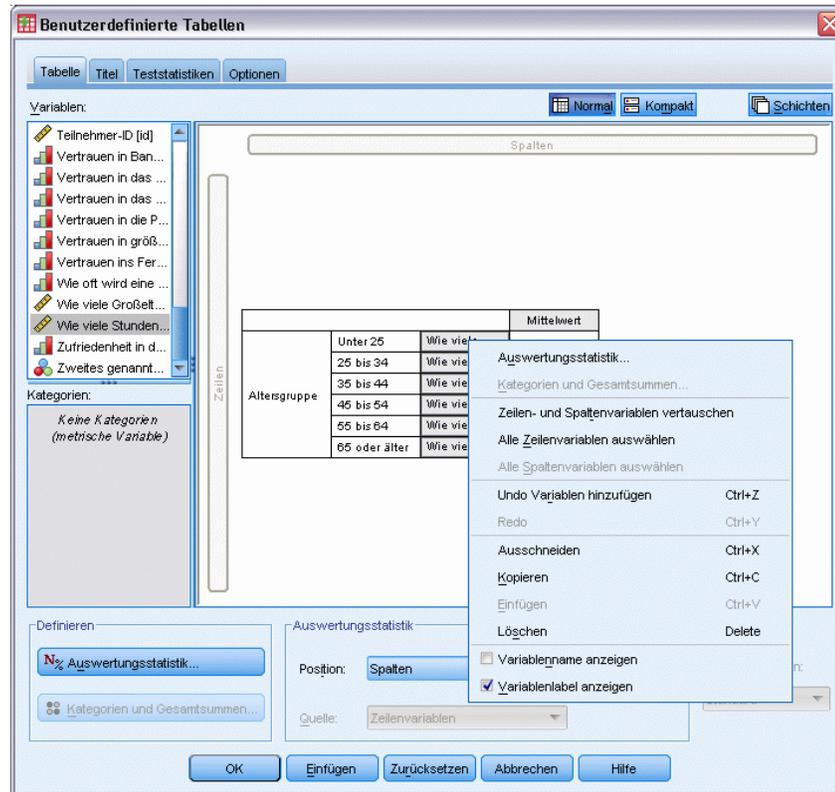


In der Standardeinstellung für die Auswertungsstatistik wird bei metrischen Variablen der Mittelwert berechnet. Sie können dies ändern und zusätzliche Auswertungsstatistiken hinzufügen.

- Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf die metrische Variable und wählen Sie aus dem Kontextmenü die Option Auswertungsstatistik aus.

Abbildung 1-26

Kontextmenü für metrische Variablen in der Tabellenvorschau



- Wählen Sie im Dialogfeld "Auswertungsstatistik" in der Liste "Statistik" die Option Median aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste "Anzeigen" zu übertragen.

Jetzt werden sowohl der Mittelwert als auch der Median in der Tabelle dargestellt.

Abbildung 1-27

Dialogfeld "Auswertungsstatistik" für metrische Variablen

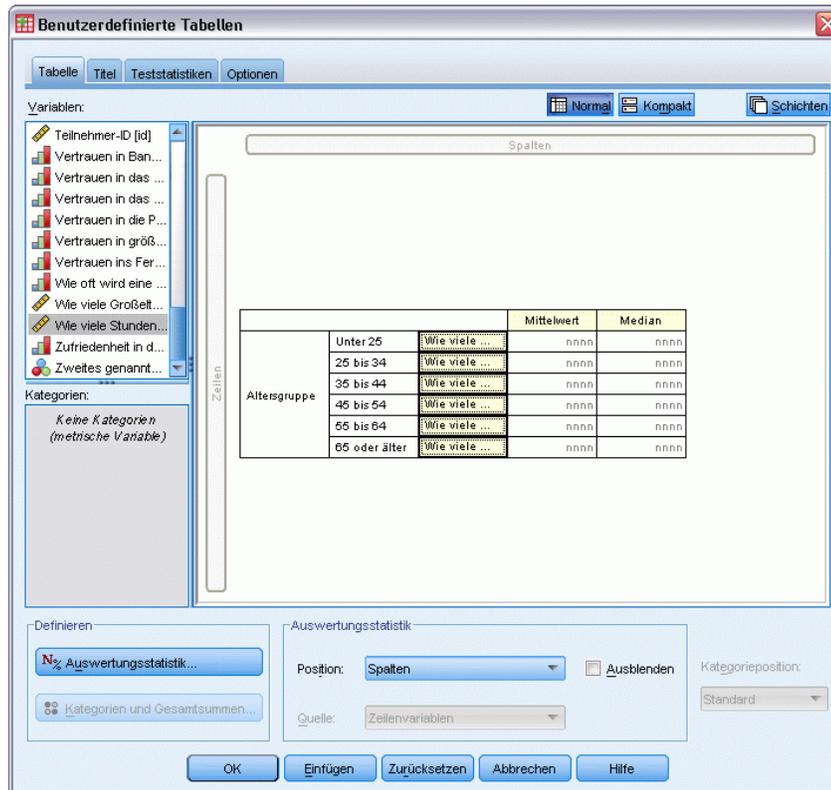


- Klicken Sie auf Der Auswahl zuweisen, um diese Einstellungen zu speichern und zur Tabellenkonstruktion zurückzukehren.

Auf der Zeichenfläche wird jetzt gezeigt, dass sowohl der Mittelwert als auch der Median in der Tabelle dargestellt werden.

Abbildung 1-28

Mittelwert und Median als Zusammenfassungen einer metrischen Variablen auf der Zeichenfläche

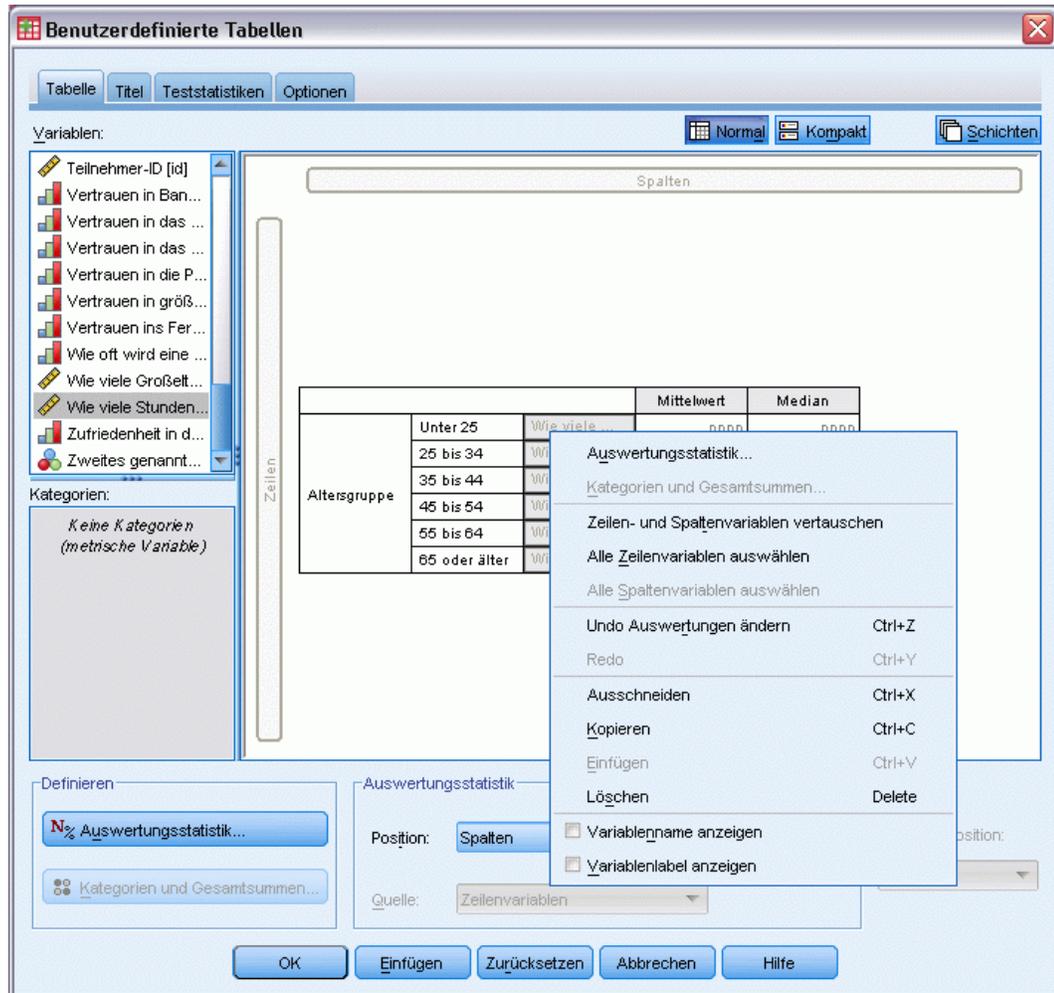


Ehe Sie diese Tabelle erstellen, können Sie die Darstellung ein wenig verbessern.

- Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Täglicher Fernsehkonsum...* und deaktivieren Sie im Kontextmenü die Option *Variablenlabel anzeigen*.

Abbildung 1-29

Unterdrücken der Anzeige von Variablenlabels



Die Spalte wird nach wie vor in der Tabellenvorschau angezeigt (wobei die Variablenlabels grau angezeigt werden), erscheint jedoch nicht in der resultierenden Tabelle.

- Klicken Sie im Dialogfeld zur Tabellenkonstruktion auf die Registerkarte *Titel*.

- Geben Sie einen aussagekräftigen Titel für die Tabelle ein, beispielsweise Mittlerer täglicher Fernsehkonsum in Stunden nach Alterskategorie.

Abbildung 1-30

Dialogfeld "Benutzerdefinierte Tabellen," Registerkarte "Titel"

The screenshot shows a dialog box titled "Benutzerdefinierte Tabellen" with a close button (X) in the top right corner. The dialog has four tabs: "Tabelle", "Titel", "Teststatistiken", and "Optionen". The "Titel" tab is selected. In the top right of the dialog, there are three buttons: "Datum", "Uhrzeit", and "Tabellenausdruck". The main area of the dialog is divided into three sections: "Titel:", "Erklärung:", and "Ecke:". The "Titel:" section contains the text "Mittlerer täglicher Fernsehkonsum in Stunden nach Alterskategorie". The "Erklärung:" and "Ecke:" sections are empty text boxes. At the bottom of the dialog, there is an information icon (i) followed by a note: "Um sicherzustellen, dass der Eckentext in der Tabelle angezeigt wird, prüfen Sie, dass in den Tabelleneigenschaften die Option 'Beschriftung für Zeilendimension' auf 'Verschachtelt' gesetzt ist. Dies kann in der Standard-Tabellenvorlage voreingestellt werden." Below the note are five buttons: "OK", "Einfügen", "Zurücksetzen", "Abbrechen", and "Hilfe".

- Klicken Sie auf OK, um die Tabelle zu erstellen.

Die Tabelle wird im Fenster des Viewers angezeigt.

Abbildung 1-31

*Mittelwert und Median des Fernsehkonsums in Stunden nach Alterskategorie*

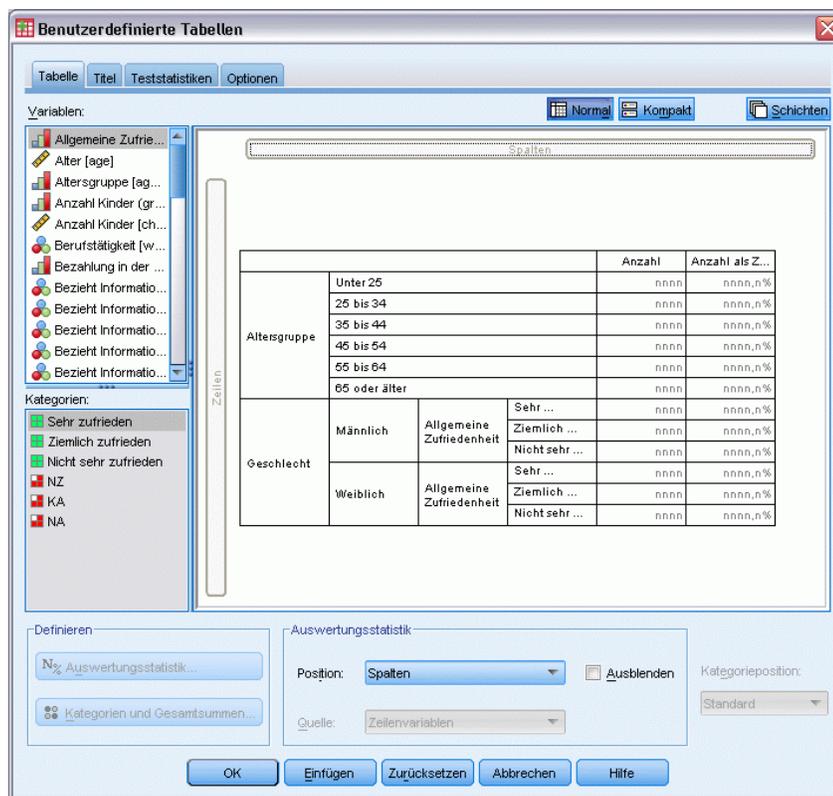
<b>Mittlerer täglicher Fernsehkonsum in Stunden nach Alters kategorie.</b>		<b>Mittelwert</b>	<b>Median</b>
Age category	Less than 25	2,85	2,00
	25 to 34	2,78	2,00
	35 to 44	2,56	2,00
	45 to 54	2,58	2,00
	55 to 64	3,02	2,50
	65 or older	3,58	3,00

# Benutzeroberfläche zum Konstruieren von Tabellen

Mit dem Befehl “Benutzerdefinierte Tabellen” öffnen Sie eine Benutzeroberfläche zum Konstruieren von Tabellen, die mit Ziehen und Ablegen bedient werden kann und in der Sie eine Vorschau der Tabelle erhalten, während Sie Variablen und Optionen auswählen. Außerdem finden Sie hier einen Grad an Flexibilität, den ein normales Dialogfeld nicht bietet, beispielsweise die Möglichkeit, die Größe des Fensters und die Größe der einzelnen Fensterbereiche zu ändern.

## Konstruieren von Tabellen

Abbildung 2-1  
Dialogfeld “Benutzerdefinierte Tabellen,” Registerkarte “Tabelle”



Auf der Registerkarte “Tabelle” des Dialogfelds für die Tabellenkonstruktion wählen Sie die Variablen und Auswertungsmaße aus, die in der Tabelle angezeigt werden sollen.

**Liste "Variablen"**. Die Variablen aus der Datendatei werden oben links im Fenster angezeigt. Bei benutzerdefinierten Tabellen werden zwei Messniveaus unterschieden, nach denen Variablen behandelt werden:

**Kategorial**. Daten mit einer begrenzten Anzahl von eindeutigen Werten bzw. Kategorien (beispielsweise Geschlecht oder Religion). Bei kategorialen Variablen kann es sich um String-Variablen (alphanumerisch) oder um numerische Variablen handeln, bei denen numerische Codes zum Darstellen der Kategorien verwendet werden (beispielsweise 0 = *Männlich* und 1 = *Weiblich*). Auch als qualitative Daten bezeichnet. Kategoriale Variablen können **nominal** oder **ordinal** sein.

- **Nominal**. Eine Variable kann als nominal behandelt werden, wenn ihre Kategorien sich nicht in eine natürliche Reihenfolge bringen lassen, z. B. die Firmenabteilung, in der eine Person arbeitet. Beispiele für nominale Variablen sind Region, Postleitzahl oder Religionszugehörigkeit.
- **Ordinal**. Eine Variable kann als ordinal behandelt werden, wenn ihre Werte für Kategorien stehen, die eine natürliche Reihenfolge aufweisen (z. B. Grad der Zufriedenheit mit Kategorien von sehr unzufrieden bis sehr zufrieden). Ordinale Variablen treten beispielsweise bei Einstellungsmessungen (Zufriedenheit oder Vertrauen) und bei Präferenzbeurteilungen auf.

**Metrisch**. Daten, die auf einer Intervall- oder Verhältnisskala gemessen werden und bei denen die Datenwerte sowohl die Reihenfolge der Werte als auch die Distanz zwischen den Werten festlegen. So ist beispielsweise ein Gehalt von \$ 72.195 höher als ein Gehalt von \$ 52.398 und die Distanz zwischen den Werten beträgt \$ 19.797. Auch als quantitative Daten bezeichnet.

Kategoriale Variablen definieren Kategorien (Zeilen, Spalten, Schichten) in der Tabelle. In der Standardeinstellung wird als Auswertungsstatistik die Anzahl (Anzahl der Fälle pro Kategorie) berechnet. In einer Standardtabelle für eine kategoriale Variable "Geschlecht" wird beispielsweise einfach die Anzahl der Männer und die Anzahl der Frauen aufgeführt.

Metrische Variablen werden normalerweise in den Kategorien von kategorialen Variablen ausgewertet. In der Standardeinstellung wird als Auswertungsstatistik der Mittelwert berechnet. So wird beispielsweise in einer Standardtabelle, in der das Einkommen nach Geschlecht kategorisiert ist, das durchschnittliche Einkommen von Männern und das durchschnittliche Einkommen von Frauen aufgeführt.

Es ist darüber hinaus möglich, metrische Variablen auszuwerten, ohne mithilfe von kategorialen Variablen Gruppen zu definieren. Dies ist in erster Linie zum **Stapeln** der Auswertungen von mehreren metrischen Variablen sinnvoll. [Für weitere Informationen siehe Thema Stapeln von Variablen auf S. 31.](#)

### **Mehrfachantworten-Sets**

In benutzerdefinierten Tabellen kann außerdem eine besondere Art von "Variable" verwendet werden, die als **Mehrfachantworten-Set** bezeichnet wird. Bei Mehrfachantworten-Sets handelt es sich nicht um Variablen im üblichen Sinn. Mehrfachantworten-Sets können nicht im Daten-Editor angezeigt werden, sie werden von anderen Prozeduren nicht erkannt. Mehrfachantworten-Sets verwenden mehrere Variablen, um Antworten auf Fragen aufzuzeichnen, auf welche der Befragte mehr als eine Antwort geben kann. Sie werden wie kategoriale Variablen behandelt und bieten

weitestgehend dieselben Möglichkeiten wie kategoriale Variablen. [Für weitere Informationen siehe Thema Mehrfachantworten-Sets in Kapitel 11 auf S. 169.](#)

Der Variablentyp ist durch ein Symbol neben der jeweiligen Variablen in der Variablenliste gekennzeichnet.

	Numerisch	Zeichenfolge	Datum	Zeit
Metrisch (stetig)		entfällt		
Ordinal				
Nominal				

Mehrfachantworten-Set, mehrere Kategorien	
Mehrfachantworten-Set, mehrere Dichotomien	

Sie können das Messniveau einer Variablen ändern, indem Sie im Dialogfeld zur Tabellenkonstruktion mit der rechten Maustaste auf die Variable in der Variablenliste klicken und im Kontextmenü **Kategorial** oder **Metrisch** auswählen. In der Variablenansicht des Daten-Editors kann das Messniveau einer Variablen dauerhaft geändert werden. Variablen, die als **nominal** oder **ordinal** definiert sind, werden bei benutzerdefinierten Tabellen als kategoriale Variablen behandelt.

**Kategorien.** Wenn Sie in der Variablenliste eine kategoriale Variable auswählen, werden die für diese Variable definierten Kategorien in der Liste “Kategorien” angezeigt. Die Kategorien werden außerdem auf der Zeichenfläche angezeigt, wenn Sie die Variable in einer Tabelle verwenden. Wenn für die Variable keine Kategorien definiert wurden, werden in der Liste “Kategorien” und auf der Zeichenfläche zwei Kategorienplatzhalter angezeigt: *Kategorie 1* und *Kategorie 2*.

Die definierten Kategorien, die in der Tabellenkonstruktion angezeigt werden, beruhen auf **Wertelabels**. Hierbei handelt es sich um aussagekräftige Beschreibungen, die verschiedenen Datenwerten zugeordnet sind (z. B. die beiden numerischen Werte 0 und 1 mit den jeweiligen Wertelabels *männlich* und *weiblich*). Wertelabels können in der Variablenansicht des Daten-Editors oder mit dem Befehl “Variableneigenschaften definieren” aus dem Menü “Daten” definiert werden.

**Zeichenfläche.** Tabellen werden durch Ziehen und Ablegen von Variablen auf den Zeilen und Spalten auf der Zeichenfläche konstruiert. Auf der Zeichenfläche wird eine Vorschau der Tabelle angezeigt, die erstellt wird. Die Zellen auf der Zeichenfläche enthalten keine tatsächlichen Datenwerte. Die Darstellung auf der Zeichenfläche bietet jedoch eine relativ genaue Layoutansicht der resultierenden Tabelle. Bei kategorialen Variablen enthält die tatsächliche

Tabelle möglicherweise mehr Kategorien als die Tabelle in der Vorschau, wenn sich in der Datendatei eindeutige Werte befinden, für die keine Wertelabels definiert wurden.

- In der Normalansicht werden alle Zeilen und Spalten angezeigt, die in der Tabelle enthalten sein werden, einschließlich der Zeilen bzw. Spalten für Auswertungsstatistiken und für die Kategorien der kategorialen Variablen.
- In der Kompaktansicht werden nur die Variablen angezeigt, die in der Tabelle enthalten sein werden, nicht jedoch die Vorschau der Zeilen und Spalten.

### **Grundregeln und Einschränkungen für das Konstruieren von Tabellen**

- Bei kategorialen Variablen beruhen Auswertungsstatistiken auf der innersten Variablen in der Quelldimension der Statistik.
- Die Standard-Quelldimension der Statistik (Zeile oder Spalte) für kategoriale Variablen beruht auf der Reihenfolge, in der Sie die Variablen auf die Zeichenfläche ziehen. Wenn Sie z. B. eine Variable zuerst in den Bereich für die Zeilen ziehen, wird die Zeilendimension als Standard-Quelldimension für die Statistik verwendet.
- Metrische Variablen können nur innerhalb der Kategorien der innersten Variablen in der Zeilen- oder der Spaltendimension ausgewertet werden. (Die metrische Variable kann auf einer beliebigen Ebene der Tabelle platziert werden, ihre Auswertung erfolgt jedoch auf der innersten Ebene.)
- Metrische Variablen können nicht innerhalb von anderen metrischen Variablen ausgewertet werden. Sie können Auswertungen von mehreren metrischen Variablen stapeln und metrische Variablen innerhalb der Kategorien von kategorialen Variablen auswerten. Es ist nicht möglich, metrische Variablen ineinander zu verschachteln oder eine metrische Variable in der Zeilendimension und eine andere in der Spaltendimension anzuordnen.
- Wenn eine Variable in der Arbeitsdatei mehr als 12.000 definierte Wertelabels enthält, können Sie nicht die Tabellenkonstruktion zum Erstellen von Tabellen verwenden. Wenn Sie keine Variablen, die diese Beschränkung überschreiten, in Ihre Tabellen aufnehmen müssen, können Sie Variablen-Sets definieren und anwenden, die diese Variablen ausschließen. Wenn Sie Variablen mit mehr als 12.000 definierten Wertelabeln aufnehmen müssen, können Sie die betreffenden Tabellen mit der Befehlssyntax `CTABLES` generieren.

### **So konstruieren Sie eine Tabelle:**

- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
Analysieren > Tabellen > Benutzerdefinierte Tabellen...
- ▶ Ziehen Sie eine oder mehrere Variablen auf die Bereiche für Zeilen bzw. Spalten auf der Zeichenfläche.
- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.  
So löschen Sie eine Variable aus der Zeichenfläche im Dialogfeld zum Konstruieren von Tabellen:
- ▶ Wählen Sie die Variable aus, indem Sie auf der Zeichenfläche darauf klicken.
- ▶ Ziehen Sie die Variable aus der Zeichenfläche heraus oder drücken Sie ENTF.

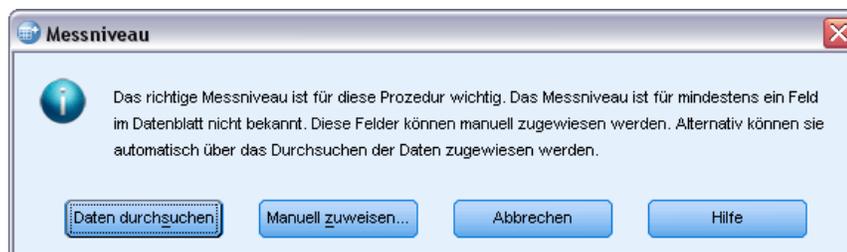
So ändern Sie das Messniveau einer Variablen:

- ▶ Klicken Sie mit der rechten Maustaste in der Variablenliste auf die Variable (dies ist nur in der Variablenliste, nicht jedoch auf der Zeichenfläche möglich).
- ▶ Wählen Sie im Kontextmenü die Option *Kategorial* oder die Option *Metrisch*.

### **Felder mit unbekanntem Messniveau**

Die Messniveau-Warnmeldung wird angezeigt, wenn das Messniveau für mindestens eine Variable (ein Feld) im Datenblatt unbekannt ist. Da sich das Messniveau auf die Berechnung der Ergebnisse für diese Prozedur auswirkt, müssen alle Variablen ein definiertes Messniveau aufweisen.

Abbildung 2-2  
Messniveau-Warnmeldung



- **Daten durchsuchen.** Liest die Daten im aktiven Datenblatt (Arbeitsdatei) und weist allen Feldern, deren Messniveau zurzeit nicht bekannt ist, das Standardmessniveau zu. Bei großen Datenblättern kann dieser Vorgang einige Zeit in Anspruch nehmen.
- **Manuell zuweisen.** Öffnet ein Dialogfeld, in dem alle Felder mit unbekanntem Messniveau aufgeführt werden. Mit diesem Dialogfeld können Sie diesen Feldern ein Messniveau zuweisen. Außerdem können Sie in der Variablenansicht des Daten-Editors ein Messniveau zuweisen.

Da das Messniveau für diese Prozedur bedeutsam ist, können Sie erst dann auf das Dialogfeld zur Ausführung dieser Prozedur zugreifen, wenn für alle Felder ein Messniveau definiert wurde.

## **Stapeln von Variablen**

Beim Stapeln werden separate Tabellen in dieselbe Anzeige eingefügt. So können beispielsweise Informationen über das *Geschlecht* und die *Alterskategorie* in unterschiedlichen Abschnitten derselben Tabelle angezeigt werden.

### **So stapeln Sie Variablen:**

- ▶ Wählen Sie in der Variablenliste alle zu stapelnden Variablen aus, ziehen Sie diese gemeinsam auf die Zeichenfläche, und legen Sie sie im Bereich für die Zeilen oder die Spalten ab.

*ODER*

- ▶ Ziehen Sie jede Variable einzeln, und legen Sie diese oberhalb oder unterhalb von vorhandenen Variablen in den Zeilen bzw. links oder rechts von vorhandenen Variablen in den Spalten ab.

Abbildung 2-3  
Gestapelte Variablen

		Anzahl
Gender	Male	1232
	Female	1600
Age category	Less than 25	242
	25 to 34	627
	35 to 44	679
	45 to 54	481
	55 to 64	320
	65 or older	479

Für weitere Informationen siehe Thema Stapeln von kategorialen Variablen in Kapitel 4 auf S. 70.

## Verschachteln von Variablen

Ähnlich wie bei Kreuztabellen kann mit Verschachtelung die Beziehung zwischen zwei kategorialen Variablen aufgezeigt werden. Bei der Verschachtelung wird eine der Variablen jedoch in derselben Dimension innerhalb der anderen Variablen verschachtelt. So können Sie beispielsweise *Geschlecht* in *Alterskategorie* in der Zeilendimension verschachteln, um die Anzahl der Männer und Frauen in jeder Alterskategorie darzustellen.

Es ist außerdem möglich, metrische Variablen in kategorialen Variablen zu verschachteln. So können Sie beispielsweise *Einkommen* in *Geschlecht* verschachteln, sodass das mittlere Einkommen (oder der Median bzw. ein anderes Auswertungsmaß des Einkommens) für Frauen und Männer getrennt aufgeführt wird.

### So verschachteln Sie Variablen:

- ▶ Ziehen Sie eine kategoriale Variable auf die Zeichenfläche, und legen Sie sie im Bereich für die Zeilen oder die Spalten ab.
- ▶ Ziehen Sie eine kategoriale oder eine metrische Variable auf die Zeichenfläche, und legen Sie sie links oder rechts von der kategorialen Zeilenvariablen bzw. oberhalb oder unterhalb der kategorialen Spaltenvariablen ab.

Abbildung 2-4  
Verschachtelte kategoriale Variablen

			Anzahl
Age category	Less than 25	Male	108
		Female	134
	25 to 34	Male	276
		Female	351
	35 to 44	Male	309
		Female	370
	45 to 54	Male	221
		Female	260
	55 to 64	Male	136
		Female	184
	65 or older	Male	178
		Female	301

Abbildung 2-5

Eine metrische Variable wurde in einer kategorialen Variablen verschachtelt.

**Mittlerer täglicher Fernsehkonsum in Stunden  
nach  
Alterskategorie.**

		Mittelwert	Median
Age category	Less than 25	2,85	2,00
	25 to 34	2,78	2,00
	35 to 44	2,56	2,00
	45 to 54	2,58	2,00
	55 to 64	3,02	2,50
	65 or older	3,58	3,00

*Hinweis:* Formal gesehen stellt die vorangegangene Tabelle ein Beispiel für die Verschachtelung einer kategorialen Variablen innerhalb einer metrischen dar. Die Tabelle vermittelt jedoch im wesentlichen dieselben Informationen wie bei der Verschachtelung der metrischen Variablen innerhalb der kategorialen, enthält aber keine überflüssigen Beschriftungen für die metrische Variable. (Wenn Sie es andersherum versuchen, wird diese Sachlage verdeutlicht.)

Für weitere Informationen siehe [Thema Verschachteln von kategorialen Variablen in Kapitel 4 auf S. 73](#).

*Anmerkung:* Bei benutzerdefinierten Tabellen wird die Verarbeitung geschichteter aufgeteilter Dateien nicht berücksichtigt. Um dasselbe Ergebnis wie bei geschichteten aufgeteilten Dateien zu erzielen, müssen Sie die Dateiaufteilungs-Variablen in den äußersten Verschachtelungsebenen der Datei platzieren.

## Schichten

Mit Schichten können Tabellen um eine Tiefendimension erweitert werden, um dreidimensionale "Würfel" zu erstellen. Schichten ähneln dem Stapeln und Verschachteln. Der wichtigste Unterschied besteht darin, dass stets nur eine Schichtenkategorie sichtbar ist. Wenn Sie beispielsweise *Alterskategorie* als Zeilenvariable und *Geschlecht* als Schichtvariable verwenden, erhalten Sie eine Tabelle, in der die Informationen für Männer und Frauen in unterschiedlichen Schichten dargestellt werden.

### So erstellen Sie Schichten:

- ▶ Klicken Sie auf der Registerkarte "Tabelle" des Dialogfelds zum Konstruieren von Tabellen auf Schichten, um die Liste "Schichten" anzuzeigen.
- ▶ Ziehen Sie die zum Definieren der Schichten vorgesehenen metrischen oder kategorialen Variablen in die Liste "Schichten", und legen Sie diese dort ab.

Abbildung 2-6  
Variablen in Schichten

Geschlecht Weiblich		
Alterskategorie		Anzahl
Weniger als 25		134
25 bis 34		351
35 bis 44		370

Geschlecht Männlich		
Alterskategorie		Anzahl
Weniger als 25		108
25 bis 34		276
35 bis 44		309
45 bis 54		221
55 bis 64		136
65 oder älter		178

Metrische und kategoriale Variablen können in der Liste “Schichten” nicht kombiniert werden. Alle Variablen müssen denselben Typ aufweisen. Mehrfachantworten-Sets werden in der Liste “Schichten” wie kategoriale Variablen behandelt. Metrische Variablen in den Schichten werden immer gestapelt.

Wenn mehrere kategoriale Schichtvariablen vorliegen, können Schichten gestapelt oder verschachtelt werden.

- Jede Kategorie als Schicht zeigen entspricht einer Stapelung. Für jede Kategorie jeder Schichtvariablen wird eine eigene Schicht angezeigt. Die Gesamtzahl der Schichten ergibt sich einfach aus der *Summe* der Anzahl Kategorien jeder Schichtvariablen. Wenn beispielsweise drei Schichtvariablen vorhanden sind, von denen jede drei Kategorien besitzt, enthält die Tabelle neun Schichten.
- Jede Kombination von Kategorien als Schicht zeigen entspricht der Verschachtelung bzw. dem Erstellen von Kreuztabellen für Schichten. Die Gesamtzahl der Schichten ergibt sich aus dem *Produkt* der Anzahl Kategorien jeder Schichtvariablen. Wenn beispielsweise drei Variablen vorhanden sind, von denen jede drei Kategorien besitzt, enthält die Tabelle 27 Schichten.

### **Einblenden und Ausblenden von Variablennamen bzw. -labels**

Für die Anzeige von Variablennamen und -labels sind die folgenden Optionen verfügbar:

- Nur Variablenlabels anzeigen. Bei Variablen, für die kein Variablenlabel definiert ist, wird der Variablenname angezeigt. Dies ist die Standardeinstellung.
- Nur Variablennamen anzeigen.
- Sowohl Variablenlabels als auch Variablennamen anzeigen.
- Weder Variablennamen noch Variablenlabels anzeigen. Die Spalte bzw. Zeile mit dem Variablenlabel oder -namen wird zwar weiterhin in der Tabellenvorschau auf der Zeichenfläche angezeigt, in der tatsächlichen Tabelle ist diese Spalte bzw. Zeile dann jedoch ausgeblendet.

So zeigen Sie Variablenlabels und -namen an oder blenden sie aus:

- ▶ Klicken Sie mit der rechten Maustaste auf die Tabellenvorschau auf der Zeichenfläche.

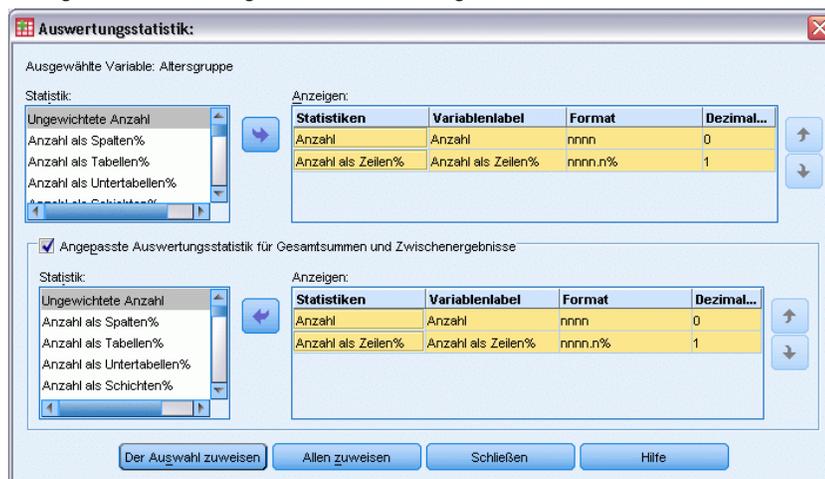
- Wählen Sie im Kontextmenü die Option Variablenlabel anzeigen oder die Option Variablenname anzeigen, um zwischen ein- und ausgeblendeten Labels und Namen zu wechseln. Ein Häkchen neben einer Option zeigt an, dass die entsprechenden Angaben angezeigt werden.

## Auswertungsstatistik

Das Dialogfeld “Auswertungsstatistik” bietet die folgenden Funktionen:

- Hinzufügen und Entfernen von Auswertungsstatistiken zu bzw. aus einer Tabelle
- Ändern der Beschriftung von Statistiken
- Ändern der Reihenfolge der Statistiken
- Ändern des Formats einer Statistik, einschließlich der Anzahl Dezimalstellen

Abbildung 2-7  
Dialogfeld “Auswertungsstatistik” für kategoriale Variablen



Es hängt vom Messniveau der jeweiligen Quellvariablen ab, welche Auswertungsstatistiken (und andere Optionen) verfügbar sind, wie oben im Dialogfeld dargestellt. Die Quelle der Auswertungsstatistiken (die Variable, auf der die Auswertungsstatistiken beruhen) wird durch die folgenden Faktoren bestimmt:

- **Messniveau.** Wenn eine Tabelle (oder ein Tabellenabschnitt in einer gestapelten Tabelle) eine metrische Variable enthält, beruhen Auswertungsstatistiken auf der metrischen Variablen.
- **Reihenfolge der Variablenauswahl.** Die Standard-Quelldimension der Statistik (Zeile oder Spalte) für kategoriale Variablen beruht auf der Reihenfolge, in der Sie die Variablen auf die Zeichenfläche ziehen. Wenn Sie z. B. eine Variable zuerst in den Bereich für die Zeilen ziehen, wird die Zeilendimension als Standard-Quelldimension für die Statistik verwendet.
- **Verschachtelung.** Bei kategorialen Variablen beruhen Auswertungsstatistiken auf der innersten Variablen in der Quelldimension der Statistik.

Eine gestapelte Tabelle kann mehrere Quellvariablen für Auswertungsstatistiken aufweisen (sowohl metrisch als auch kategorial); jeder Tabellenabschnitt besitzt jedoch nur eine Quelle für die Auswertungsstatistik.

**So ändern Sie die Quelldimension der Auswertungsstatistik:**

- ▶ Wählen Sie auf der Registerkarte “Tabelle” im Gruppenfeld “Auswertungsstatistik” in der Dropdown-Liste Quelle die Dimension (Zeilen, Spalten oder Schichten) aus.

**So legen Sie die in einer Tabelle angezeigte Auswertungsstatistik fest:**

- ▶ Wählen Sie die Quellvariable für die Auswertungsstatistik aus, indem Sie auf der Zeichenfläche der Registerkarte “Tabelle” darauf klicken.
- ▶ Klicken Sie auf der Registerkarte “Tabelle” im Gruppenfeld “Definieren” auf Auswertungsstatistik.

*ODER*

- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf die Quellvariable für die Auswertungsstatistik und wählen Sie im Kontextmenü die Option Auswertungsstatistik aus.
- ▶ Wählen Sie die Auswertungsstatistik aus, die Sie in die Tabelle aufnehmen möchten. Mit der Pfeilschaltfläche können Sie ausgewählte Statistiken aus der Liste “Statistik” in die Liste “Anzeigen” übertragen. Sie können die Statistiken auch aus der Liste “Statistik” in die Liste “Anzeigen” ziehen und dort ablegen.
- ▶ Klicken Sie auf die nach oben oder unten weisende Pfeilschaltfläche, um die Anzeigeposition der aktuell ausgewählten Auswertungsstatistik zu ändern.
- ▶ Wählen Sie aus der Dropdown-Liste “Format” ein Anzeigeformat für die ausgewählte Auswertungsstatistik aus.
- ▶ Geben Sie in der Zelle “Dezimalstellen” die Anzahl der anzuzeigenden Dezimalstellen für die ausgewählte Auswertungsstatistik ein.
- ▶ Klicken Sie auf Der Auswahl zuweisen, um die ausgewählten Auswertungsstatistiken für die aktuell ausgewählten Variablen auf der Zeichenfläche in die Tabelle aufzunehmen.
- ▶ Klicken Sie auf Allen zuweisen, um die ausgewählten Auswertungsstatistiken für alle gestapelten Variablen des gleichen Typs auf der Zeichenfläche in die Tabelle aufzunehmen.

*Hinweis:* Allen zuweisen unterscheidet sich nur bei gestapelten Variablen gleichen Typs, die sich bereits auf der Zeichenfläche befinden, von Der Auswahl zuweisen. In beiden Fällen werden die ausgewählten Auswertungsstatistiken automatisch auf alle weiteren gestapelten Variablen gleichen Typs angewendet, die Sie der Tabelle hinzufügen.

**Auswertungsstatistiken für kategoriale Variablen**

Als grundlegende Statistiken für kategoriale Variablen sind Häufigkeiten und Prozentwerte verfügbar. Zusätzlich können angepaßte Auswertungsstatistiken für Gesamtsummen und Zwischenergebnisse festgelegt werden. Diese angepaßten Auswertungsstatistiken umfassen Lagemaße (wie Mittelwert und Median) und Streuungsmaße (wie Standardabweichung), die sich für einige ordinale kategoriale Variablen eignen. [Für weitere Informationen siehe Thema Angepasste Auswertungsstatistik für Gesamtsummen bei kategorialen Variablen auf S. 40.](#)

**Häufigkeiten.** Anzahl der Fälle in jeder Zelle einer Tabelle bzw. Anzahl der Antworten bei Mehrfachantworten-Sets.

**Ungewichtete Anzahl.** Ungewichtete Anzahl von Fällen in jeder Zelle der Tabelle.

**Spaltenprozent.** Prozentwerte in jeder Spalte. Die Summe der Prozentwerte in jeder Spalte einer Untertabelle beträgt 100 % (bei einfachen Prozentwerten). Spaltenprozent sind in der Regel nur dann sinnvoll, wenn eine kategoriale *Zeilenvariable* vorhanden ist.

**Zeilenprozent.** Prozentwerte in jeder Zeile. Die Summe der Prozentwerte in jeder Zeile einer Untertabelle beträgt 100 % (bei einfachen Prozentwerten). Zeilenprozent sind in der Regel nur dann sinnvoll, wenn eine kategoriale *Spaltenvariable* vorhanden ist.

**Zeilen- und Spaltenprozent in Schichten.** Die Summe der Zeilen- oder Spaltenprozent über alle Untertabellen in einer verschachtelten Tabelle beträgt 100 % (bei einfachen Prozentwerten). Wenn die Tabelle Schichten enthält, ergibt die Summe der Zeilen- oder Spaltenprozent über alle verschachtelten Untertabellen jeder Schicht 100 %.

**Schichtenprozent.** Prozentwerte in jeder Schicht. Bei einfachen Prozentwerten ergibt die Summe der Prozentwerte für Zellen in der aktuell sichtbaren Schicht 100 %. Wenn keine Schichtvariablen vorhanden sind, entspricht dies den Tabellenprozent.

**Tabellenprozent.** Die Prozentwerte für jede Zelle beziehen sich auf die gesamte Tabelle. Alle Prozentwerte für Zellen beruhen auf derselben Gesamtanzahl von Fällen; die Summe der Prozentwerte über die gesamte Tabelle beträgt 100 % (bei einfachen Prozentwerten).

**Untertabellenprozent.** Die Prozentwerte für jede Zelle beziehen sich auf die Untertabelle. Alle Prozentwerte für Zellen in der Untertabelle beruhen auf derselben Gesamtanzahl von Fällen; die Summe der Prozentwerte über die gesamte Untertabelle beträgt 100 %. In verschachtelten Tabelle werden die Untertabellen durch die Variable definiert, die der innersten Verschachtelungsebene vorangeht. So werden beispielsweise die Untertabellen in einer Tabelle mit *Familienstand* innerhalb von *Geschlecht* innerhalb von *Alterskategorie* durch die Variable *Geschlecht* definiert.

Die Prozentwerte von Mehrfachantworten-Sets können sich auf Fälle, Antworten oder Häufigkeiten beziehen. [Für weitere Informationen siehe Thema Auswertungsstatistiken für Mehrfachantworten-Sets auf S. 38.](#)

### ***Gestapelte Tabellen***

Bei der Berechnung der Prozentwerte wird jeder von einer gestapelten Variablen definierte Tabellenabschnitt als eigene Tabelle behandelt. Die Summe von Zeilen- und Spaltenprozent in Schichten sowie Tabellenprozent in jedem gestapelten Tabellenabschnitt beträgt 100 % (bei einfachen Prozentwerten). Bei anderen Prozentberechnungen basieren die Prozentwerte auf den Fällen in jedem gestapelten Tabellenabschnitt.

### **Berechnungsgrundlage für Prozentwerte**

Abhängig von der Behandlung fehlender Werten in der Berechnungsgrundlage können Prozentwerte auf drei verschiedenen Arten berechnet werden:

**Einfacher Prozentwert.** Die Prozentwerte beziehen sich auf die Anzahl der Fälle, die in der Tabelle verwendet werden. Die Summe der Prozentwerte beträgt stets 100 %. Wenn eine Kategorie aus der Tabelle ausgeschlossen ist, werden die Fälle dieser Kategorie bei der Berechnung nicht berücksichtigt. Fälle mit systemdefiniert fehlenden Werten werden immer aus der Berechnungsgrundlage ausgeschlossen. Fälle mit benutzerdefiniert fehlenden Werten werden bei der Berechnung nicht berücksichtigt, wenn benutzerdefiniert fehlende Kategorien aus der Tabelle ausgeschlossen sind (Standardeinstellung). Sie werden berücksichtigt, wenn benutzerdefiniert fehlende Kategorien in der Tabelle enthalten sind. Alle Prozentwerte ohne *Gültige N* oder *Gesamtanzahl* in der Bezeichnung sind einfache Prozentwerte.

**Prozentwerte auf Grundlage der Gesamtanzahl.** Der Berechnungsgrundlage für einfache Prozentwerte werden die Fälle mit systemdefiniert fehlenden und benutzerdefiniert fehlenden Werten hinzugefügt. Die Summe der Prozentwerte kann weniger als 100 % betragen.

**Prozentwerte auf Grundlage der gültigen N.** Fälle mit benutzerdefiniert fehlenden Werten werden aus der Berechnungsgrundlage für einfache Prozentwerte entfernt, auch wenn benutzerdefiniert fehlende Kategorien in der Tabelle enthalten sind.

*Hinweis:* Fälle in manuell ausgeschlossenen Kategorien, die keine benutzerdefiniert fehlenden Kategorien sind, werden bei der Berechnung nie berücksichtigt.

### **Auswertungsstatistiken für Mehrfachantworten-Sets**

Für Mehrfachantworten-Sets sind die folgenden zusätzlichen Auswertungsstatistiken verfügbar.

**Antworten als Spalten%/Zeilen%/Schichten%.** Die Prozentwerte basieren auf Antworten.

**Antworten als Spalten%/Zeilen%/Schichten% (Basis: Anzahl).** Der Zähler enthält die Antworten, der Nenner die Gesamtanzahl.

**Anzahl als Spalten%/Zeilen%/Schichten% (Basis: Antworten).** Der Zähler enthält die Anzahl, der Nenner die Gesamtanzahl der Antworten.

**Antworten als Spalten%/Zeilen% in Schicht.** Prozentwert für mehrere Untertabellen. Die Prozentwerte basieren auf Antworten.

**Antworten als Spalten%/Zeilen% in Schicht (Basis: Anzahl).** Prozentwerte für mehrere Untertabellen. Der Zähler enthält die Antworten, der Nenner die Gesamtanzahl.

**Antworten als Spalten%/Zeilen% in Schicht (Basis: Antworten).** Prozentwerte für mehrere Untertabellen. Der Zähler enthält die Anzahl, der Nenner die Gesamtanzahl der Antworten.

**Antworten.** Anzahl der Antworten.

**Antworten als Untertabellen%/Tabellen%.** Die Prozentwerte basieren auf Antworten.

**Antworten als Untertabellen%/Tabellen% (Basis: Anzahl).** Der Zähler enthält die Antworten, der Nenner die Gesamtanzahl.

**Anzahl als Untertabellen%/Tabellen% (Basis: Antworten).** Der Zähler enthält die Anzahl, der Nenner die Gesamtanzahl der Antworten.

### **Auswertungsstatistiken für metrische Variablen und angepasste Gesamtsummen für kategoriale Variablen**

Zusätzlich zu den für kategoriale Variablen verfügbaren Häufigkeiten und Prozentwerten sind für metrische Variablen die folgenden Auswertungsstatistiken verfügbar, die auch als angepasste Auswertungen für Gesamtsummen und Zwischenergebnisse bei kategorialen Variablen verwendet werden können. Diese Auswertungsstatistiken sind für Mehrfachantworten-Sets oder (alphanumerische) String-Variablen nicht verfügbar.

**Mittelwert.** Arithmetisches Mittel; die Summe geteilt durch die Anzahl der Fälle.

**Median.** Wert, über und unter dem jeweils die Hälfte der Fälle liegt; 50. Perzentil.

**Modalwert.** Häufigster Wert. Wenn zwei oder mehr Werte gleich häufig auftreten, wird der kleinste Wert angezeigt.

**Minimum.** Kleinster (niedrigster) Wert.

**Maximum.** Größter (höchster) Wert.

**Fehlend.** Anzahl der fehlenden Werte (sowohl benutzerdefiniert fehlende als auch systemdefiniert fehlende Werte).

**Perzentil.** Sie können das 5., 25., 75., 95. und/oder 99. Perzentil abrufen.

**Bereich.** Differenz zwischen größtem und kleinstem Wert.

**Standardfehler des Mittelwerts.** Ein Maß dafür, wie stark der Mittelwert von Stichprobe zu Stichprobe in derselben Verteilung variieren kann. Dieser Wert kann für einen ungefähren Vergleich des beobachteten Mittelwerts mit einem hypothetischen Wert verwendet werden. (Sie können schließen, dass die beiden Werte unterschiedlich sind, wenn der Quotient von Differenz und Standardfehler kleiner als  $-2$  oder größer als  $+2$  ist.)

**Standardabweichung** Ein Maß für die Streuung um den Mittelwert. In einer Normalverteilung liegen 68 % der Fälle innerhalb von einer Standardabweichung des Mittelwerts und 95 % der Fälle innerhalb von zwei Standardabweichungen. Wenn beispielsweise das durchschnittliche Alter 45 und die Standardabweichung 10 beträgt, liegen bei einer Normalverteilung 95 % der Fälle im Bereich zwischen 25 und 65 Jahren (Quadratwurzel der Varianz).

**Summe.** Die Summe der Werte.

**Prozent der Summe.** Prozentwerte bezogen auf Summen. Verfügbar für Zeilen und Spalten (von Untertabellen), ganze Zeilen und Spalten (über mehrere Untertabellen), Schichten, Untertabellen und ganze Tabellen.

**Gesamtanzahl.** Anzahl der nichtfehlenden, benutzerdefiniert fehlenden und systemdefiniert fehlenden Werten. Fälle in manuell ausgeschlossenen Kategorien, die nicht zu den benutzerdefiniert fehlenden Kategorien gehören, werden hierbei nicht berücksichtigt.

**Gültige N.** Anzahl der nichtfehlenden Werte. Fälle in manuell ausgeschlossenen Kategorien, die nicht zu den benutzerdefiniert fehlenden Kategorien gehören, werden hierbei nicht berücksichtigt.

**Varianz.** Ein Maß der Streuung um den Mittelwert, gleich der Summe der quadrierten Abweichungen vom Mittelwert geteilt durch eins weniger als die Anzahl der Fälle. Die Einheit der Varianz entspricht der quadrierten Einheit der eigentlichen Variablen (quadrierte Standardabweichung).

### **Gestapelte Tabellen**

Jeder von einer gestapelten Variablen definierte Tabellenabschnitt wird als eigene Tabelle behandelt, und die Auswertungsstatistiken werden dementsprechend berechnet.

### **Angepasste Auswertungsstatistik für Gesamtsummen bei kategorialen Variablen**

Tabellen mit kategorialen Variablen, die Gesamtsummen oder Zwischenergebnisse enthalten, können andere Auswertungsstatistiken als die Zusammenfassungen enthalten, die für jede Kategorie angezeigt werden. So können beispielsweise Häufigkeiten und Spaltenprozentage für eine ordinale kategoriale Zeilenvariable und der Median für die Gesamtsummen angezeigt werden.

So erstellen Sie eine Tabelle für eine kategoriale Variable mit einer angepaßten Gesamtsumme:

- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
Analysieren > Tabellen > Benutzerdefinierte Tabellen...  
Das Dialogfeld zum Konstruieren von Tabellen wird geöffnet.
- ▶ Ziehen Sie eine kategoriale Variable auf die Zeichenfläche, und legen Sie sie im Bereich für die Zeilen oder die Spalten ab.
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf die Variable und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.
- ▶ Klicken Sie auf das Kontrollkästchen Gesamt (aktivieren Sie dieses) und klicken Sie anschließend auf Zuweisen.
- ▶ Klicken Sie erneut mit der rechten Maustaste auf der Zeichenfläche auf die Variable und wählen Sie im Kontextmenü die Option Auswertungsstatistik aus.
- ▶ Klicken Sie auf das Kontrollkästchen Angepaßte Auswertungsstatistik für Gesamtsummen und Zwischenergebnisse (aktivieren Sie dieses) und wählen Sie anschließend die gewünschte angepaßte Auswertungsstatistik.

In der Standardeinstellung werden alle Auswertungsstatistiken einschließlich der angepaßten in der entgegengesetzten Dimension der kategorialen Variablen dargestellt. Wenn beispielsweise eine kategoriale Zeilenvariable vorhanden ist, definieren die Auswertungsstatistiken wie im folgenden Beispiel die Spalten der Tabelle:

**Abbildung 2-8**

*Standardposition der Auswertungsstatistiken*

		Anzahl	Mittelwert
Vertrauen in das Fernsehen	Sehr viel	196	
	Mäßig	936	
	Fast kein	744	
	Gesamt	1876	2,29

So zeigen Sie Auswertungsstatistiken in derselben Dimension wie die kategoriale Variable an:

- Wählen Sie im Dialogfeld zum Konstruieren von Tabellen auf der Registerkarte “Tabelle” im Gruppenfeld “Auswertungsstatistik” in der Dropdown-Liste “Position” die Dimension.

Wenn die kategoriale Variable beispielsweise in den Zeilen angezeigt wird, wählen Sie in der Dropdown-Liste die Option Zeilen.

Abbildung 2-9

Kategoriale Variable und Auswertungsstatistiken in derselben Dimension

Confidence in television	1 A great deal	Anzahl	196
	2 Only some	Anzahl	936
	3 Hardly any	Anzahl	744
	Gesamt	Anzahl	1876
		Median	2.29

### Anzeigeformate für Auswertungsstatistiken

Für die Anzeigeformate sind die folgenden Optionen verfügbar:

**nnnn.** Einfach numerisch.

**nnnn%.** Wert mit nachgestelltem Prozentzeichen.

**Auto.** Definiertes Anzeigeformat für die Variable (einschließlich der Dezimalstellen).

**N=nnnn.** Zeigt den Wert mit vorangestelltem  $N=$  an. Dies ist bei Häufigkeiten, *Gültige N* und *Gesamtanzahl* in Tabellen sinnvoll, in denen die Beschriftungen für Auswertungsstatistiken nicht angezeigt werden.

**(nnnn).** Alle Werte werden in Klammern angezeigt.

**(nnnn)(neg. value).** Nur negative Werte werden in Klammern angezeigt.

**(nnnn%).** Alle Werte werden in Klammern und mit nachgestelltem Prozentzeichen angezeigt.

**n,nnn.n.** Kommaformat. Unabhängig von den regionalen Einstellungen bzw. Ländereinstellungen von Windows wird das Komma als Tausender- und der Punkt als Dezimaltrennzeichen verwendet.

**n.nnn,n.** Punktformat. Unabhängig von den regionalen Einstellungen bzw. Ländereinstellungen von Windows wird der Punkt als Tausender- und das Komma als Dezimaltrennzeichen verwendet.

**\$n,nnn.n.** Dollarformat. Unabhängig von den regionalen Einstellungen bzw. Ländereinstellungen von Windows wird das Dollarzeichen dem Wert vorangestellt und das Komma als Tausender- sowie der Punkt als Dezimaltrennzeichen verwendet.

**CCA, CCB, CCC, CCD, CCE.** Währungsformate. In der Liste werden die aktuell definierten Formate jeder Währung angezeigt. Diese Formate werden im Dialogfeld “Optionen” (Menü “Bearbeiten”, Befehl “Optionen”) auf der Registerkarte “Währung” definiert.

### Allgemeine Regeln und Einschränkungen

- Außer beim Format “Auto” richtet sich die Anzahl der Dezimalstellen nach der Einstellung in der Spalte “Dezimalstellen”.

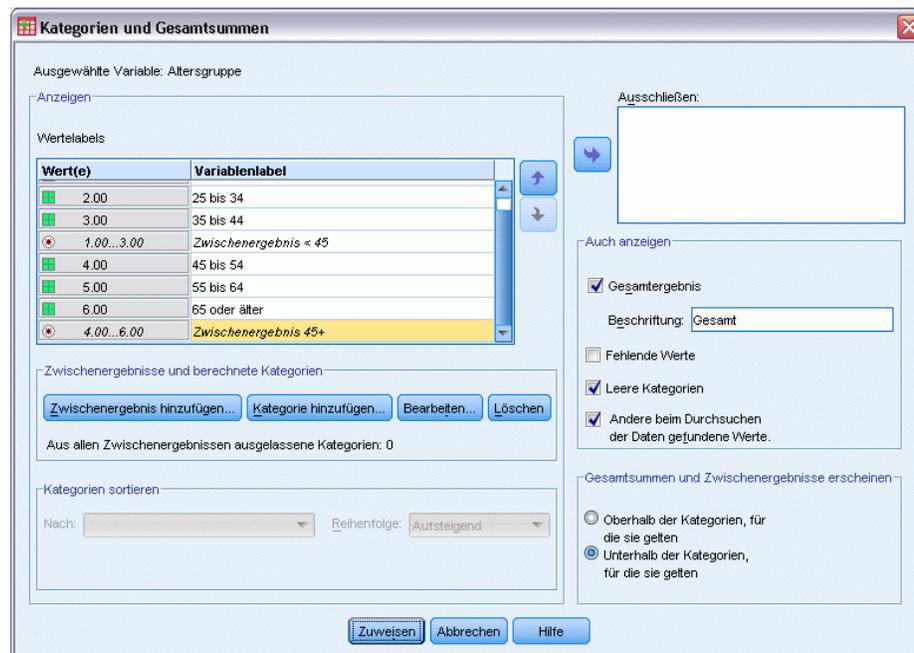
- Außer bei den Formaten “Komma”, “Dollar” und “Punkt” wird als Dezimaltrennzeichen das in der Systemsteuerung von Windows unter “Ländereinstellungen” bzw. “Regionale Einstellungen” im aktuellen Gebietschema festgelegte Zeichen verwendet.
- Bei den Formaten “Komma”, “Dollar” und “Punkt” wird als Tausendertrennzeichen ein Komma oder ein Punkt verwendet. Beim Konstruieren einer Tabelle ist jedoch kein Anzeigeformat verfügbar, mit dem das Tausendertrennzeichen aus dem aktuellen Gebietschema der regionalen Einstellungen bzw. Ländereinstellungen von Windows (Systemsteuerung, Option “Ländereinstellungen” bzw. “Regionale Einstellungen”) übernommen wird.

## Kategorien und Gesamtsummen

Das Dialogfeld “Kategorien und Gesamtsummen” bietet die folgenden Funktionen:

- Umordnen und Ausschließen von Kategorien
- Einfügen von Gesamtsummen und Zwischenergebnissen
- Einfügen von berechneten Kategorien.
- Ein- oder Ausschließen von leeren Kategorien
- Ein- oder Ausschließen von als fehlend definierten Kategorien
- Ein- oder Ausschließen von Kategorien, für die keine Wertelabels definiert wurden

Abbildung 2-10  
Dialogfeld “Kategorien und Gesamtsummen”



- Dieses Dialogfeld ist nur für kategoriale Variablen und Mehrfachantworten-Sets verfügbar. Es ist nicht für metrische Variablen verfügbar.

- Bei mehreren ausgewählten Variablen mit unterschiedlichen Kategorien können keine Zwischenergebnisse und berechneten Kategorien eingefügt, Kategorien ausgeschlossen oder Kategorien manuell umgeordnet werden. Dies ist aber nur dann der Fall, wenn Sie in der Vorschau auf der Zeichenfläche mehrere Variablen auswählen und anschließend das Dialogfeld für alle ausgewählten Variablen öffnen. Diese Vorgänge können jedoch für jede Variable einzeln durchgeführt werden.
- Für Variablen ohne Wertelabels können nur Kategorien sortiert und Gesamtsummen eingefügt werden.

***So öffnen Sie das Dialogfeld "Kategorien und Gesamtsummen":***

- ▶ Ziehen Sie eine kategoriale Variable bzw. ein Mehrfachantworten-Set auf die Zeichenfläche, und legen Sie diese bzw. dieses dort ab.
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf die Variable und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.

*ODER*

- ▶ Wählen Sie die Variable auf der Zeichenfläche aus, indem Sie darauf klicken, und klicken Sie anschließend auf der Registerkarte "Tabelle" im Gruppenfeld "Definieren" auf Kategorien und Gesamtsummen.

Sie können außerdem auf der Zeichenfläche mehrere kategoriale Variablen in derselben Dimension auswählen:

- ▶ Klicken Sie bei gedrückter STRG-Taste auf der Zeichenfläche auf die Variablen.

*ODER*

- ▶ Klicken Sie auf der Zeichenfläche auf einen Bereich außerhalb der Tabellenvorschau. Wählen Sie anschließend durch Klicken und Ziehen einen Bereich aus, der die gewünschten Variablen enthält.

*ODER*

- ▶ Klicken Sie mit der rechten Maustaste auf eine Variable in einer Dimension und wählen Sie die Option Alle [Dimension]variablen auswählen, um alle Variablen in dieser Dimension auszuwählen.

***So ordnen Sie Kategorien um:***

So ordnen Sie Kategorien manuell um:

- ▶ Wählen Sie in der Liste eine Kategorie aus, indem Sie darauf klicken.
- ▶ Klicken Sie auf die nach oben oder die nach unten weisende Pfeilschaltfläche, um die Kategorie in der Liste nach oben oder unten zu verschieben.

*ODER*

- ▶ Klicken Sie auf die Spalte "Wert(e)" der Kategorie, und verändern Sie durch Ziehen und Ablegen deren Position.

**So schließen Sie Kategorien aus:**

- ▶ Wählen Sie in der Liste eine Kategorie aus, indem Sie darauf klicken.
- ▶ Klicken Sie auf die Pfeilschaltfläche neben der Liste “Ausschließen”.

*ODER*

- ▶ Klicken Sie auf die Spalte “Wert(e)” der Kategorie, und ziehen Sie sie aus der Liste.

Wenn Sie eine Kategorie ausschließen, werden alle Kategorien, für die keine Wertelabels definiert sind, ebenfalls ausgeschlossen.

**So sortieren Sie Kategorien:**

Kategorien können nach Datenwerten, Wertelabels, Zellenhäufigkeit oder Auswertungsstatistik in absteigender oder aufsteigender Reihenfolge sortiert werden.

- ▶ Wählen Sie im Gruppenfeld “Kategorien sortieren” in der Dropdown-Liste “Nach” das gewünschte Sortierkriterium aus: Wert, Label, Anzahl oder Auswertungsstatistik (z. B. Mittelwert, Median, Modalwert). Welche Auswertungsstatistik für die Sortierung verfügbar ist, hängt davon ab, welche Auswertungsstatistik Sie für die Anzeige in der Tabelle ausgewählt haben.
- ▶ Wählen Sie in der Dropdown-Liste “Reihenfolge” die Sortierreihenfolge (aufsteigend, absteigend) aus.

Wenn Sie Kategorien ausgeschlossen haben, ist die Funktion zum Sortieren von Kategorien nicht verfügbar.

**Zwischenergebnisse**

- ▶ Klicken Sie in der Liste auf die letzte Kategorie, die im Zwischenergebnis berücksichtigt werden soll.
- ▶ Klicken Sie auf Zwischenergebnis hinzufügen....
- ▶ Ändern Sie im Dialogfeld “Zwischenergebnis definieren” bei Bedarf den Bezeichnungstext des Zwischenergebnisses.
- ▶ Um nur ein Zwischenergebnis anzuzeigen und die Anzeige der Kategorien, durch die das Zwischenergebnis definiert wird, zu unterdrücken, wählen Sie Zwischenergebniskategorien aus Tabelle ausblenden.
- ▶ Klicken Sie auf Weiter, um das Zwischenergebnis hinzuzufügen.

**Gesamt**

- ▶ Klicken Sie auf das Kontrollkästchen Gesamt. Sie können außerdem den Beschriftungstext für die Gesamtsumme ändern.

Wenn die ausgewählte Variable innerhalb einer anderen Variablen verschachtelt ist, werden Gesamtsummen für jede Untertabelle eingefügt.

### **Anzeigeoption für Gesamtsummen und Zwischenergebnisse**

Gesamtsummen und Zwischenergebnisse können über oder unter den in der jeweiligen Gesamtsumme enthaltenen Kategorien angezeigt werden.

- Wenn im Gruppenfeld “Gesamtsummen und Zwischenergebnisse erscheinen” die Option Unter ausgewählt ist, werden Gesamtsummen über jeder Untertabelle angezeigt, und bei der Berechnung der Zwischenergebnisse werden alle Kategorien über und einschließlich der ausgewählten Kategorie (aber unter allen vorhergehenden Zwischenergebnissen) berücksichtigt.
- Wenn im Gruppenfeld “Gesamtsummen und Zwischenergebnisse erscheinen” die Option Über ausgewählt ist, werden Gesamtsummen unter jeder Untertabelle angezeigt, und bei der Berechnung der Zwischenergebnisse werden alle Kategorien unter und einschließlich der ausgewählten Kategorie (aber über allen vorhergehenden Zwischenergebnissen) berücksichtigt.

*Wichtiger Hinweis:* Die Anzeigeoption für die Zwischenergebnisse sollte vor dem Definieren der Zwischenergebnisse ausgewählt werden. Wenn Sie die Anzeigeoption ändern, hat dies Auswirkungen auf alle Zwischenergebnisse (nicht nur auf das gegenwärtig ausgewählte). Außerdem werden *die in den Zwischenergebnissen enthaltenen Kategorien geändert*.

### **Berechnete Kategorien**

Sie können Kategorien anzeigen, die aus Auswertungsstatistiken, Gesamtergebnissen, Zwischenergebnissen und/oder Konstanten berechnet wurden. [Für weitere Informationen siehe Thema Berechnete Kategorien auf S. 46.](#)

### **Angepaßte Auswertungsstatistiken für Gesamtsummen und Zwischenergebnisse**

Mithilfe des Dialogfelds “Auswertungsstatistik” können Sie neben Gesamtsummen auch andere Statistiken in den Tabellenbereich für Gesamtsummen und Zwischenergebnisse aufnehmen. [Für weitere Informationen siehe Thema Auswertungsstatistiken für kategoriale Variablen auf S. 36.](#)

*Hinweis:* Wenn Sie mehrere benutzerdefinierte Gesamtstatistiken auswählen, die sich auch im Tabellenkörper befinden, und die Beschriftungen für die Statistiken ausblenden, werden die Gesamtsummen in derselben Reihenfolge sortiert, in der sie sich im Tabellenkörper befinden. Da die Beschriftungen ausgeblendet sind, können Sie möglicherweise nicht erkennen, was jede dieser Gesamtstatistiken darstellt. Im Allgemeinen empfiehlt es sich nicht, mehrere Statistiken auszuwählen und die Beschriftungen für die Statistiken auszublenden.

### **Gesamtsummen, Zwischenergebnisse und ausgeschlossenen Kategorien**

Fälle in ausgeschlossenen Kategorien werden bei der Berechnung der Gesamtsummen nicht berücksichtigt.

### ***Fehlende Werte, leere Kategorien und Werte ohne Wertelabels***

**Fehlende Werte.** Hiermit wird die Anzeige von **benutzerdefiniert fehlenden** Werten bzw. von Werten gesteuert, die als fehlende Werte definiert wurden (z. B. der Code "99", der für die Schwangerschaft bei Männern "nicht zutreffend" angibt). In der Standardeinstellung werden benutzerdefiniert fehlende Werte ausgeschlossen. Wählen Sie diese Option aus (aktivieren Sie diese), um benutzerdefiniert fehlende Werte in die Tabellen aufzunehmen. Obwohl die Variable mehrere Kategorien mit fehlenden Werten enthalten kann, wird in der Tabellenvorschau auf der Zeichenfläche nur eine generische Kategorie für fehlende Werte angezeigt. Alle benutzerdefiniert fehlenden Kategorien werden in die Tabelle aufgenommen. **Systemdefiniert fehlende Werte**(leere Zelle für numerische Variablen im Daten-Editor) werden immer ausgeschlossen.

**Leere Kategorien.** Bei leeren Kategorien handelt es sich um Kategorien mit definierten Wertelabels, die aber in der betreffenden Kategorie keine Fälle für eine bestimmte Tabelle oder Untertabelle enthalten. In der Standardeinstellung werden leere Kategorien in Tabellen aufgenommen. Deaktivieren Sie diese Option, um fehlende Kategorien aus der Tabelle auszuschließen.

**Andere beim Durchsuchen der Daten gefundene Werte.** In der Standardeinstellung werden Kategoriewerte in der Datendatei ohne definierte Wertelabels automatisch in die Tabelle aufgenommen. Deaktivieren Sie diese Option, um Werte ohne definierte Wertelabels aus der Tabelle auszuschließen. Wenn Sie Kategorien mit definierten Wertelabels ausschließen, werden auch Kategorien ohne definierte Wertelabels ausgeschlossen.

### ***Berechnete Kategorien***

Neben aggregierten Ergebnissen von Auswertungsstatistiken kann eine Tabelle auch eine oder mehrere Kategorien anzeigen, die aus diesen aggregierten Ergebnissen berechnet wurden: aus konstanten Werten, Zwischenergebnissen, Gesamtsummen oder einer Kombination davon. Die Ergebnisse werden als berechnete Kategorien oder Postcomputes bezeichnet. Eine berechnete Kategorie verhält sich wie eine Kategorie in einer einzigen Variablen, wobei es folgende Ähnlichkeiten und Unterschiede gibt:

- Eine berechnete Kategorie ist wie die anderen Kategorien positioniert.
- Eine berechnete Kategorie nutzt dieselben Statistiken wie die anderen Kategorien.
- Berechnete Kategorien wirken sich nicht auf Zwischenergebnisse, Gesamtsummen oder Signifikanztests aus.
- Standardmäßig verwenden die Werte berechneter Kategorien dieselbe Formatierung für Auswertungsstatistiken wie die anderen Kategorien. Sie können das Format bei der Definition der berechneten Kategorie überschreiben.

Da berechnete Kategorien verwendet werden können, um aggregierte Ergebnisse zu summieren, können sie Zwischenergebnissen ähneln. Berechnete Kategorien haben jedoch im Vergleich zu Zwischenergebnissen folgende Vorteile:

- Berechnete Kategorien können aus den Ergebnissen anderer Zwischenergebnisse errechnet werden.
- Berechnete Kategorien können sich gegenseitig überschneiden, da sie dieselben (oder einige derselben) Kategorien nutzen.

- Berechnete Kategorien müssen nicht die Werte aller anderen Kategorien oberhalb oder unterhalb der berechneten Kategorie einschließen. Das heißt, dass berechnete Kategorien nicht abschließend sind.
- Berechnete Kategorien können Werte aus nicht benachbarten Kategorien enthalten.

Im Gegensatz zu Gesamtsummen und Zwischenergebnissen werden berechnete Kategorien anhand der aggregierten Daten berechnet, nicht anhand der ursprünglichen Daten. Deshalb ist es möglich, dass die Werte berechneter Kategorien nicht mit den Ergebnissen von Gesamtsummen und Zwischenergebnissen übereinstimmen. Da Sie außerdem die Möglichkeit haben, Quellkategorien bei der Definition der berechneten Kategorie auszublenden, könnte es schwierig sein, die Zwischenergebnisse in der daraus resultierenden Tabelle zu interpretieren. Wenn Sie berechnete Kategorien verwenden, empfiehlt es sich, benutzerdefinierte Beschriftungen für Zwischenergebnisse anzugeben.

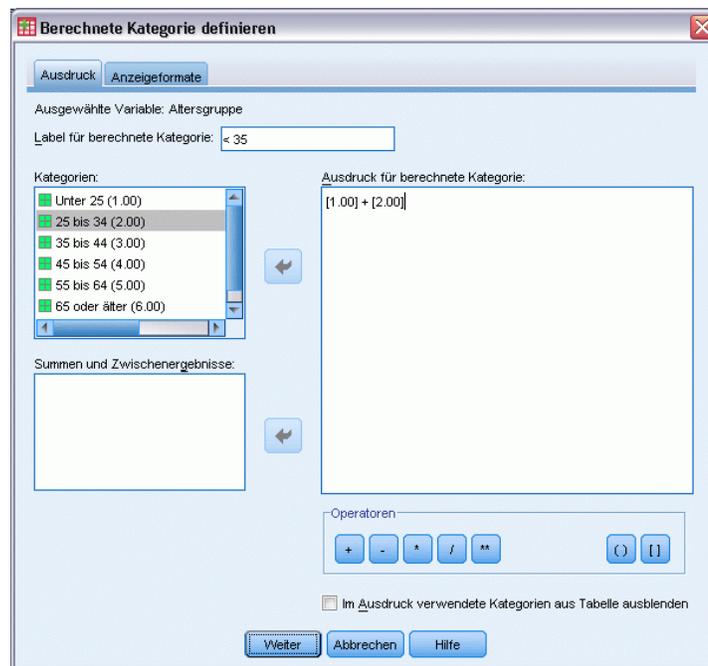
### **So definieren Sie eine berechnete Kategorie:**

Berechnete Kategorien werden anhand des Dialogfelds “Kategorien und Gesamtsummen” hinzugefügt. Weitere Informationen über den Zugriff auf dieses Dialogfeld finden Sie unter [Kategorien und Gesamtsummen](#) auf S. 42.

- Klicken Sie im Dialogfeld “Kategorien und Gesamtsummen” auf Kategorie hinzufügen....

Abbildung 2-11

Dialogfeld “Berechnete Kategorie definieren”



- Geben Sie unter Label für berechnete Kategorie eine Beschriftung für die berechnete Kategorie an. Sie können Kategorien aus der Liste “Kategorien” ziehen, um Beschriftungen für sie einzuschließen.

- ▶ Erstellen Sie einen Ausdruck, indem Sie Kategorien und/oder Gesamtsummen und Zwischenergebnisse auswählen und mithilfe von Operatoren die berechneten Kategorien definieren. Sie können auch konstante Werte (z. B. 500) in den Ausdruck aufnehmen.
- ▶ Um nur eine berechnete Kategorie anzuzeigen und die Anzeige der Kategorien, durch die die berechnete Kategorie definiert wird, zu unterdrücken, wählen Sie im Ausdruck verwendete Kategorien aus Tabelle ausblenden.
- ▶ Klicken Sie auf die Registerkarte **Anzeigeformate**, um das Anzeigeformat und die Anzahl an Dezimalstellen für die berechnete Kategorie zu ändern. [Für weitere Informationen siehe Thema \*\*Anzeigeformate für berechnete Kategorien\*\* auf S. 48.](#)
- ▶ Klicken Sie auf **Weiter**, um die berechnete Kategorie hinzuzufügen.

### ***Anzeigeformate für berechnete Kategorien***

Standardmäßig verwendet eine berechnete Kategorie das gleiche Anzeigeformat und die gleiche Anzahl an Dezimalstellen wie die andere Kategorie in der Variable. Sie können diese Einstellungen auf der Registerkarte "Anzeigeformate" im Dialogfeld "Berechnete Kategorie" überschreiben. Die Registerkarte "Anzeigeformate" enthält neben den Anzeigeformaten und der Anzahl an Dezimalstellen für diese Statistik auch eine Liste der aktuellen Auswertungsstatistik, die die berechnete Kategorie nutzt.

Sie haben in jeder Auswertungsstatistik folgende Möglichkeiten:

- ▶ Wählen Sie aus der Dropdown-Liste "Format" ein Anzeigeformat für die Auswertungsstatistik aus. Eine vollständige Liste der Anzeigeformate finden Sie im Thema [Anzeigeformate für Auswertungsstatistiken](#) auf S. 41.
- ▶ Geben Sie in der Zelle "Dezimalstellen" die Anzahl der anzuzeigenden Dezimalstellen für die ausgewählte Auswertungsstatistik ein.

### ***Tabellen für Variablen mit gemeinsamen Kategorien (Comperimeter-Tabellen)***

Oft enthalten Umfragen viele Fragen mit übereinstimmenden möglichen Antworten. Sie können diese zusammenhängenden Variablen stapeln, damit sie in einer Tabelle dargestellt werden, und die gemeinsamen Antwortkategorien können als Spalten der Tabelle erscheinen.

***So erstellen Sie eine Tabelle für mehrere Variablen mit gemeinsamen Kategorien:***

- ▶ Ziehen Sie die kategorialen Variablen aus der Variablenliste auf die Zeichenfläche, und legen Sie sie im Zeilenbereich ab. Die Variablen sollten **gestapelt** sein. [Für weitere Informationen siehe Thema \*\*Stapeln von Variablen\*\* auf S. 31.](#)
- ▶ Wählen Sie in der Dropdown-Liste "Kategorieposition" die Option **Zeilenlabels in Spalten**.

Abbildung 2-12  
Gestapelte Variablen mit gemeinsamen Antwortkategorien in Spalten

	Sehr viel	Mäßig	Fast kein
Vertrauen in die Medizin	844	864	167
Vertrauen in Banken & Geldinstitute	490	1068	306
Vertrauen in das Fernsehen	196	936	744
Vertrauen in die Bildung	511	1055	315
Vertrauen in die Presse	176	878	808
Vertrauen in Großkonzerne	500	1078	243

Für weitere Informationen siehe Thema Tabellen für Variablen mit gemeinsamen Kategorien in Kapitel 7 auf S. 111.

### Anpassen des Dialogfelds zum Konstruieren von Tabellen

Im Gegensatz zu herkömmlichen Dialogfeldern kann die Größe des Dialogfelds zum Konstruieren von Tabellen wie bei Fenstern geändert werden.

- Klicken Sie auf den oberen, unteren oder einen der seitlichen Ränder oder eine beliebige Ecke des Dialogfelds zum Konstruieren von Tabellen, und ziehen Sie, um die Größe zu ändern.

Darüber hinaus können Sie auf der Registerkarte "Tabelle" die Größe der Variablenliste, der Liste "Kategorien" und der Zeichenfläche ändern.

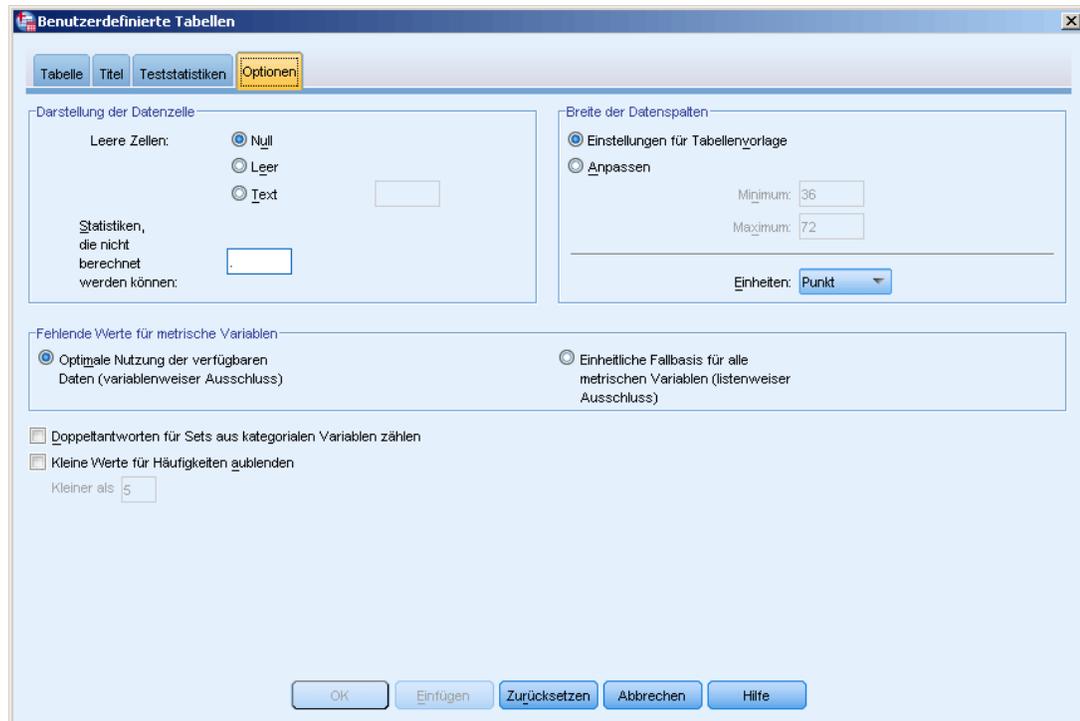
- Klicken Sie auf den horizontalen Balken zwischen der Variablenliste und der Liste "Kategorien", und ziehen Sie diesen, um einen längeren oder kürzeren Listenausschnitt anzuzeigen. Wenn Sie den Balken nach unten ziehen, wird der angezeigte Ausschnitt der Variablenliste länger und der Ausschnitt der Liste "Kategorien" kürzer. Wenn Sie den Balken nach oben ziehen, geschieht das Gegenteil.
- Klicken Sie auf den vertikalen Balken zwischen der Variablenliste und der Liste "Kategorien" auf der einen und der Zeichenfläche auf der anderen Seite, und ziehen Sie diesen, um die Listenausschnitte breiter oder schmaler zu machen. Die Größe der Zeichenfläche wird automatisch entsprechend dem verbleibenden Platz angepaßt.

### Benutzerdefinierte Tabellen: Registerkarte "Optionen"

Die Registerkarte "Optionen" bietet die folgenden Funktionen:

- Festlegen der Anzeige von leeren Zellen und Zellen, für die keine Statistiken berechnet werden können
- Steuern der Behandlung von fehlenden Werten bei der Berechnung von Statistiken für metrische Variablen
- Festlegen der minimalen und/oder maximalen Breite der Datenspalten
- Steuern der Behandlung von Doppeltantworten in Sets aus kategorialen Variablen

Abbildung 2-13  
Dialogfeld "Benutzerdefinierte Tabellen," Registerkarte "Optionen"



**Darstellung der Datenzelle.** Legt die Anzeige von leeren Zellen und Zellen fest, für die keine Statistiken berechnet werden können.

- **Leere Zellen.** Für Tabellenzellen, die keine Fälle enthalten (Zellenhäufigkeit 0), kann eine der folgenden drei Anzeigoptionen ausgewählt werden: "Null", "Leer" oder ein von Ihnen eingegebener Text. Der Text kann bis zu 255 Zeichen enthalten.
- **Statistiken, die nicht berechnet werden können.** Text, der angezeigt wird, wenn eine Statistik nicht berechnet werden kann (z. B. der Mittelwert für eine Kategorie ohne Fälle). Der Text kann bis zu 255 Zeichen enthalten. Der Standardwert ist ein Punkt (.).

**Breite der Datenspalten.** Steuert die minimale und maximale Breite der Datenspalten. Diese Einstellungen haben keinen Einfluß auf die Spaltenbreite der Zeilenbeschriftung.

- **Einstellungen für Tabellenvorlage.** Verwendet die in der aktuelle Standard-Tabellenvorlage. Sie können angepaßte Standardtabellenvorlagen zur Verwendung mit neuen Tabellen erstellen sowie die Breite von Spalten für die Zeilenbeschriftung und Datenspalten mit einer Tabellenvorlage steuern.
- **Benutzerdefiniert.** Überschreibt die Einstellungen der Standardtabellenvorlage für die Breite der Datenspalten. Legt die minimale und maximale Breite der Datenspalten für die Tabelle sowie die Maßeinheit fest: Punkt, Zoll oder Zentimeter.

**Fehlende Werte für metrische Variablen.** Steuert die Behandlung von fehlenden Daten bei der Statistik für metrische Variablen in Tabellen mit zwei oder mehr metrischen Variablen.

- **Optimale Nutzung der verfügbaren Daten (variablenweiser Ausschluss).** Alle Fälle mit gültigen Werten für die einzelnen metrischen Variablen werden bei den Auswertungsstatistiken für diese metrischen Variablen berücksichtigt.
- **Einheitliche Fallbasis für alle metrischen Variablen (listenweiser Ausschluss).** Fälle mit fehlenden Werten für metrische Variablen in der Tabelle werden bei den Auswertungsstatistiken für alle metrischen Variablen in der Tabelle ausgeschlossen.

**Doppeltantworten für Sets aus kategorialen Variablen zählen.** Bei einer Doppeltantwort wurde für zwei oder mehr Variablen in einem Set aus kategorialen Variablen die gleiche Antwort gegeben. In der Standardeinstellung werden Doppeltantworten nicht gezählt. Diese können aber durchaus gültige Antworten sein, die berücksichtigt werden müssen (z. B. in einem Set aus kategorialen Variablen, das die Hersteller der letzten drei Autos darstellt, die der Befragte erworben hat).

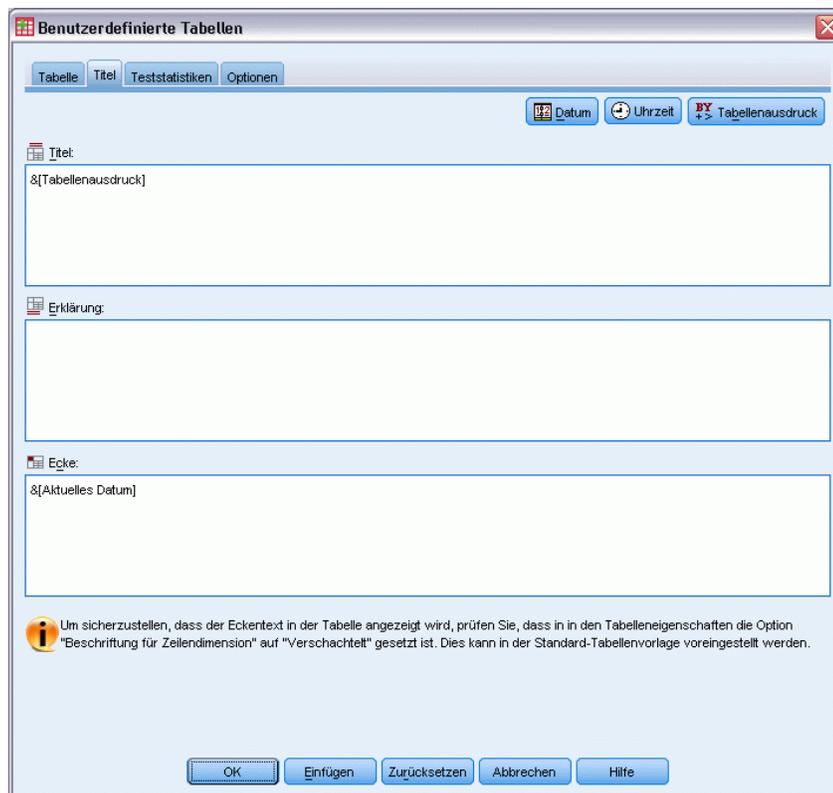
**Kleine Werte für Häufigkeiten ausblenden.** Sie können festlegen, dass Häufigkeiten ausgeblendet werden, wenn sie einen bestimmten ganzzahligen Wert unterschreiten. Ausgeblendete Werte werden als <N angezeigt. Dabei ist N die angegebene Ganzzahl. Die angegebene Ganzzahl muss größer oder gleich 2 sein.

## Benutzerdefinierte Tabellen: Registerkarte "Titel"

Auf der Registerkarte "Titel" wird die Anzeige von Titeln, Erklärungen und Eckenbeschriftungen gesteuert.

Abbildung 2-14

Dialogfeld "Benutzerdefinierte Tabellen"; Registerkarte "Titel"



**Titel.** Text, der über der Tabelle angezeigt wird.

**Erklärung.** Text, der unter der Tabelle und, sofern vorhanden, über den Fußnoten angezeigt wird.

**Ecke.** Text, der links oben in der Ecke der Tabelle angezeigt wird. Der Eckentext wird nur dann angezeigt, wenn die Tabelle Zeilenvariablen enthält und die Eigenschaft “Beschriftung für Zeilendimension” für die Pivot-Tabelle auf Verschachtelt gesetzt ist. In der Standardtabellenvorlage ist dies *nicht* der Fall.

Die folgenden automatisch erstellten Werte können in den Tabellentitel, die Erklärung oder den Eckenbeschriftung aufgenommen werden:

**Datum.** Aktuelles Datum mit Jahr, Monat und Tag, angezeigt in dem Format, das in den regionalen Einstellungen bzw. Ländereinstellungen von Windows festgelegt ist.

**Uhrzeit.** Aktuelle Uhrzeit mit Stunde, Minute und Sekunde, angezeigt in dem Format, das in den regionalen Einstellungen bzw. Ländereinstellungen von Windows festgelegt ist.

**Tabellenausdruck.** In der Tabelle verwendete Variablen und deren Verwendung in der Tabelle. Wenn eine Variable über ein definiertes Variablenlabel verfügt, wird das Label angezeigt. Die folgenden Symbole zeigen in der erstellten Tabelle an, wie Variablen in dieser verwendet werden.

- + zeigt gestapelte Variablen an.
- > zeigt Verschachtelung an.
- **Nach** zeigt Kreuztabellen oder Schichten an.

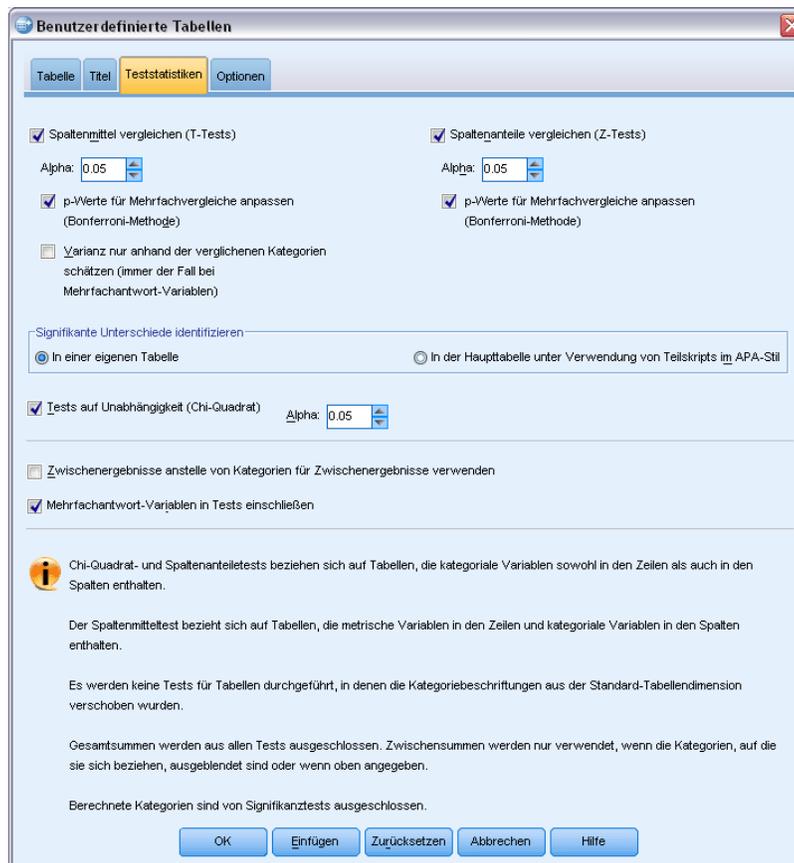
## ***Benutzerdefinierte Tabellen: Registerkarte “Teststatistiken”***

Auf der Registerkarte “Teststatistiken” können Sie verschiedene Signifikanztests für benutzerdefinierte Tabellen anfordern. Dazu gehören:

- Chi-Quadrat-Tests auf Unabhängigkeit
- Tests auf Gleichheit der Spaltenmittel
- Test auf Gleichheit der Spaltenanteile
- Signifikanztests für Mehrfachantworten-Sets und Zwischenergebnisse (Informationen zu Signifikanztests für Mehrfachantworten-Sets finden Sie unter [Signifikanztests bei Mehrfachantworten-Sets](#) auf S. 178.)

Diese Tests sind für Tabellen mit Kategorienbeschriftungen, die aus ihrer vorgegebenen Tabellendimension verschoben wurden, nicht verfügbar.

Abbildung 2-15  
Dialogfeld "Benutzerdefinierte Tabellen," Registerkarte "Teststatistiken"



**Spaltenmittel vergleichen (T-Tests).** Mit dieser Option werden paarweise Tests auf Gleichheit der Spaltenmittel für Tabellen erstellt, die mindestens eine kategoriale Variable in den Spalten und mindestens eine metrische Variable in den Zeilen enthalten. Die Tabelle muss den Mittelwert als Auswertungsstatistik beinhalten. Sie können festlegen, ob die  $p$ -Werte des Tests mit der Bonferroni-Methode korrigiert werden sollen. Sie können außerdem das Alpha-Niveau des Tests festlegen. Dieses muss einen Wert größer als 0 und kleiner als 1 aufweisen. Des Weiteren kann die Varianz, die beim Spaltenmitteltest stets ausschließlich auf den für Mehrfachantworten-Tests verglichenen Kategorien basiert, für reguläre kategoriale Variablen auf der Grundlage der verglichenen Kategorien oder aller Kategorien geschätzt werden.

**Spaltenanteile vergleichen (Z-Tests).** Mit dieser Option werden paarweise Tests auf Gleichheit von Spaltenanteilen für Tabellen erstellt, in denen sowohl in den Zeilen als auch in den Spalten mindestens eine kategoriale Variable vorhanden ist. Die Tabelle muss Häufigkeiten oder einfache Spaltenprozentage beinhalten. Sie können festlegen, ob die  $p$ -Werte des Tests mit der Bonferroni-Methode korrigiert werden sollen. Sie können außerdem das Alpha-Niveau des Tests festlegen. Dieses sollte einen Wert größer als 0 und kleiner als 1 aufweisen.

**Identifizieren signifikanter Differenzen.** Wenn Sie Spaltenmittel vergleichen oder Spaltenanteile vergleichen auswählen, können Sie festlegen, wie signifikante Differenzen angegeben werden.

- **In einer separaten Tabelle.** Ergebnisse von Signifikanztests werden in einer separaten Tabelle angezeigt. Falls zwei Werte deutlich voneinander abweichen, wird in der dem höheren Wert zugehörigen Tabelle ein Schlüssel angezeigt, der auf die Spalte mit dem kleineren Wert verweist. Hier ein Beispiel:

Abbildung 2-16  
Vergleich der Spaltenanteile in einer separaten Tabelle

		Familienstand				
		Verheiratet	Verwitwet	Geschieden	Getrennt lebend	Ledig
		(A)	(B)	(C)	(D)	(E)
Beschäftigungsstatus	Vollzeit	B		A B	B	B
	Teilzeit					A B C
	Zeitweise unbeschäftigt				. <sup>a</sup>	A B
	Arbeitslos, entlassen					A B
	Ruhestand	E	A C D E	E		
	Ausbildung					A B C
	Arbeit im Haushalt	C E	C E		C E	
	Andere					

- **In der Haupttabelle unter Verwendung von Teilskripts im APA-Stil.** In der Haupttabelle werden deutliche Abweichungen anhand von Formatierung im APA-Stil mit tiefgestellten Buchstaben gekennzeichnet. Falls zwei Werte deutlich voneinander abweichen, weisen diese Werte *unterschiedliche* tiefgestellte Buchstaben auf. Diese tiefgestellten Buchstaben sind keine Fußnoten. Ist diese Option aktiviert, wird der in der aktuellen Tabellenvorlage definierte Fußnoten-Stil überschrieben und tatsächliche Fußnoten werden als hochgestellte Zahlen angezeigt. Hier ein Beispiel:

Abbildung 2-17  
Vergleiche von Spaltenanteilen anhand von tiefgestellten Buchstaben im APA-Stil

		Familienstand				
		Verheiratet	Verwitwet	Geschieden	Getrennt	Nie verheiratet
		Anzahl der Spalten (%)				
Berufstätigkeit	Vollzeit berufstätig	57.8% <sub>a</sub>	15.5% <sub>b</sub>	66.1% <sub>c</sub>	62.4% <sub>a,c</sub>	59.1% <sub>a,c</sub>
	Teilzeit berufstätig	10.3% <sub>a</sub>	7.1% <sub>a</sub>	7.8% <sub>a</sub>	9.7% <sub>a,b</sub>	15.4% <sub>b</sub>
	Zeitweilig nicht berufstätig	1.7% <sub>a</sub>	.7% <sub>a</sub>	2.0% <sub>a</sub>	1.1% <sub>a</sub>	1.7% <sub>a</sub>
	Arbeitslos, entlassen	1.0% <sub>a</sub>	1.1% <sub>a</sub>	2.2% <sub>b</sub>	.0%	4.8% <sub>b</sub>
	Im Ruhestand	12.5% <sub>a</sub>	53.0% <sub>b</sub>	11.9% <sub>a</sub>	6.5% <sub>a,c</sub>	2.6% <sub>c</sub>
	Schule	.7% <sub>a</sub>	.4% <sub>a</sub>	1.6% <sub>a</sub>	2.2% <sub>a,b</sub>	9.0% <sub>b</sub>
	Haushalt	14.9% <sub>a</sub>	19.4% <sub>a</sub>	5.6% <sub>b</sub>	14.0% <sub>a</sub>	5.3% <sub>b</sub>
	Sonstiges	1.2% <sub>a</sub>	2.8% <sub>a</sub>	2.7% <sub>a</sub>	4.3% <sub>a</sub>	2.1% <sub>a</sub>

Das vollständige Beispiel, das zeigt, wie diese Tabellen erstellt und interpretiert werden können, finden Sie unter [Vergleichen der Spaltenanteile](#) auf S. 160.

*Hinweis:* Die Tabelle im APA-Stil beinhaltet eine Erklärung, die erläutert, wie die tiefgestellten Zeichen in der Tabelle zu interpretieren sind. Wenn Sie eine Erklärung auf der Registerkarte "Titel" angeben, wird diese Erklärung angezeigt und nicht die Erklärung für die tiefgestellten Zeichen.

**Tests auf Unabhängigkeit (Chi-Quadrat).** Mit dieser Option wird ein Chi-Quadrat-Test auf Unabhängigkeit für Tabellen erstellt, in denen sowohl in den Zeilen als auch in den Spalten mindestens eine kategoriale Variable vorhanden ist. Sie können außerdem das Alpha-Niveau des Tests festlegen. Dieses sollte einen Wert größer als 0 und kleiner als 1 aufweisen.

**Zwischenergebnisse anstelle von Kategorien für Zwischenergebnisse verwenden.** Wenn diese Option aktiviert ist, werden in Signifikanztests statt den Kategorien für Zwischenergebnisse die Zwischenergebnisse selbst verwendet. Andernfalls werden nur die für Zwischenergebnisse verwendeten ausgeblendeten Kategorien in den Tests durch die Zwischenergebnisse selbst ersetzt.

**Mehrfachantwort-Variablen in Tests einschließen.** Wenn diese Option aktiviert ist, werden Tests unter Verwendung der Kategorien von Mehrfachantworten-Sets durchgeführt. Andernfalls werden Mehrfachantworten-Sets beim Durchführen von Tests ignoriert.

# Einfache Tabellen für kategoriale Variablen

Die meisten Tabellen, die Sie erstellen, enthalten wahrscheinlich mindestens eine **kategoriale Variable**. Eine kategoriale Variable ist eine Variable mit einer bestimmten Anzahl von eindeutigen Werten bzw. Kategorien (beispielsweise Geschlecht oder Religion). Kategoriale Variablen können **nominal** oder **ordinal** sein.

- **Nominal**. Eine Variable kann als nominal behandelt werden, wenn ihre Kategorien sich nicht in eine natürliche Reihenfolge bringen lassen, z. B. die Firmenabteilung, in der eine Person arbeitet. Beispiele für nominale Variablen sind Region, Postleitzahl oder Religionszugehörigkeit.
- **Ordinal**. Eine Variable kann als ordinal behandelt werden, wenn ihre Werte für Kategorien stehen, die eine natürliche Reihenfolge aufweisen (z. B. Grad der Zufriedenheit mit Kategorien von sehr unzufrieden bis sehr zufrieden). Ordinale Variablen treten beispielsweise bei Einstellungsmessungen (Zufriedenheit oder Vertrauen) und bei Präferenzbeurteilungen auf.

Der Variablentyp ist durch ein Symbol neben der jeweiligen Variablen in der Variablenliste gekennzeichnet.

	Numerisch	Zeichenfolge	Datum	Zeit
Metrisch (stetig)		entfällt		
Ordinal				
Nominal				

Die Funktion zum Erstellen von benutzerdefinierten Tabellen wurde für kategoriale Variablen mit definierten **Wertelabels** optimiert. [Für weitere Informationen siehe Thema Konstruieren von Tabellen in Kapitel 2 auf S. 27.](#)

## Beispiel-Datendatei

Für die Beispiele in diesem Kapitel wird die Datendatei *survey\_sample.sav* verwendet. [Für weitere Informationen siehe Thema Beispieldateien in Anhang A auf S. 201.](#)

In allen angegebenen Beispielen werden die Variablenlabels in Dialogfeldern in alphabetischer Reihenfolge angezeigt. Die Anzeigeeigenschaften für Variablenlisten werden im Dialogfeld "Optionen" auf der Registerkarte "Allgemein" festgelegt (Menü "Bearbeiten", Befehl "Optionen").

## Eine einzelne kategoriale Variable

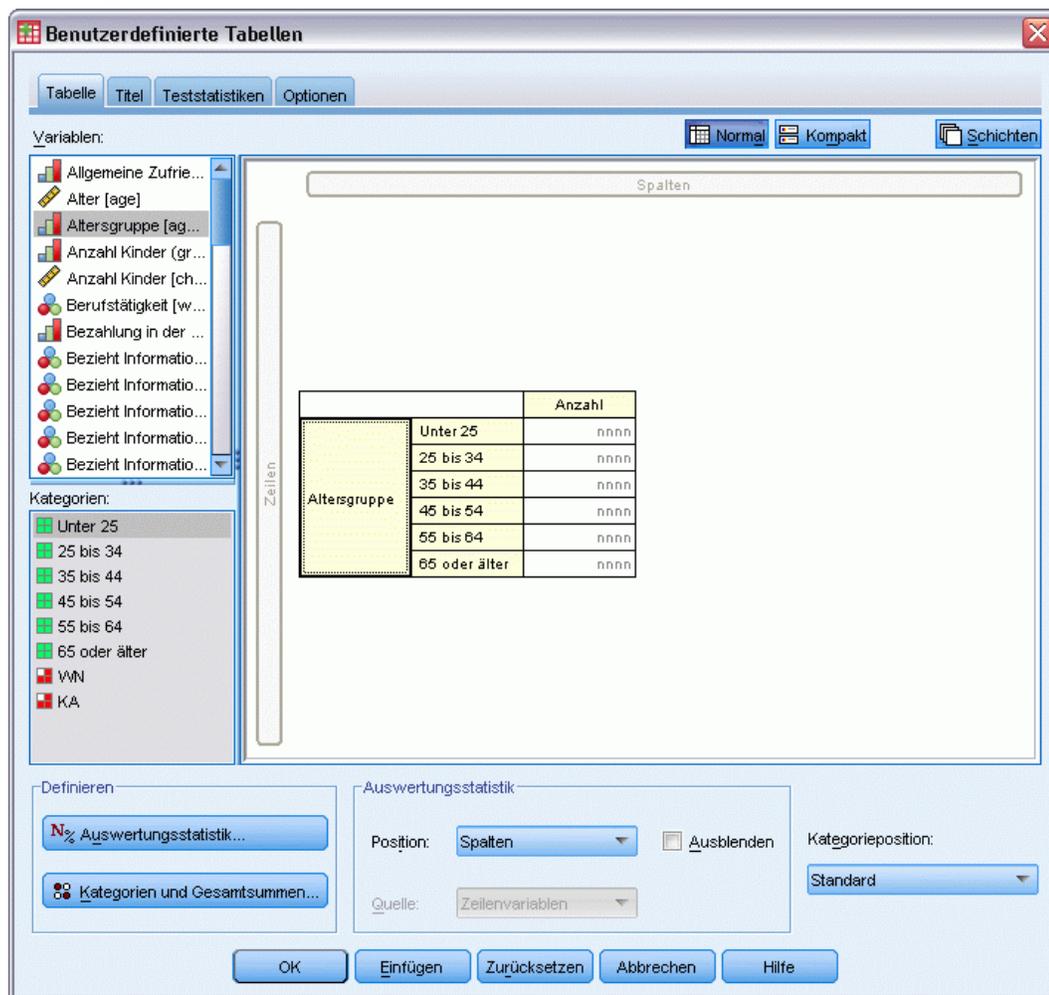
Eine Tabelle mit einer einzelnen kategorialen Variablen ist wahrscheinlich eine der einfachsten Tabellen, die Sie erstellen können. Oft werden Sie jedoch keine komplexere Tabelle benötigen.

- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
Analysieren > Tabellen > Benutzerdefinierte Tabellen...
- ▶ Ziehen Sie die Variable *Alterskategorie* aus der Variablenliste der Tabellenkonstruktion in den Bereich "Zeilen" auf der Zeichenfläche.

Auf der Zeichenfläche wird eine Vorschau der Tabelle angezeigt. In der Vorschau werden nicht die tatsächlichen Datenwerte angezeigt, sondern Platzhalter anstelle der eigentlichen Daten.

Abbildung 3-1

Eine kategoriale Variable in Zeilen in der Vorschau auf der Zeichenfläche der Tabellenkonstruktion



- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.  
Die Tabelle wird im Fenster des Viewers angezeigt.

**Abbildung 3-2**  
*Eine kategoriale Variable in Zeilen*

		Anzahl
Alterskategorie	Weniger als 25	242
	25 bis 34	627
	35 bis 44	679
	45 bis 54	481
	55 bis 64	320
	65 oder älter	479

In dieser einfachen Tabelle ist die Spaltenüberschrift *Anzahl* eigentlich nicht notwendig, und die Tabelle kann auch ohne diese Spaltenüberschrift erstellt werden.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen erneut (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Aktivieren Sie im Gruppenfeld “Auswertungsstatistik” neben der Dropdown-Liste “Position” die Option Ausblenden.
- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

**Abbildung 3-3**  
*Eine kategoriale Variable ohne Spaltenbeschriftung für die Auswertungsstatistik*

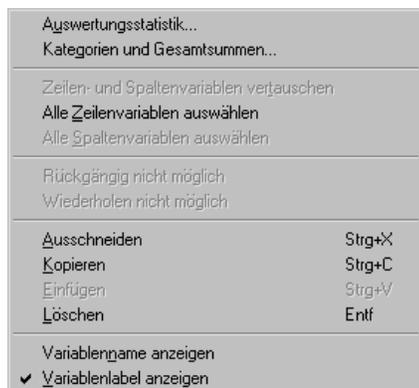
Alterskategorie	Weniger als 25	242
	25 bis 34	627
	35 bis 44	679
	45 bis 54	481
	55 bis 64	320
	65 oder älter	479

## **Prozente**

Zusätzlich zu Häufigkeiten können Sie auch Prozentwerte anzeigen lassen. Bei einer einfachen Tabelle mit einer einzelnen kategorialen Variablen, die in Zeilen angezeigt wird, werden wahrscheinlich die Spaltenprozentwerte interessant sein. Wenn die Variable hingegen in Spalten angezeigt wird, interessieren Sie wahrscheinlich eher die Zeilenprozentwerte.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen erneut (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Deaktivieren Sie im Gruppenfeld “Auswertungsstatistik” neben der Dropdown-Liste “Position” die Option Ausblenden. Da diese Tabelle zwei Spalten enthält, sollen die Spaltenbeschriftungen angezeigt werden, sodass Sie wissen, welche Größen in den Spalten dargestellt werden.
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option Auswertungsstatistik aus.

Abbildung 3-4  
Kontextmenü (rechte Maustaste) der Zeichenfläche



- ▶ Wählen Sie im Dialogfeld “Auswertungsstatistik” in der Liste “Statistik” die Option Anzahl als Spalten% aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” zu übertragen.
- ▶ Löschen Sie in der Liste “Anzeigen” in der Zelle mit der Beschriftung die Standardbeschriftung und geben Sie Prozent ein.

Abbildung 3-5  
Dialogfeld “Auswertungsstatistik” für kategoriale Variablen



- ▶ Klicken Sie auf Der Auswahl zuweisen und anschließend in der Tabellenkonstruktion auf OK, um die Tabelle zu erstellen.

Abbildung 3-6  
Häufigkeiten und Spaltenprozentwerte

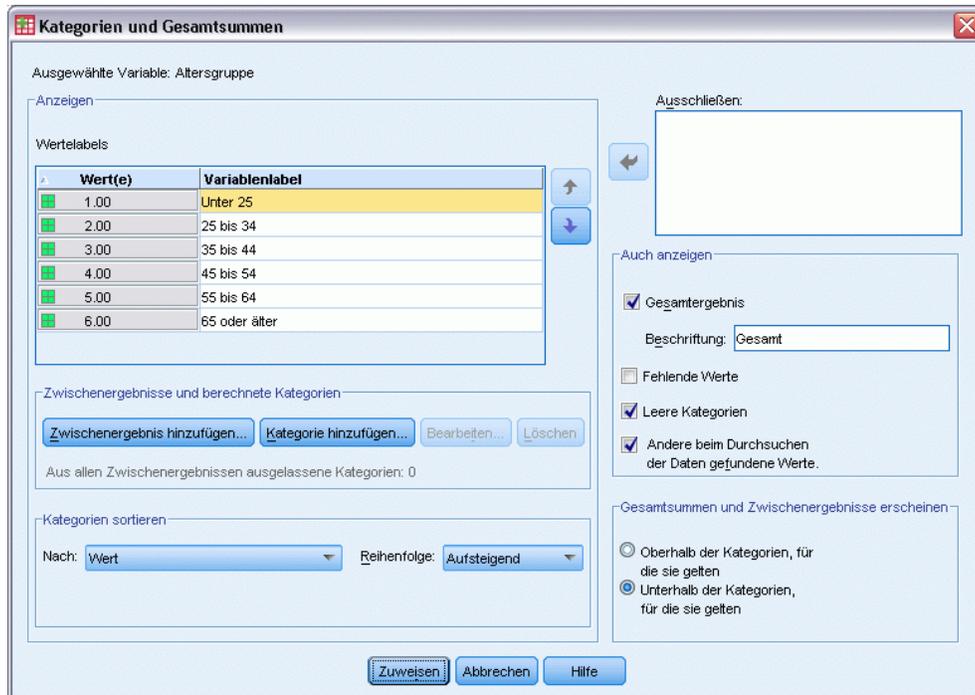
	Anzahl	Prozent
Alterskategorie		
Weniger als 25	242	8,6%
25 bis 34	627	22,2%
35 bis 44	679	24,0%
45 bis 54	481	17,0%
55 bis 64	320	11,3%
65 oder älter	479	16,9%

## Gesamt

Gesamtsummen werden nicht automatisch in benutzerdefinierte Tabellen aufgenommen. Gesamtsummen können einer Tabelle jedoch problemlos hinzugefügt werden.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen erneut (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.
- ▶ Wählen Sie im Dialogfeld “Kategorien und Gesamtsummen” die Option Gesamt aus, indem Sie darauf klicken.

Abbildung 3-7  
Dialogfeld “Kategorien und Gesamtsummen”



- ▶ Klicken Sie auf Zuweisen und anschließend in der Tabellenkonstruktion auf OK, um die Tabelle zu erstellen.

Abbildung 3-8  
Häufigkeiten, Spaltenprozentwerte und Gesamtsummen

		Anzahl	Prozent
Alterskategorie	Weniger als 25	242	8,6%
	25 bis 34	627	22,2%
	35 bis 44	679	24,0%
	45 bis 54	481	17,0%
	55 bis 64	320	11,3%
	65 oder älter	479	16,9%
	Gesamt	2828	100,0%

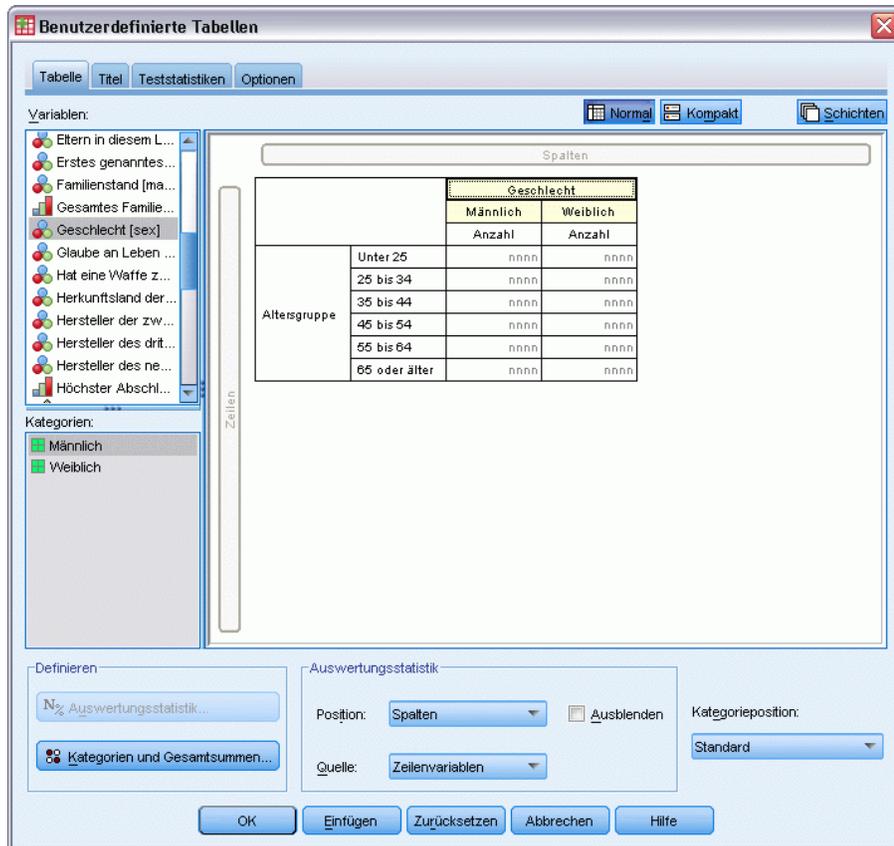
Für weitere Informationen siehe Thema Gesamtsummen und Zwischenergebnisse für kategoriale Variablen in Kapitel 5 auf S. 87.

## Kreuztabelle

Eine Kreuztabelle stellt ein grundlegendes Verfahren dar, mit dem die Beziehung zwischen zwei kategorialen Variablen untersucht werden kann. Wenn Sie beispielsweise *Alterskategorie* als Zeilenvariable und *Geschlecht* als Spaltenvariable verwenden, können Sie eine zweidimensionale Kreuztabelle erstellen, mit der die Anzahl an Männern und Frauen in jeder Alterskategorie dargestellt wird.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie auf Zurücksetzen, um frühere Einstellungen in der Tabellenkonstruktion zu löschen.
- ▶ Ziehen Sie die Variable *Alterskategorie* aus der Variablenliste der Tabellenkonstruktion in den Bereich “Zeilen” auf der Zeichenfläche.
- ▶ Ziehen Sie die Variable *Geschlecht* aus der Variablenliste in den Spaltenbereich auf der Zeichenfläche. (Möglicherweise müssen Sie einen Bildlauf durch die Variablenliste durchführen, um diese Variable zu finden.)

Abbildung 3-9  
Kreuztabelle in der Vorschau auf der Zeichenfläche der Tabellenkonstruktion



- Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 3-10  
Kreuztabelle von Alterskategorie und Geschlecht

		Geschlecht	
		Männlich	Weiblich
		Anzahl	Anzahl
Alterskategorie	Weniger als 25	108	134
	25 bis 34	276	351
	35 bis 44	309	370
	45 bis 54	221	260
	55 bis 64	136	184
	65 oder älter	178	301
Gesamt		1228	1600

### Prozentwerte in Kreuztabellen

In einer zweidimensionalen Kreuztabelle sind möglicherweise sowohl die Zeilen- als auch die Spaltenprozentwerte nützliche Informationen.

- Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen erneut (Menü "Analysieren", Untermenü "Tabellen", Befehl "Benutzerdefinierte Tabellen").

- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Geschlecht*.  
Möglicherweise fällt Ihnen hierbei auf, dass die Option Auswertungsstatistik im Kontextmenü deaktiviert ist. Der Grund hierfür ist, dass Sie Auswertungsstatistiken nur für die innerste Variable in der Quelldimension der Statistik auswählen können. Die Standard-Quelldimension der Statistik (Zeile oder Spalte) für kategoriale Variablen beruht auf der Reihenfolge, in der Sie die Variablen auf die Zeichenfläche ziehen. In diesem Beispiel wurde *Alterskategorie* zuerst in die Zeilendimension gezogen. Da dort keine weiteren Variablen vorliegen, ist *Alterskategorie* die Quellvariable der Statistik. Sie können die Quelldimension der Statistik ändern. In diesem Beispiel ist dies jedoch nicht nötig. [Für weitere Informationen siehe Thema Auswertungsstatistik in Kapitel 2 auf S. 35.](#)
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option Auswertungsstatistik aus.
- ▶ Wählen Sie im Dialogfeld “Auswertungsstatistik” in der Liste “Statistik” die Option Anzahl als Spalten% aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” zu übertragen.
- ▶ Wählen Sie in der Liste “Statistik” die Option Anzahl als Zeilen% aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” zu übertragen.
- ▶ Klicken Sie auf Der Auswahl zuweisen und anschließend in der Tabellenkonstruktion auf OK, um die Tabelle zu erstellen.

Abbildung 3-11  
Kreuztabelle mit Zeilen- und Spaltenprozentwerten

		Geschlecht					
		Männlich			Weiblich		
		Anzahl	Zeilen%	Spalten%	Anzahl	Zeilen%	Spalten%
Alterskategorie	Weniger als 25	108	44,6%	8,8%	134	55,4%	8,4%
	25 bis 34	276	44,0%	22,5%	351	56,0%	21,9%
	35 bis 44	309	45,5%	25,2%	370	54,5%	23,1%
	45 bis 54	221	45,9%	18,0%	260	54,1%	16,3%
	55 bis 64	136	42,5%	11,1%	184	57,5%	11,5%
	65 oder älter	178	37,2%	14,5%	301	62,8%	18,8%
	Gesamt	1228	43,4%	100,0%	1600	56,6%	100,0%

## Steuern des Anzeigeformats

Sie können das Anzeigeformat einschließlich der Anzahl der in Auswertungsstatistiken angezeigten Dezimalstellen steuern. Prozentwerte werden in der Standardeinstellung beispielsweise mit einer Dezimalstelle und einem Prozentzeichen angezeigt. Möglicherweise möchten Sie die Zellenwerte jedoch mit zwei Dezimalstellen aber ohne Prozentzeichen darstellen.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option Auswertungsstatistik aus.

- Wählen Sie für die beiden ausgewählten Auswertungsstatistiken mit Prozentwerten (Anzahl als Spalten% und Anzahl als Zeilen%) in der Dropdown-Liste “Format” die Option nnn.n aus und geben Sie für beide in der Zelle “Dezimalstellen” den Wert 2 ein.

Abbildung 3-12  
Dialogfeld “Auswertungsstatistik”



- Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 3-13  
Formatierte Zellenanzeige für Zeilen- und Spaltenprozentwerte

		Geschlecht					
		Männlich			Weiblich		
		Anzahl	Zeilen%	Spalten%	Anzahl	Zeilen%	Spalten%
Alterskategorie	Weniger als 25	108	44,63%	8,79%	134	55,37%	8,38%
	25 bis 34	276	44,02%	22,48%	351	55,98%	21,94%
	35 bis 44	309	45,51%	25,16%	370	54,49%	23,13%
	45 bis 54	221	45,95%	18,00%	260	54,05%	16,25%
	55 bis 64	136	42,50%	11,07%	184	57,50%	11,50%
	65 oder älter	178	37,16%	14,50%	301	62,84%	18,81%

## Randsummen

In Kreuztabellen sollen häufig **Randsummen** (Gesamtsummen für jede Zeile und Spalte) angezeigt werden. Da diese in der Standardeinstellung nicht in benutzerdefinierten Tabellen enthalten sind, müssen sie den Tabellen ausdrücklich hinzugefügt werden.

- Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- Klicken Sie auf Zurücksetzen, um frühere Einstellungen in der Tabellenkonstruktion zu löschen.
- Ziehen Sie die Variable *Alterskategorie* aus der Variablenliste der Tabellenkonstruktion in den Bereich “Zeilen” auf der Zeichenfläche.

- ▶ Ziehen Sie die Variable *Geschlecht* aus der Variablenliste in den Spaltenbereich auf der Zeichenfläche. (Möglicherweise müssen Sie einen Bildlauf durch die Variablenliste durchführen, um diese Variable zu finden.)
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.
- ▶ Wählen Sie im Dialogfeld “Kategorien und Gesamtsummen” die Option Gesamt aus, indem Sie darauf klicken, und klicken Sie anschließend auf Zuweisen.
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Geschlecht* und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.
- ▶ Wählen Sie im Dialogfeld “Kategorien und Gesamtsummen” die Option Gesamt aus, indem Sie darauf klicken, und klicken Sie anschließend auf Zuweisen.
- ▶ Aktivieren Sie im Gruppenfeld “Auswertungsstatistik” neben der Dropdown-Liste “Position” die Option Ausblenden. (Da nur Häufigkeiten angezeigt werden, muss die in den Datenzellen der Tabelle angezeigte “Statistik” nicht angegeben werden.)
- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 3-14  
Kreuztabelle mit Randsummen

		Geschlecht		
		Männlich	Weiblich	Gesamt
Alterskategorie	Weniger als 25	108	134	242
	25 bis 34	276	351	627
	35 bis 44	309	370	679
	45 bis 54	221	260	481
	55 bis 64	136	184	320
	65 oder älter	178	301	479
	Gesamt	1228	1600	2828

## Sortieren und Ausschließen von Kategorien

In der Standardeinstellung werden die Kategorien in aufsteigender Reihenfolge der Datenwerte, die durch die Wertelabels der Kategorien dargestellt werden, sortiert. Obwohl für Alterskategorien beispielsweise die Wertelabels *Unter 25*, *25 bis 34*, *35 bis 44* usw. angezeigt werden, lauten die tatsächlich zugrunde liegenden Datenwerte 1, 2, 3 usw. Die Standard-Anzeigereihenfolge der Kategorien wird durch diese zugrunde liegenden Datenwerte festgelegt.

Die Reihenfolge der Kategorien kann leicht geändert werden, und Kategorien, die in der Tabelle nicht angezeigt werden sollen, können ausgeschlossen werden.

### Sortieren von Kategorien

Sie können Kategorien manuell neu anordnen oder Kategorien in aufsteigender oder absteigender Reihenfolge nach folgenden Elementen sortieren:

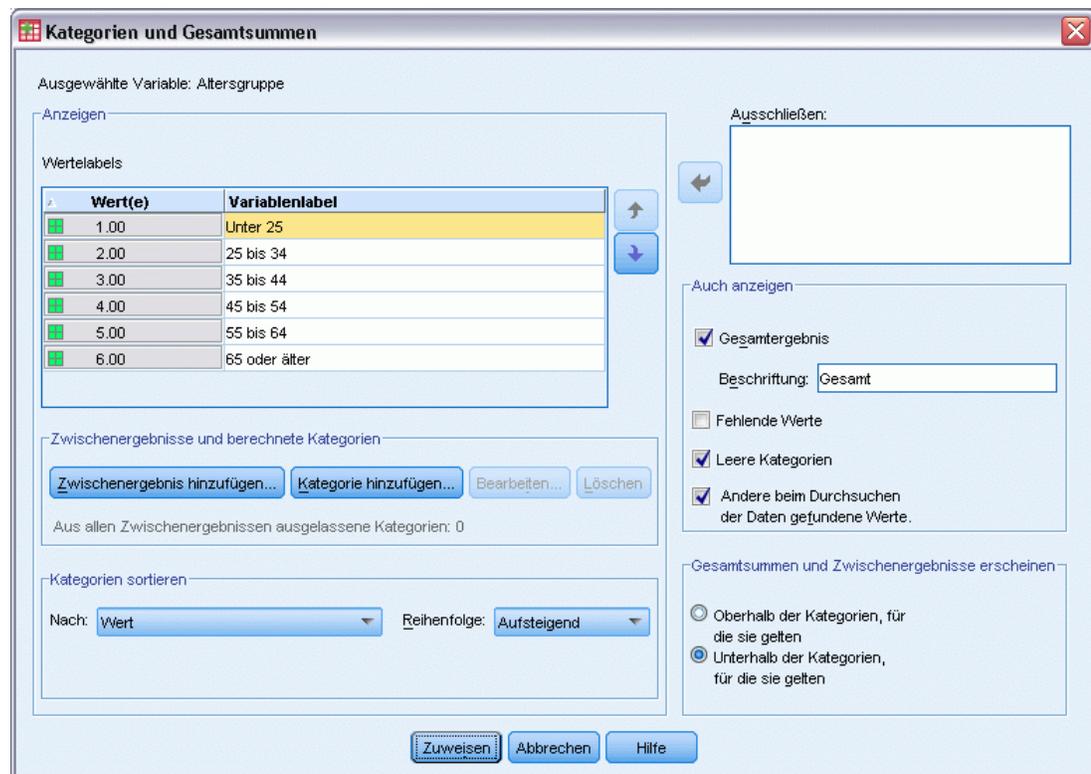
- Datenwerte.
- Wertelabels.

- Zellenhäufigkeit.
- Auswertungsstatistik. Welche Auswertungsstatistik für die Sortierung verfügbar ist, hängt davon ab, welche Auswertungsstatistik Sie für die Anzeige in der Tabelle ausgewählt haben.
- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Wenn die Variable *Alterskategorie* nicht bereits auf der Zeichenfläche im Bereich für die Zeilen angezeigt wird, ziehen Sie sie dorthin und legen sie dort ab.
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.

Sowohl die Datenwerte als auch die damit verbundenen Wertelabels werden in der aktuellen Sortierreihenfolge angezeigt. In diesem Fall handelt es hierbei noch um die aufsteigende Reihenfolge der Datenwerte.

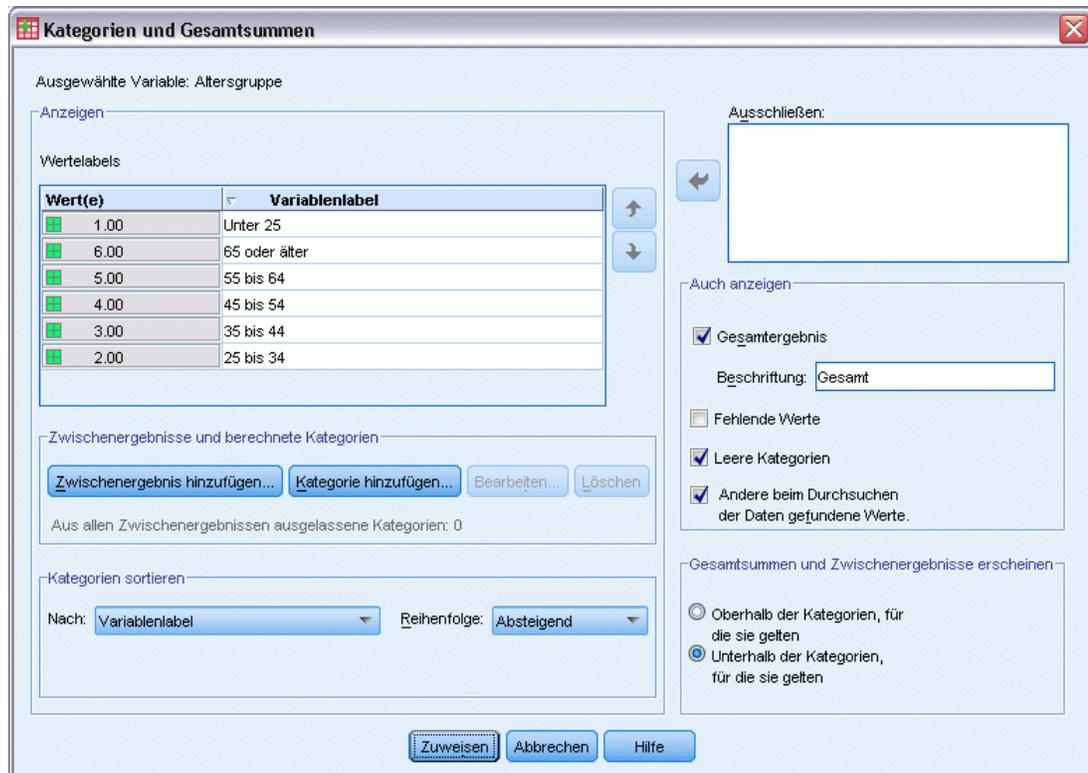
Abbildung 3-15

Standard-Kategorienreihenfolge (aufsteigend nach Datenwerten)



- ▶ Wählen Sie im Gruppenfeld “Kategorien sortieren” in der Dropdown-Liste “Reihenfolge” die Option Absteigend aus.  
Die Sortierreihenfolge wurde nun umgekehrt.
- ▶ Wählen Sie in der Dropdown-Liste “Nach” die Option Labels aus.  
Die Kategorien werden nun in absteigender alphabetischer Reihenfolge der Wertelabels sortiert.

Abbildung 3-16  
Absteigende alphabetische Sortierreihenfolge

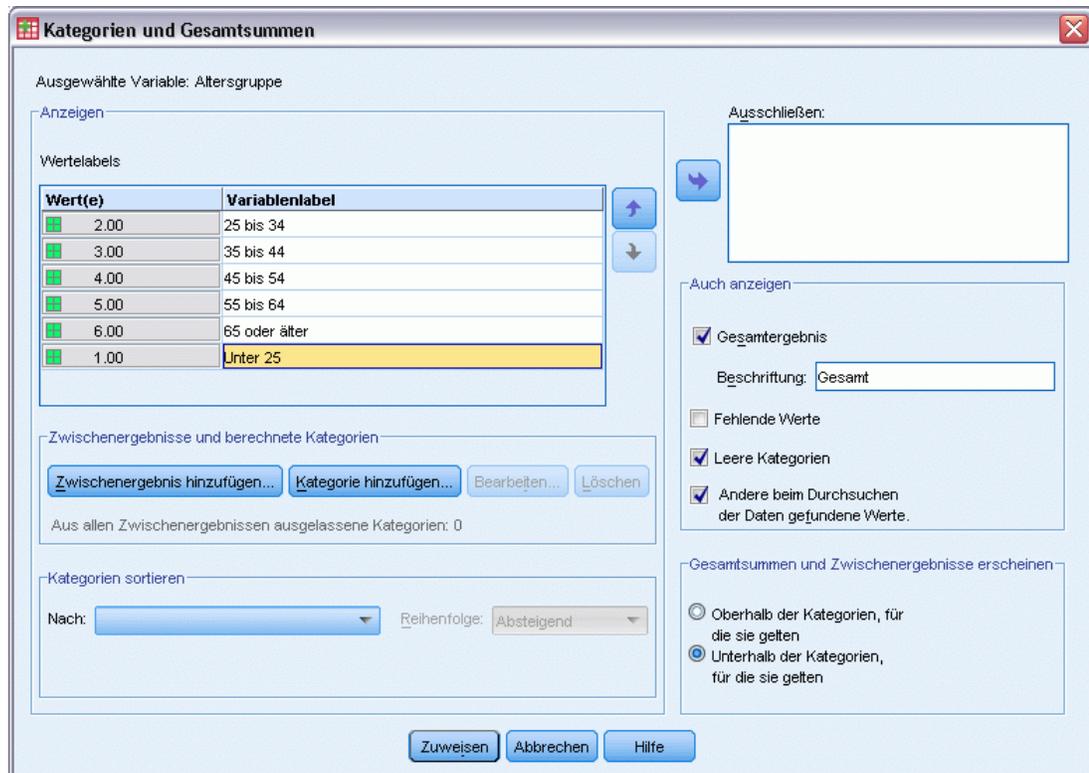


Beachten Sie, dass sich die Kategorie mit der Beschriftung *Unter 25* am Anfang der Liste befindet. Bei der alphabetischen Reihenfolge werden Buchstaben nach Ziffern angezeigt. Da dies die einzige Beschriftung ist, die mit einem Buchstaben anfängt, und die Liste in absteigender (umgekehrter) Reihenfolge sortiert wird, wird diese Kategorie am Anfang der Liste angezeigt.

Wenn eine bestimmte Kategorie an anderer Stelle in der Liste angezeigt werden soll, können Sie sie leicht dorthin verschieben.

- ▶ Klicken Sie in der Liste "Label" auf die Kategorie mit der Beschriftung *Unter 25*.
- ▶ Klicken Sie auf den Abwärtspfeil rechts neben der Liste. Die Kategorie wird nun in der Liste um eine Zeile nach unten verschoben.
- ▶ Klicken Sie so oft auf den Abwärtspfeil, bis sich die Kategorie am Ende der Liste befindet.

Abbildung 3-17  
Manuell angeordnete Kategorien



### **Ausschließen von Kategorien**

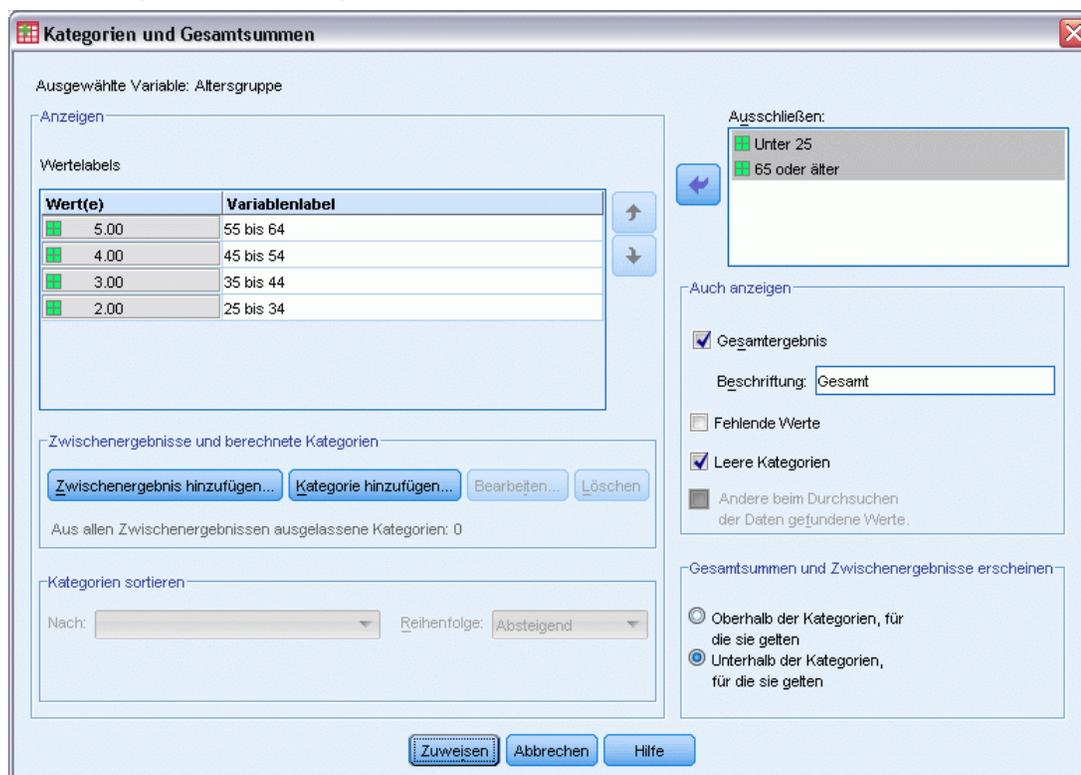
Kategorien, die nicht in der Tabelle angezeigt werden sollen, können ausgeschlossen werden.

- ▶ Klicken Sie in der Liste "Label" auf die Kategorie mit der Beschriftung *Unter 25*.
- ▶ Klicken Sie auf die Pfeilschaltfläche links neben der Liste "Ausschließen".
- ▶ Klicken Sie in der Liste "Label" auf die Kategorie mit der Beschriftung *65 oder älter*.

- Klicken Sie erneut auf die Pfeilschaltfläche links neben der Liste “Ausschließen”.

Daraufhin werden die beiden Kategorien aus der Liste “Anzeigen” in die Liste “Ausschließen” verschoben. Wenn Sie dies rückgängig machen möchten, können Sie sie leicht wieder in die Liste “Anzeigen” aufnehmen.

Abbildung 3-18  
Manuell ausgeschlossene Kategorien



- Klicken Sie auf Zuweisen und anschließend in der Tabellenkonstruktion auf OK, um die Tabelle zu erstellen.

Abbildung 3-19  
Nach absteigendem Wertelabel sortierte Tabelle mit einigen ausgeschlossenen Kategorien

		Geschlecht		
		Männlich	Weiblich	Gesamt
Alterskategorie	25 bis 34	276	351	627
	35 bis 44	309	370	679
	45 bis 54	221	260	481
	55 bis 64	136	184	320
	Gesamt	942	1165	2107

Beachten Sie, dass die Gesamtsummen niedriger als vor dem Ausschluss der beiden Kategorien sind. Dies liegt daran, dass Gesamtsummen auf den in der Tabelle enthaltenen Kategorien beruhen. Ausgeschlossene Kategorien werden auch aus der Berechnung der Gesamtsummen ausgeschlossen. [Für weitere Informationen siehe Thema Gesamtsummen und Zwischenergebnisse für kategoriale Variablen in Kapitel 5 auf S. 87.](#)

# ***Stapeln, Verschachteln und Schichten von kategorialen Variablen***

Stapeln, Verschachteln und Schichten sind Methoden, um in derselben Tabelle mehrere Variablen gleichzeitig anzuzeigen. In diesem Kapitel wird die Verwendung dieser Verfahren bei kategorialen Variablen behandelt. Die Verfahren können jedoch auch bei metrischen Variablen verwendet werden.

## ***Beispiel-Datendatei***

Für die Beispiele in diesem Kapitel wird die Datendatei *survey\_sample.sav* verwendet. [Für weitere Informationen siehe Thema Beispieldateien in Anhang A auf S. 201.](#)

In allen angegebenen Beispielen werden die Variablenlabels in Dialogfeldern in alphabetischer Reihenfolge angezeigt. Die Anzeigeeigenschaften für Variablenlisten werden im Dialogfeld “Optionen” auf der Registerkarte “Allgemein” festgelegt (Menü “Bearbeiten”, Befehl “Optionen”).

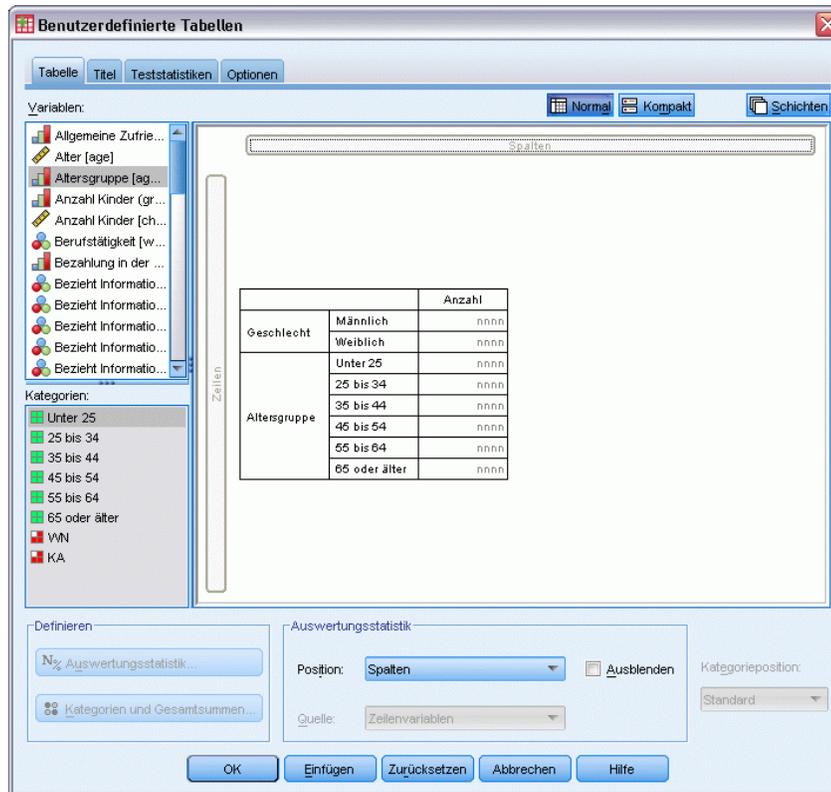
## ***Stapeln von kategorialen Variablen***

Beim Stapeln werden separate Tabellen in dieselbe Anzeige eingefügt. So können beispielsweise Informationen über das *Geschlecht* und die *Alterskategorie* in unterschiedlichen Abschnitten derselben Tabelle angezeigt werden.

- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
Analysieren > Tabellen > Benutzerdefinierte Tabellen...
- ▶ Ziehen Sie die Variable *Geschlecht* aus der Variablenliste der Tabellenkonstruktion in den Zeilenbereich auf der Zeichenfläche.
- ▶ Ziehen Sie die Variable *Alterskategorie* aus der Variablenliste, und legen Sie sie im Zeilenbereich unter *Geschlecht* ab.

Die beiden Variablen sind nun in der Zeilendimension gestapelt.

Abbildung 4-1  
Auf der Zeichenfläche werden gestapelte kategoriale Variablen angezeigt.



- Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 4-2  
Tabelle mit kategorialen Variablen, die in den Zeilen gestapelt sind

		Anzahl
Gender	Male	1232
	Female	1600
Age category	Less than 25	242
	25 to 34	627
	35 to 44	679
	45 to 54	481
	55 to 64	320
	65 or older	479

Auf die gleiche Weise können Variablen in den Spalten gestapelt werden.

### Stapeln bei Kreuztabellen

Eine gestapelte Tabelle kann andere Variablen in anderen Dimensionen enthalten. So können beispielsweise zwei in den Zeilen gestapelte Variablen mit einer dritten Variable, die in der Spaltendimension dargestellt ist, zu einer Kreuztabelle kombiniert werden.

- Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen erneut (Menü "Analysieren", Untermenü "Tabellen", Befehl "Benutzerdefinierte Tabellen").

- ▶ Falls *Alterskategorie* und *Geschlecht* nicht bereits in den Zeilen gestapelt sind, folgen Sie den vorangegangenen Anweisungen zum Stapeln von Variablen.
- ▶ Ziehen Sie die Variable *Nachrichten aus dem Internet* aus der Variablenliste auf die Zeichenfläche, und legen Sie im Spaltenbereich ab.
- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 4-3

Zwei gestapelte Zeilenvariablen wurden mit einer Spaltenvariablen zu einer Kreuztabelle kombiniert.

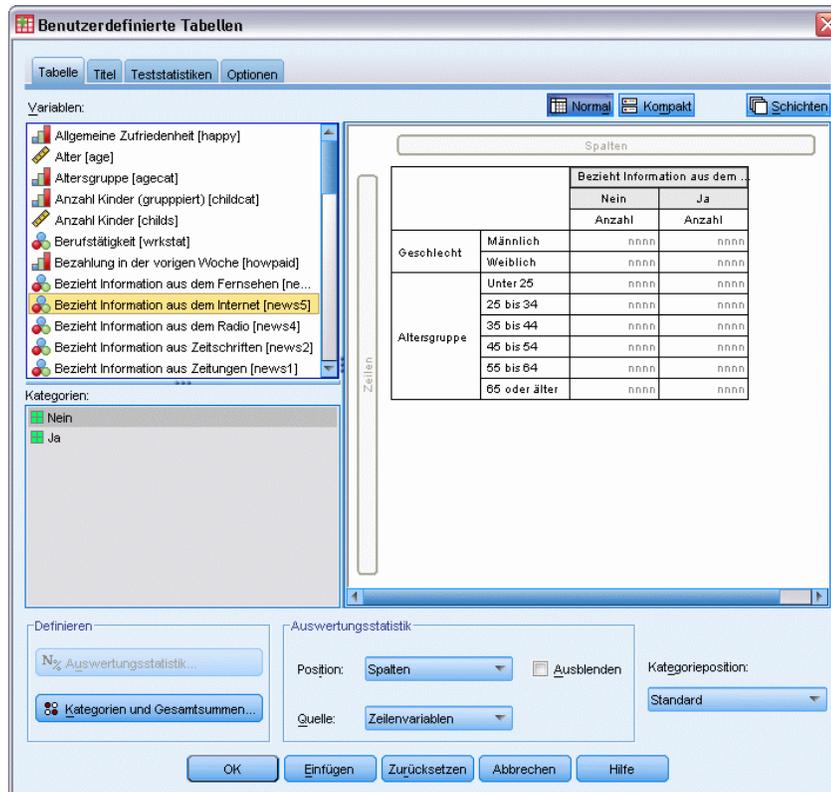
		Nachrichten aus dem Internet	
		Nein	Ja
		Anzahl	Anzahl
Geschlecht	Männlich	873	359
	Weiblich	1092	508
Alterskategorie	Weniger als 25	146	96
	25 bis 34	368	259
	35 bis 44	435	244
	45 bis 54	346	135
	55 bis 64	252	68
	65 oder älter	416	63

*Hinweis:* Es sind mehrere Variablen vorhanden, deren Labels mit dem Wort *Nachrichten* beginnen, sodass es schwierig sein könnte, diese in der Variablenliste zu unterscheiden (da die Labels für eine vollständige Anzeige in der Variablenliste möglicherweise zu lang sind). Es gibt zwei Möglichkeiten, um die Labels der Variablen vollständig anzuzeigen:

- Zeigen Sie mit dem Mauszeiger auf eine Variable in der Liste, um das vollständige Label als QuickInfo einzublenden.
- Klicken Sie auf den vertikalen Balken zwischen der Variablenliste und der Liste “Kategorien” auf der einen und der Zeichenfläche auf der anderen Seite, und ziehen Sie diesen, um die Listenausschnitte breiter zu machen.

Abbildung 4-4

Die Variablenliste wurde verbreitert, damit mehr von den Variablenlabels angezeigt wird.



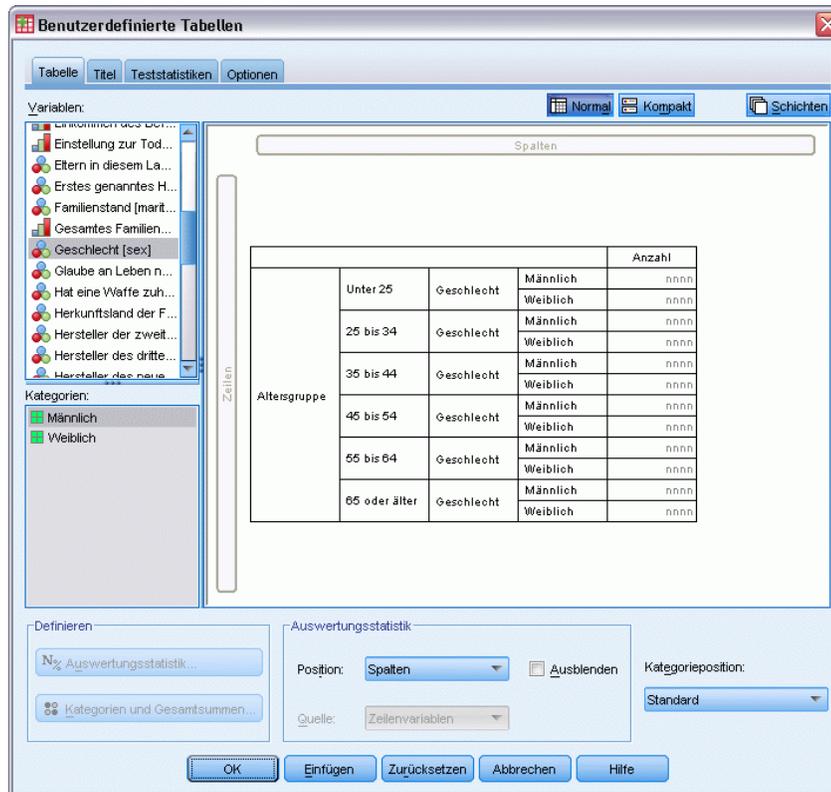
## Verschachteln von kategorialen Variablen

Ähnlich wie bei Kreuztabellen kann mit Verschachtelung die Beziehung zwischen zwei kategorialen Variablen aufgezeigt werden. Bei der Verschachtelung wird eine der Variablen jedoch in derselben Dimension innerhalb der anderen Variablen verschachtelt. So können Sie beispielsweise *Geschlecht* in *Alterskategorie* in der Zeilendimension verschachteln, um die Anzahl der Männer und Frauen in jeder Alterskategorie darzustellen.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü "Analysieren", Untermenü "Tabellen", Befehl "Benutzerdefinierte Tabellen").
- ▶ Klicken Sie auf Zurücksetzen, um frühere Einstellungen in der Tabellenkonstruktion zu löschen.
- ▶ Ziehen Sie die Variable *Alterskategorie* aus der Variablenliste der Tabellenkonstruktion in den Bereich "Zeilen" auf der Zeichenfläche.
- ▶ Ziehen Sie die Variable *Geschlecht* aus der Variablenliste, und legen Sie sie im Zeilenbereich rechts neben *Alterskategorie* ab.

In der Vorschau auf der Zeichenfläche wird nun eine verschachtelte Tabelle angezeigt, die eine einzelne Spalte mit Häufigkeiten enthält. Die Zellen der Spalte enthalten die Anzahl der Frauen und Männer in der jeweiligen Alterskategorie.

Abbildung 4-5  
Das Geschlecht ist in der Alterskategorie verschachtelt.



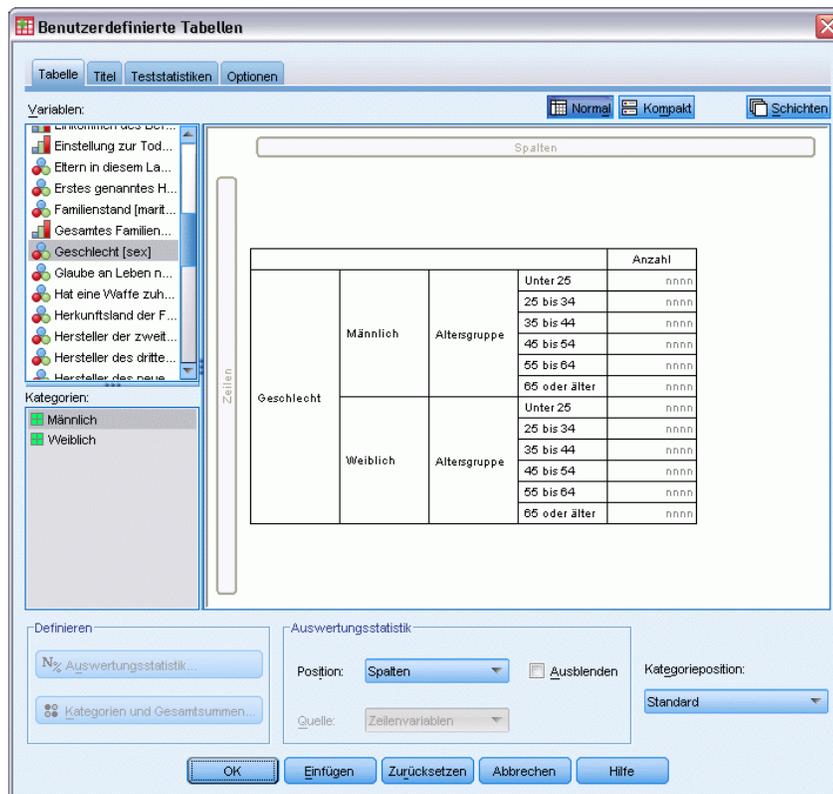
Es wird Ihnen auffallen, dass das Variablenlabel *Geschlecht* bei jeder Alterskategorie angezeigt wird. Solche Wiederholungen können reduziert werden, wenn Sie die Variable mit der geringsten Anzahl von Kategorien auf der äußersten Verschachtelungsebene anordnen.

- ▶ Klicken Sie auf der Zeichenfläche auf das Variablenlabel *Geschlecht*.
- ▶ Ziehen Sie die Variable, und legen Sie sie so weit links wie möglich im Zeilenbereich ab.

Das Label *Geschlecht* wird nun nicht mehr sechsmal wiederholt, *Alterskategorie* dafür zweimal. Diese Tabelle ist übersichtlicher und stellt im Wesentlichen die gleichen Ergebnisse dar.

Abbildung 4-6

In der Vorschau der Tabellenkonstruktion wird die Alterskategorie verschachtelt im Geschlecht gezeigt.



- Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 4-7

Tabelle der Alterskategorie verschachtelt im Geschlecht.

				Anzahl
Geschlecht	Männlich	Alterskategorie	Weniger als 25	108
			25 bis 34	276
			35 bis 44	309
			45 bis 54	221
			55 bis 64	136
			65 oder älter	178
	Weiblich	Alterskategorie	Weniger als 25	134
			25 bis 34	351
			35 bis 44	370
			45 bis 54	260
			55 bis 64	184
			65 oder älter	301

*Anmerkung:* Bei benutzerdefinierten Tabellen wird die Verarbeitung geschichteter aufgeteilter Dateien nicht berücksichtigt. Um dasselbe Ergebnis wie bei geschichteten aufgeteilten Dateien zu erzielen, müssen Sie die Dateiaufteilungs-Variablen in den äußersten Verschachtelungsebenen der Datei platzieren.

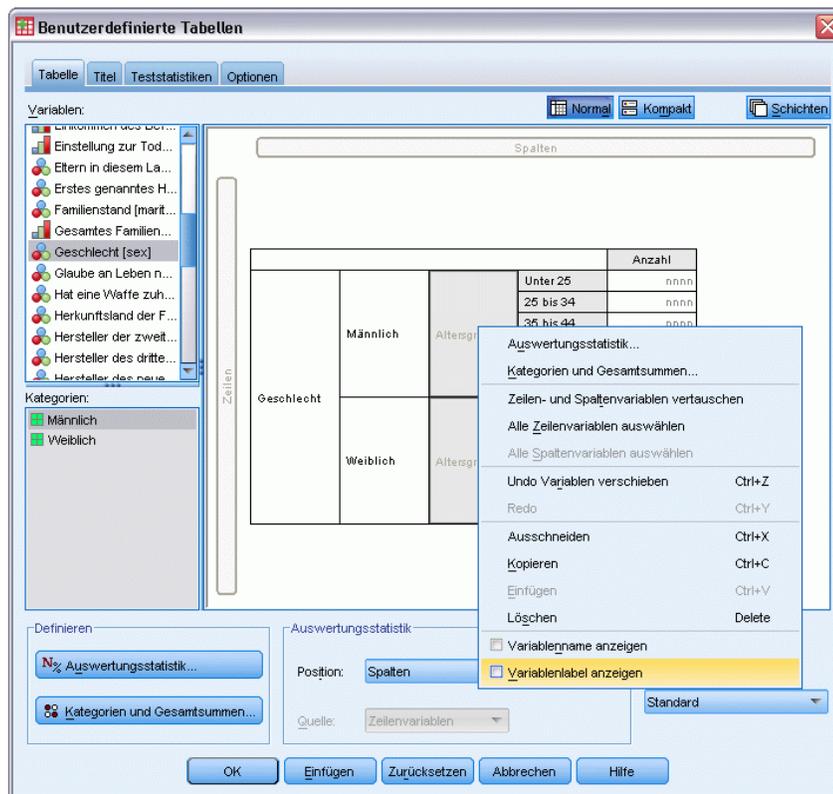
## Unterdrücken von Variablenlabels

Eine andere Möglichkeit zum Vermeiden von überflüssigen Variablenlabels in verschachtelten Tabellen besteht darin, die Anzeige von Variablennamen oder -labels einfach zu unterdrücken. Da die Wertelabels von *Geschlecht* und *Alterskategorie* selbsterklärend sein dürften, können die Labels für beide Variablen entfernt werden.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie*, und deaktivieren Sie im Kontextmenü die Option Variablenlabel anzeigen.
- ▶ Wiederholen Sie diesen Vorgang für *Geschlecht*.

Abbildung 4-8

Unterdrücken von Variablenlabels über das Kontextmenü im Dialogfeld zum Konstruieren von Tabellen



Die Variablenlabels werden weiterhin in der Tabellenvorschau angezeigt, sind aber nicht mehr in der Tabelle enthalten.

- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 4-9  
Verschachtelte Tabelle ohne Variablenlabels

		Anzahl
Männlich	Weniger als 25	108
	25 bis 34	276
	35 bis 44	309
	45 bis 54	221
	55 bis 64	136
	65 oder älter	178
Weiblich	Weniger als 25	134
	25 bis 34	351
	35 bis 44	370
	45 bis 54	260
	55 bis 64	184
	65 oder älter	301

Wenn die Variablenlabels in die Tabelle aufgenommen werden sollen, ohne dass diese mehrfach wiederholt im Tabellenkorpus angezeigt werden, können Sie sie im Tabellentitel oder in der Eckenbeschriftung unterbringen.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie auf die Registerkarte Titel.
- ▶ Klicken Sie im Textfeld “Titel” auf eine beliebige Stelle.
- ▶ Klicken Sie auf Tabellenausdruck. Im Textfeld “Titel” wird der Text  $\&[Tabellenausdruck]$  angezeigt. Dies ergibt einen Tabellentitel, der die Variablenlabels für die in der Tabelle verwendeten Variablen enthält.
- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 4-10  
Variablenlabels im Tabellentitel  
**Geschlecht > Alterskategorie**

		Anzahl
Männlich	Weniger als 25	108
	25 bis 34	276
	35 bis 44	309
	45 bis 54	221
	55 bis 64	136
	65 oder älter	178
Weiblich	Weniger als 25	134
	25 bis 34	351
	35 bis 44	370
	45 bis 54	260
	55 bis 64	184
	65 oder älter	301

Das Größer-als-Zeichen (>) im Titel zeigt an, daß *Alterskategorie* in *Geschlecht* verschachtelt ist.

## ***Verschachtelte Kreuztabellen***

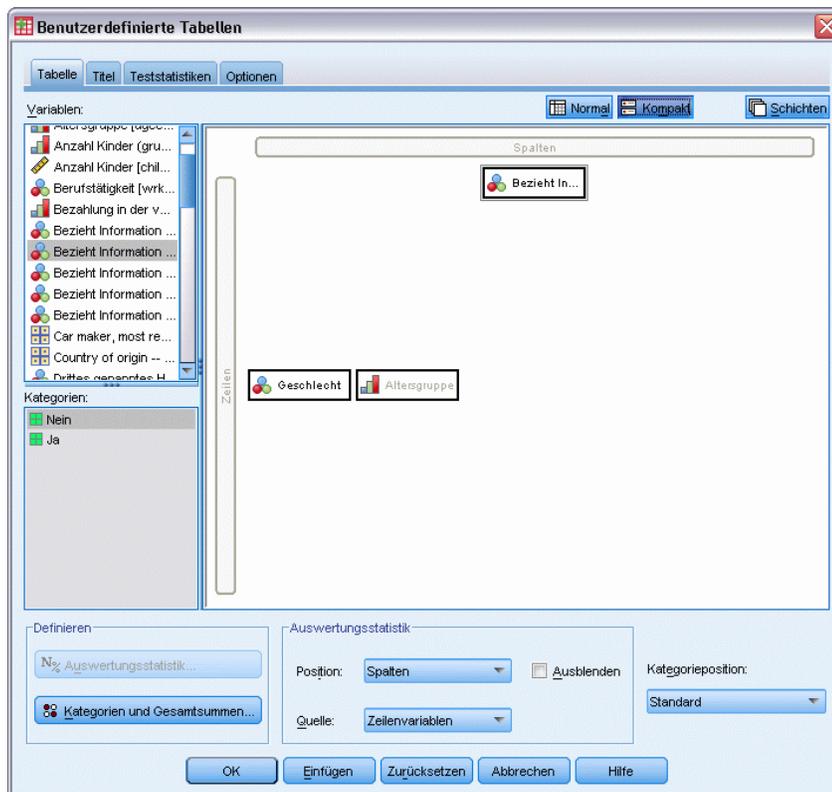
Eine verschachtelte Tabelle kann andere Variablen in anderen Dimensionen enthalten. Sie können beispielsweise *Alterskategorie* in *Geschlecht* in Zeilen verschachteln und diese verschachtelten Zeilen mit einer dritten Variablen in der Spaltendimension zu einer Kreuztabelle kombinieren.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Falls *Alterskategorie* nicht bereits in den Zeilen in *Geschlecht* verschachtelt ist, folgen Sie den vorangegangenen Anweisungen zum Verschachteln von Variablen.
- ▶ Ziehen Sie die Variable *Nachrichten aus dem Internet* aus der Variablenliste auf die Zeichenfläche, und legen Sie im Spaltenbereich ab.

Sie werden feststellen, dass die Tabelle zu groß ist, um vollständig auf der Zeichenfläche angezeigt zu werden. Sie können auf der Zeichenfläche einen Bildlauf nach oben/unten bzw. rechts/links durchführen, um andere Bereiche der Tabellenvorschau einzublenden, oder folgende Aktionen durchführen:

- Klicken Sie im Dialogfeld zum Konstruieren von Tabellen auf *Kompakt*, um eine Kompaktansicht anzuzeigen. Hierbei werden nur die Variablenlabels ohne Informationen zu den in der Tabelle enthaltenen Kategorien oder Auswertungsstatistiken angezeigt.
- Vergrößern Sie das Dialogfeld zum Konstruieren von Tabellen, indem Sie auf eine beliebige Seite oder Ecke des Dialogfeld zum Konstruieren von Tabellen klicken und diese ziehen.

Abbildung 4-11  
Kompaktansicht auf der Zeichenfläche



- Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 4-12  
Verschachtelte Kreuztabelle

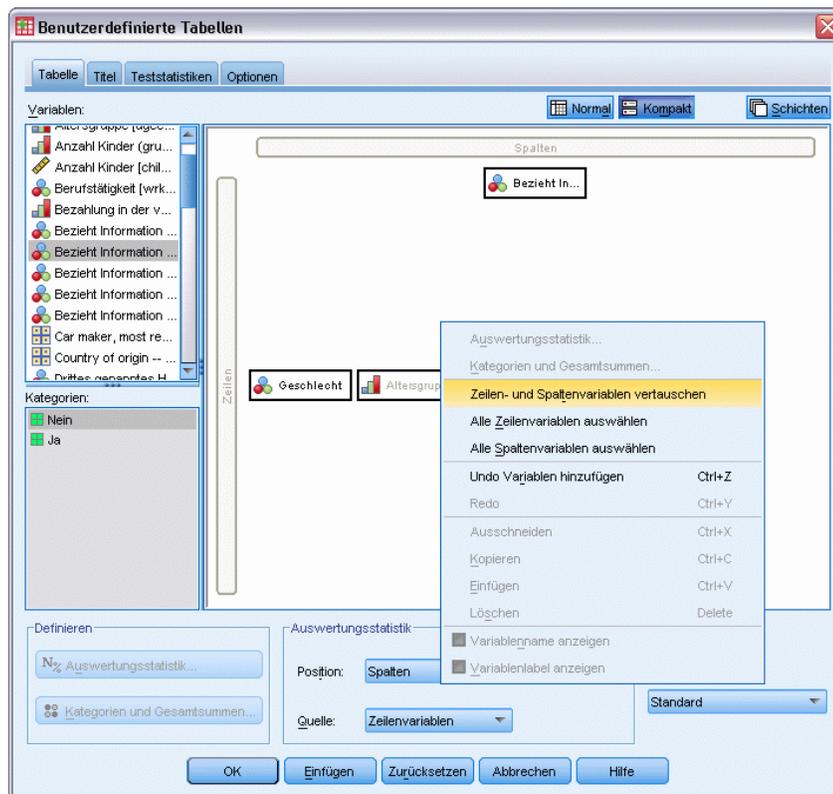
				Nachrichten aus dem Internet	
				Nein	Ja
				Anzahl	Anzahl
Geschlecht	Männlich	Alterskategorie	Weniger als 25	59	49
			25 bis 34	159	117
			35 bis 44	217	92
	Weiblich	Alterskategorie	45 bis 54	169	52
			55 bis 64	112	24
			65 oder älter	155	23
Weiblich	Alterskategorie	Weniger als 25	87	47	
		25 bis 34	209	142	
		35 bis 44	218	152	
		45 bis 54	177	83	
		55 bis 64	140	44	
		65 oder älter	261	40	

### Vertauschen von Zeilen und Spalten

Wie gehen Sie vor, wenn Sie viel Zeit aufgewendet haben, um genau die gewünschte Tabelle zu konstruieren, dann aber feststellen, dass die Orientierung der Tabelle geändert werden müsste, indem die Zeilenvariablen in die Spalten und die Spaltenvariablen in die Zeilen verschoben werden? Angenommen, Sie haben eine verschachtelte Kreuztabelle erstellt, bei der *Alterskategorie* und *Geschlecht* in den Zeilen verschachtelt sind. Nun möchten Sie diese beiden demografischen Variablen jedoch lieber in den Spalten verschachteln.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf eine beliebige Stelle und wählen Sie im Kontextmenü die Option Zeilen- und Spaltenvariablen vertauschen aus.

Abbildung 4-13  
Vertauschen von Zeilen- und Spaltenvariablen



Die Zeilen- und Spaltenvariablen sind nun vertauscht.

Vor dem Erstellen der Tabelle sollten noch einige Änderungen vorgenommen werden, um die Anzeige etwas übersichtlicher zu gestalten.

- ▶ Wählen Sie Verbergen aus, um die Spaltenbeschriftungen für Auswertungsstatistiken zu unterdrücken.

- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Geschlecht*, und deaktivieren Sie Variablenlabel anzeigen.
- ▶ Klicken Sie jetzt auf OK, um die Tabelle zu erstellen.

Abbildung 4-14

Kreuztabelle mit in Spalten verschachtelten demografischen Variablen

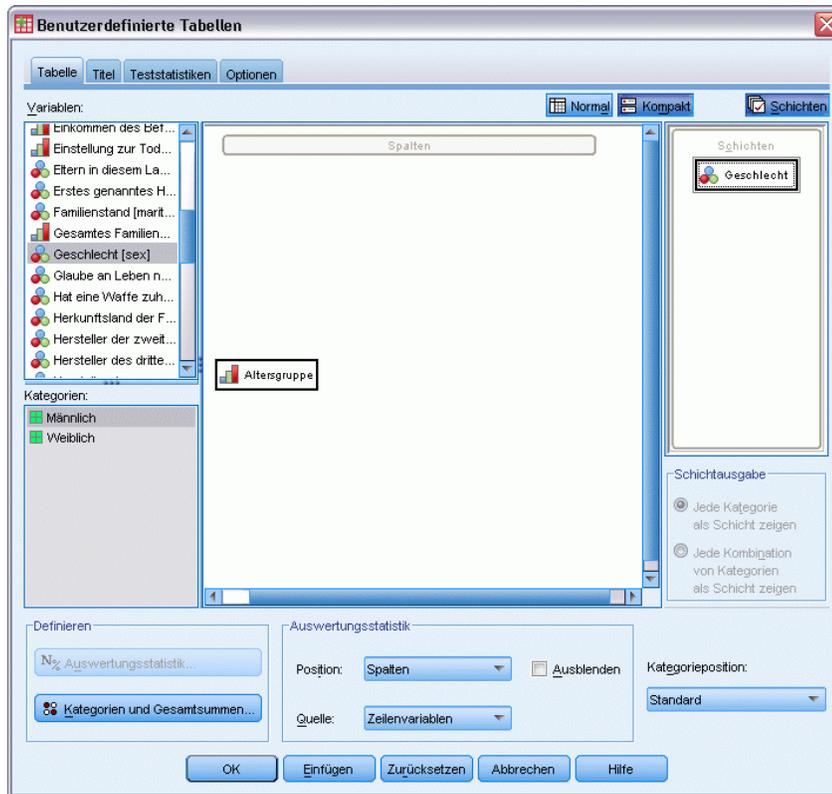
		Männlich						Weiblich					
		Weniger als 25	25 bis 34	35 bis 44	45 bis 54	55 bis 64	65 oder älter	Weniger als 25	25 bis 34	35 bis 44	45 bis 54	55 bis 64	65 oder älter
		Anzahl	Anzahl	Anzahl	Anzahl	Anzahl	Anzahl	Anzahl	Anzahl	Anzahl	Anzahl	Anzahl	Anzahl
Nachrichten aus dem Internet	Nein	59	159	217	169	112	155	87	209	218	177	140	261
	Ja	49	117	92	52	24	23	47	142	152	83	44	40

## Schichten

Mit Schichten können Tabellen um eine Tiefendimension erweitert werden, um dreidimensionale “Würfel” zu erstellen. Schichten ähneln stark der Verschachtelung oder Stapelung. Der wichtigste Unterschied besteht darin, dass stets nur eine Schichtenkategorie sichtbar ist. Wenn Sie beispielsweise *Alterskategorie* als Zeilenvariable und *Geschlecht* als Schichtvariable verwenden, erhalten Sie eine Tabelle, in der die Informationen für Männer und Frauen in unterschiedlichen Schichten dargestellt werden.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie auf Zurücksetzen, um frühere Einstellungen in der Tabellenkonstruktion zu löschen.
- ▶ Ziehen Sie die Variable *Alterskategorie* aus der Variablenliste der Tabellenkonstruktion in den Bereich “Zeilen” auf der Zeichenfläche.
- ▶ Klicken Sie im oberen Bereich des Dialogfelds zum Konstruieren von Tabellen auf Schichten, um die Liste “Schichten” anzuzeigen.
- ▶ Ziehen Sie die Variable *Geschlecht* aus der Variablenliste, und legen Sie diese in der Liste “Schichten” ab.

Abbildung 4-15  
Alterskategorie in Zeilen, Geschlecht in Schichten



Sie werden an dieser Stelle feststellen, dass sich das Hinzufügen einer Schichtvariable in der Vorschau auf der Zeichenfläche optisch nicht bemerkbar macht. Schichtvariablen haben nur dann Auswirkungen auf die Vorschau auf der Zeichenfläche, wenn die Schichtvariable die Quellvariable für die Statistik ist und Sie die Auswertungsstatistik ändern.

- Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 4-16  
Einfache Tabelle mit Schichten

Geschlecht Männlich		Anzahl
Alterskategorie	Weniger als 25	108
	25 bis 34	276
	35 bis 44	309
	45 bis 54	221
	55 bis 64	136
	65 oder älter	178

Auf den ersten Blick unterscheidet sich diese Tabelle nicht von einer einfachen Tabelle mit einer einzelnen kategorialen Variablen. Der einzige Unterschied besteht im Vorhandensein der Beschriftung *Geschlecht männlich* am oberen Rand der Tabelle.

- Doppelklicken Sie im Fenster des Viewers auf die Tabelle, um diese zu aktivieren.

- ▶ Nun wird klar, dass es sich bei der Beschriftung *Geschlecht männlich* eigentlich um einen Eintrag in einer Dropdown-Liste handelt.
- ▶ Klicken Sie auf den Abwärtspfeil der Dropdown-Liste, um die vollständige Liste der Schichten anzuzeigen.

Abbildung 4-17

Liste der Schichten in der aktivierten Pivot-Tabelle

Schicht	Geschlecht Männlich	Anzahl
Alterskategorie	Geschlecht Weiblich	108
	25 bis 34	276
	35 bis 44	309
	45 bis 54	221
	55 bis 64	136
	65 oder älter	178

Bei dieser Tabelle enthält die Liste nur einen weiteren Eintrag.

- ▶ Wählen Sie in der Dropdown-Liste den Eintrag *Geschlecht weiblich* aus.

Abbildung 4-18

Einfache Tabelle mit Schichten; angezeigt wird die andere Schicht.

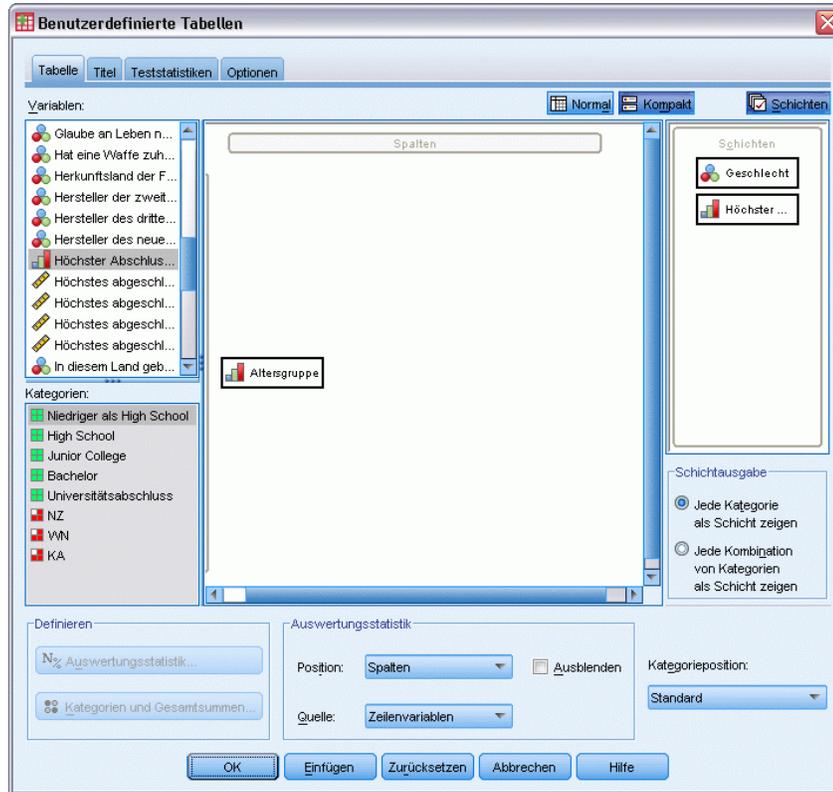
Geschlecht Weiblich		Anzahl
Alterskategorie	Weniger als 25	134
	25 bis 34	351
	35 bis 44	370
	45 bis 54	260
	55 bis 64	184
	65 oder älter	301

## Zwei gestapelte kategoriale Schichtvariablen

Wenn die Schichten mehrere kategoriale Variablen enthalten, können Sie die Schichtvariablen stapeln oder verschachteln. In der Standardeinstellung werden Schichtvariablen gestapelt. (*Hinweis:* Bei metrischen Schichtvariablen ist nur Stapeln möglich.)

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Wenn sich *Alterskategorie* nicht bereits in den Zeilen und *Geschlecht* nicht bereits in den Schichten befindet, folgen Sie den vorangegangenen Anweisungen zum Erstellen einer Tabelle mit Schichten.
- ▶ Ziehen Sie die Variable *Höchster Schulabschluss* aus der Variablenliste und legen Sie sie in der Liste “Schichten” unter *Geschlecht* ab.

Abbildung 4-19  
Gestapelte Schichtvariablen im Dialogfeld zum Konstruieren von Tabellen



Die beiden Optionsfelder im Gruppenfeld “Schichtausgabe” unter der Liste “Schichten” sind jetzt freigegeben. In der Standardeinstellung ist Jede Kategorie als Schicht zeigen aktiviert. Dies entspricht der Stapelung.

- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.
- ▶ Doppelklicken Sie im Fenster des Viewers auf die Tabelle, um diese zu aktivieren.
- ▶ Klicken Sie auf den Abwärtspfeil der Dropdown-Liste, um die vollständige Liste der Schichten anzuzeigen.

Abbildung 4-20  
Liste der gestapelten Schichten in der aktivierten Pivot-Tabelle

Schicht	Geschlecht Männlich
	Geschlecht Männlich
Alterskategorie	Geschlecht Weiblich
	Höchster Schulabschluß Hauptschule
	Höchster Schulabschluß Realschule
	Höchster Schulabschluß Gymnasium
	Höchster Schulabschluß Hochschulabschluß
	Höchster Schulabschluß Promotion

Die Tabelle enthält sieben Schichten: zwei Schichten für die beiden Kategorien von *Geschlecht* und fünf Schichten für die fünf Kategorien von *Höchster Schulabschluss*. Bei gestapelten Schichten entspricht die Gesamtzahl Schichten der Summe der Anzahl Kategorien aller Schichtvariablen (einschließlich der Kategorien für Gesamtsummen und Zwischenergebnisse, falls Sie diese für die Schichtvariablen angefordert haben).

## Zwei verschachtelte kategoriale Schichtvariablen

Beim Verschachteln von kategorialen Schichtvariablen wird für jede Kombination von Kategorien der Schichtvariablen eine separate Schicht erstellt.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Wenn Sie dies nicht bereits getan haben, befolgen Sie die vorangegangenen Anweisungen zum Erstellen einer Tabelle mit gestapelten Schichten.
- ▶ Wählen Sie im Gruppenfeld “Schichtausgabe” die Option Jede Kombination von Kategorien als Schicht zeigen aus. Dies entspricht der Verschachtelung.
- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.
- ▶ Doppelklicken Sie im Fenster des Viewers auf die Tabelle, um diese zu aktivieren.
- ▶ Klicken Sie auf den Abwärtspfeil der Dropdown-Liste, um die vollständige Liste der Schichten anzuzeigen.

Abbildung 4-21

Liste der verschachtelten Schichten in der aktivierten Pivot-Tabelle

Schicht	Geschlecht Männlich Höchster Schulabschluss Hauptschule	
	Geschlecht Männlich Höchster Schulabschluss Hauptschule	
Alterskategorie	Geschlecht Männlich Höchster Schulabschluss Realschule	3
	Geschlecht Männlich Höchster Schulabschluss Gymnasium	3
	Geschlecht Männlich Höchster Schulabschluss Hochschulabschluss	4
	Geschlecht Männlich Höchster Schulabschluss Promotion	5
	Geschlecht Weiblich Höchster Schulabschluss Hauptschule	4
	Geschlecht Weiblich Höchster Schulabschluss Realschule	6
	Geschlecht Weiblich Höchster Schulabschluss Gymnasium	

Die Tabelle enthält 10 Schichten (Sie müssen in der Liste blättern, um alle Schichten anzuzeigen): eine Schicht für jede Kombination von *Geschlecht* und *Höchster Schulabschluss*. Bei verschachtelten Schichten entspricht die Gesamtzahl der Schichten dem *Produkt* aus den Anzahlen der Kategorien jeder Schichtvariablen (im vorliegenden Beispiel ist dies  $5 \times 2 = 10$ ).

## Drucken von Tabellen mit Schichten

In der Standardeinstellung wird nur die aktuell sichtbare Schicht gedruckt. So drucken Sie alle Schichten in einer Tabelle:

- ▶ Doppelklicken Sie im Fenster des Viewers auf die Tabelle, um diese zu aktivieren.

- ▶ Wählen Sie die folgenden Befehle aus den Menüs des Viewer-Fensters aus:  
Format > Tabelleneigenschaften...
- ▶ Klicken Sie auf die Registerkarte Drucken.
- ▶ Wählen Sie Alle Schichten drucken aus.

Sie können diese Einstellung in einer Tabellenvorlage (auch der Standard-Tabellenvorlage) speichern.

# ***Gesamtsummen und Zwischenergebnisse für kategoriale Variablen***

Benutzerdefinierte Tabellen können sowohl Gesamtsummen als auch Zwischenergebnisse enthalten. Gesamtsummen und Zwischenergebnisse können kategorialen Variablen auf jeder Verschachtelungsebene und in jeder Dimension (Zeile, Spalte oder Schicht) zugewiesen werden.

## ***Beispiel-Datendatei***

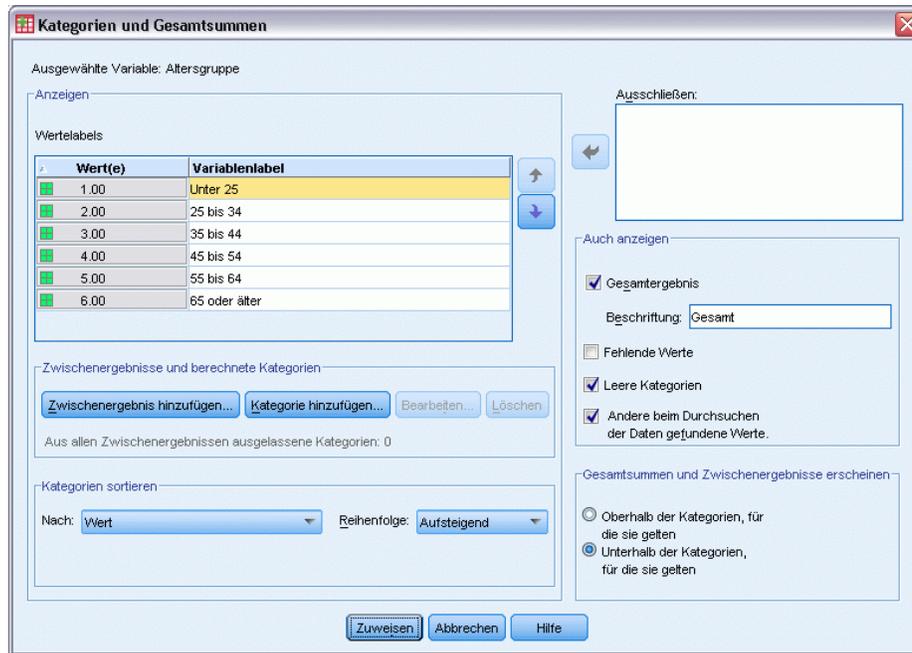
Für die Beispiele in diesem Kapitel wird die Datendatei *survey\_sample.sav* verwendet. [Für weitere Informationen siehe Thema Beispieldateien in Anhang A auf S. 201.](#)

In allen angegebenen Beispielen werden die Variablenlabels in Dialogfeldern in alphabetischer Reihenfolge angezeigt. Die Anzeigeeigenschaften für Variablenlisten werden im Dialogfeld “Optionen” auf der Registerkarte “Allgemein” festgelegt (Menü “Bearbeiten”, Befehl “Optionen”).

## ***Einfache Gesamtsumme für eine Variable***

- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
Analysieren > Tabellen > Benutzerdefinierte Tabellen...
- ▶ Ziehen Sie die Variable *Alterskategorie* aus der Variablenliste der Tabellenkonstruktion in den Bereich “Zeilen” auf der Zeichenfläche.
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option Auswertungsstatistik aus.
- ▶ Wählen Sie im Dialogfeld “Auswertungsstatistik” in der Liste “Statistik” die Option Anzahl als Spalten% aus, und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” zu übertragen.
- ▶ Löschen Sie in der Liste “Anzeigen” in der Zelle mit der Beschriftung die Standardbeschriftung und geben Sie Prozent ein.
- ▶ Klicken Sie auf Der Auswahl zuweisen.
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.
- ▶ Wählen Sie im Dialogfeld “Kategorien und Gesamtsummen” die Option Gesamt aus, indem Sie darauf klicken.

Abbildung 5-1  
Dialogfeld "Kategorien und Gesamtsummen"



- Klicken Sie auf Zuweisen und anschließend in der Tabellenkonstruktion auf OK, um die Tabelle zu erstellen.

Abbildung 5-2  
Einfache Gesamtsumme für eine kategoriale Variable

		Anzahl	Prozent
Alterskategorie	Weniger als 25	242	8,6%
	25 bis 34	627	22,2%
	35 bis 44	679	24,0%
	45 bis 54	481	17,0%
	55 bis 64	320	11,3%
	65 oder älter	479	16,9%
	Gesamt	2828	100,0%

## **Die Gesamtsumme entspricht der Summe aller angezeigten Werte**

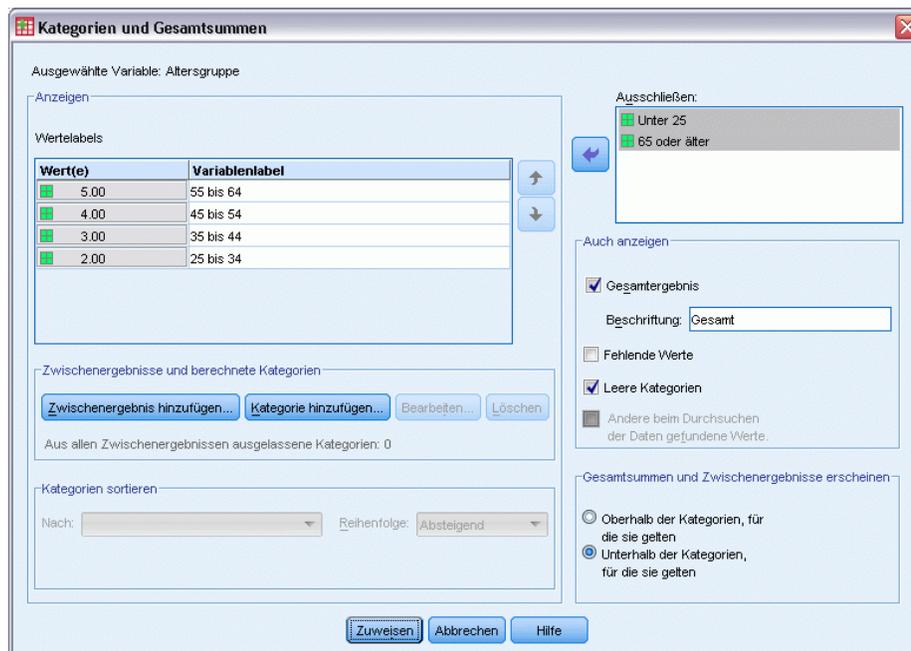
Die Gesamtsumme wird auf der Grundlage der in der Tabelle angezeigten Kategorien berechnet. Wenn Sie einige Kategorien aus der Tabelle ausschließen, werden die Fälle dieser Kategorien in der Berechnung der Gesamtsumme nicht berücksichtigt.

- Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü "Analysieren", Untermenü "Tabellen", Befehl "Benutzerdefinierte Tabellen").
- Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.
- Klicken Sie in der Liste "Label" auf die Kategorie mit der Beschriftung *Unter 25*.

- ▶ Klicken Sie auf die Pfeilschaltfläche links neben der Liste “Ausschließen”.
- ▶ Klicken Sie in der Liste “Label” auf die Kategorie mit der Beschriftung *65 oder älter*.
- ▶ Klicken Sie erneut auf die Pfeilschaltfläche links neben der Liste “Ausschließen”.

Daraufhin werden die beiden Kategorien aus der Liste “Anzeigen” in die Liste “Ausschließen” verschoben.

Abbildung 5-3  
Manuell ausgeschlossene Kategorien



- ▶ Klicken Sie auf Zuweisen und anschließend in der Tabellenkonstruktion auf OK, um die Tabelle zu erstellen.

Abbildung 5-4  
Gesamtsumme in einer Tabelle mit ausgeschlossenen Kategorien

Alterskategorie	Anzahl	Prozent
25 bis 34	627	29,8%
35 bis 44	679	32,2%
45 bis 54	481	22,8%
55 bis 64	320	15,2%
Gesamt	2107	100,0%

Die Gesamtanzahl in dieser Tabelle beträgt nur 2.107, im Vergleich zu 2.828 bei Berücksichtigung aller Kategorien. Nur die in der Tabelle verwendeten Kategorien werden bei der Berechnung der Gesamtsumme berücksichtigt. (Die Gesamtsumme für die Prozentwerte beträgt weiterhin 100 %, da die Prozentwerte anhand der Gesamtanzahl der in der Tabelle verwendeten Fälle berechnet werden, nicht anhand der Gesamtanzahl der Fälle in der Datendatei.)

## Anzeigeposition der Gesamtsummen

In der Standardeinstellung werden Gesamtsummen unter den Kategorien angezeigt, auf die sie sich beziehen. Die Gesamtsummen können jedoch auch über den jeweiligen Kategorien angezeigt werden.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.
- ▶ Wählen Sie im Gruppenfeld “Gesamtsummen und Zwischenergebnisse erscheinen” die Option Oberhalb der Kategorien, für die sie gelten aus.
- ▶ Klicken Sie auf Zuweisen und anschließend in der Tabellenkonstruktion auf OK, um die Tabelle zu erstellen.

Abbildung 5-5

Die Gesamtsumme wird über den entsprechenden Kategorien angezeigt.

		Anzahl	Prozent
Alterskategorie	Gesamt	2107	100,0%
	25 bis 34	627	29,8%
	35 bis 44	679	32,2%
	45 bis 54	481	22,8%
	55 bis 64	320	15,2%

## Gesamtsummen für verschachtelte Tabellen

Da Gesamtsummen kategorialen Variablen auf jeder Ebene der Verschachtelung zugewiesen werden können, können Sie Tabellen erstellen, die Gesamtsummen für Gruppen auf mehreren Verschachtelungsebenen enthalten.

### Gesamtsummen für Gruppen

Gesamtsummen für kategoriale Variablen, die in anderen kategorialen Variablen verschachtelt sind, stellen Gesamtsummen für Gruppen dar.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Ziehen Sie die Variable *Geschlecht* auf die Zeichenfläche, und legen Sie sie links neben *Alterskategorie* ab.
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.

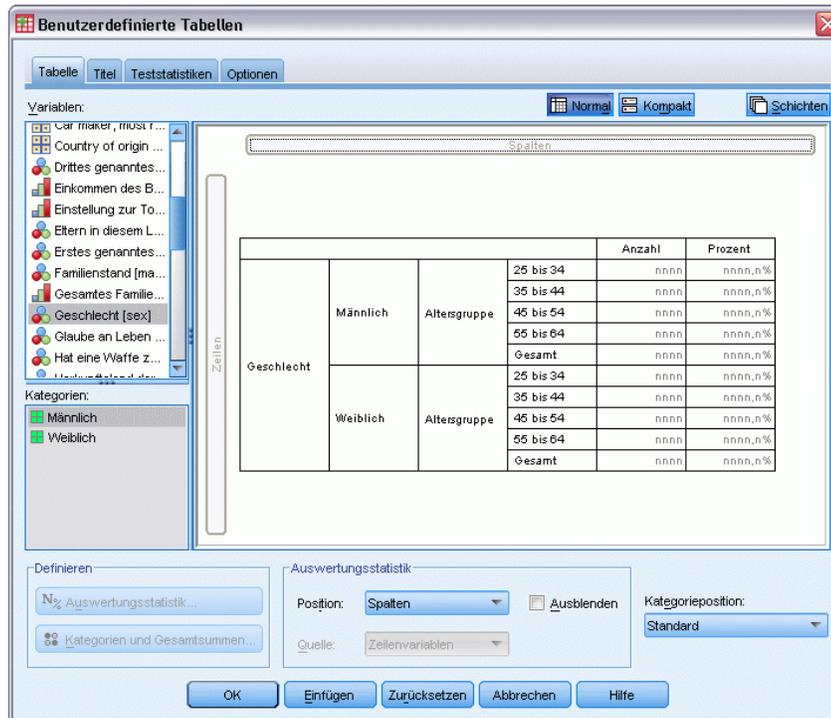
Bevor Sie die Tabelle erstellen, legen Sie erneut fest, dass die Gesamtsummen wieder unter den entsprechenden Kategorien angezeigt werden.

- ▶ Wählen Sie im Gruppenfeld “Gesamtsummen und Zwischenergebnisse erscheinen” die Option Unterhalb der Kategorien, für die sie gelten aus.

- Klicken Sie auf Zuweisen, um diese Einstellung zu speichern und zur Tabellenkonstruktion zurückzukehren.

Abbildung 5-6

Die Alterskategorie wurde in der Tabellenkonstruktion in Geschlecht verschachtelt.



- Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 5-7

Gesamtsummen von Alterskategorien innerhalb der Geschlechtskategorien

				Anzahl	Prozent
Geschlecht	Männlich	Alterskategorie	25 bis 34	276	29,3%
			35 bis 44	309	32,8%
			45 bis 54	221	23,5%
			55 bis 64	136	14,4%
			Gesamt	942	100,0%
Weiblich	Alterskategorie	25 bis 34	351	30,1%	
		35 bis 44	370	31,8%	
		45 bis 54	260	22,3%	
		55 bis 64	184	15,8%	
		Gesamt	1165	100,0%	

In der Tabelle werden nun zwei Gesamtsummen für Gruppen angezeigt: eine für Männer und eine für Frauen.

### **Gesamtergebnisse**

Bei Gesamtsummen von verschachtelten Variablen handelt es sich immer um Gesamtsummen für Gruppen, nicht um Gesamtergebnisse. Wenn Sie Gesamtsummen für die gesamte Tabelle erhalten möchten, können Sie festlegen, dass Gesamtsummen für die Variable auf der äußersten Verschachtelungsebene erstellt werden.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen erneut (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Geschlecht* und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.
- ▶ Wählen Sie im Dialogfeld “Kategorien und Gesamtsummen” die Option Gesamt aus, indem Sie darauf klicken.
- ▶ Klicken Sie auf Zuweisen und anschließend in der Tabellenkonstruktion auf OK, um die Tabelle zu erstellen.

Abbildung 5-8  
Gesamtergebnisse für eine verschachtelte Tabelle

				Anzahl	Prozent
Geschlecht	Männlich	Alterskategorie	25 bis 34	276	29,3%
			35 bis 44	309	32,8%
			45 bis 54	221	23,5%
			55 bis 64	136	14,4%
			Gesamt	942	100,0%
	Weiblich	Alterskategorie	25 bis 34	351	30,1%
			35 bis 44	370	31,8%
			45 bis 54	260	22,3%
			55 bis 64	184	15,8%
			Gesamt	1165	100,0%
	Gesamt	Alterskategorie	25 bis 34	627	29,8%
			35 bis 44	679	32,2%
			45 bis 54	481	22,8%
			55 bis 64	320	15,2%
			Gesamt	2107	100,0%

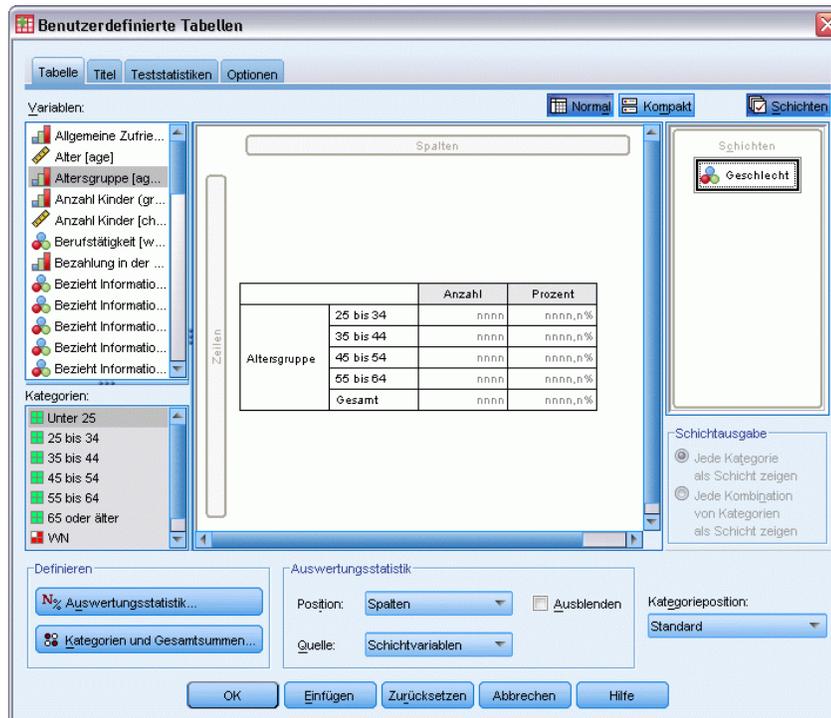
Beachten Sie, dass das Gesamtergebnis lediglich 2.107 beträgt, nicht 2.828. Zwei Alterskategorien sind noch immer aus der Tabelle ausgeschlossen, und die Fälle in diesen Kategorien werden deshalb nicht bei den Gesamtsummen berücksichtigt.

### **Gesamtsummen für Schichtvariablen**

Gesamtsummen für Schichtvariablen werden als separate Schichten in der Tabelle angezeigt.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie im Dialogfeld zum Konstruieren von Tabellen auf Schichten, um die Liste “Schichten” anzuzeigen.
- ▶ Ziehen Sie die Variable *Geschlecht* aus dem Bereich für die Zeilen auf der Zeichenfläche in die Liste “Schichten”.

Abbildung 5-9  
Schichtvariable in der Tabellenkonstruktion



*Hinweis:* Da Sie bereits festgelegt haben, dass für *Geschlecht* Gesamtsummen berechnet werden, müssen Sie dies jetzt nicht mehr tun. Das Verschieben der Variablen zwischen den Dimensionen hat keine Auswirkungen auf die Einstellungen für diese Variable.

- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.
- ▶ Doppelklicken Sie im Viewer auf die Pivot-Tabelle, um diese zu aktivieren.
- ▶ Klicken Sie in der Dropdown-Liste "Schicht" auf den Abwärts Pfeil, um eine Liste aller Schichten in der Tabelle anzuzeigen.

Die Tabelle enthält drei Schichten: *Geschlecht männlich*, *Geschlecht weiblich* und *Geschlecht gesamt*.

Abbildung 5-10  
Schicht mit Gesamtsummen in der Liste "Schichten" der aktivierten Pivot-Tabelle

Schicht	Geschlecht	Anzahl	Prozent
Alterskategorie	Geschlecht Männlich	76	29,3%
	Geschlecht Weiblich	79	32,8%
	Geschlecht Gesamt	155	32,8%
	45 bis 54	221	23,5%
	55 bis 64	136	14,4%
	Gesamt	942	100,0%

### **Anzeigeposition der Schichtsummen**

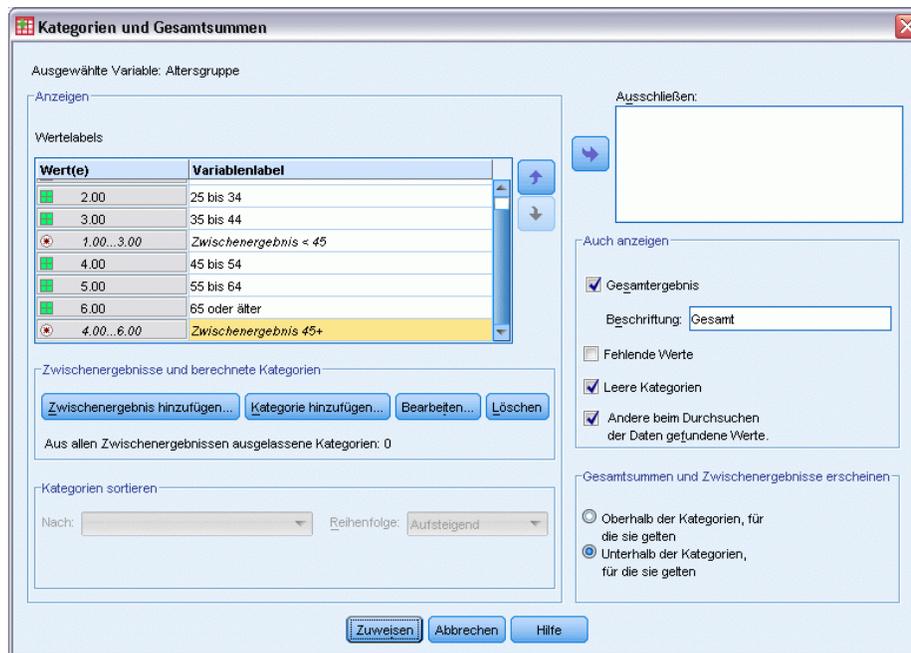
Bei Gesamtsummen für Schichtvariablen bestimmt die Anzeigeposition (über bzw. unter) der Gesamtsummen die Position der Schicht mit den Gesamtsummen. Wenn Sie beispielsweise für die Gesamtsumme einer Schichtvariablen die Option Oberhalb der Kategorien, für die sie gelten angeben, wird die Schicht mit den Gesamtsummen als erste Schicht angezeigt.

## **Zwischenergebnisse**

Sie können auch Zwischenergebnisse für Untergruppen von Kategorien einer Variablen aufnehmen. Beispielsweise können Sie Zwischenergebnisse für Alterskategorien berechnen lassen, die alle Befragten der Beispielumfrage unter und über 45 Jahren darstellt.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie auf Zurücksetzen, um frühere Einstellungen in der Tabellenkonstruktion zu löschen.
- ▶ Ziehen Sie die Variable *Alterskategorie* aus der Variablenliste der Tabellenkonstruktion in den Bereich “Zeilen” auf der Zeichenfläche.
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.
- ▶ Wählen Sie in der Liste “Wert(e)” den Eintrag 3,00 aus.
- ▶ Klicken Sie auf Zwischenergebnis hinzufügen, um das Dialogfeld “Zwischenergebnis definieren” anzuzeigen.
- ▶ Geben Sie im Textfeld “Label” Zwischenergebnis < 45 ein.
- ▶ Klicken Sie dann auf Weiter.  
Hiermit fügen Sie eine Zeile mit dem Zwischenergebnis für die ersten drei Alterskategorien ein.
- ▶ Wählen Sie in der Liste “Wert(e)” den Eintrag 6,00 aus.
- ▶ Klicken Sie auf Zwischenergebnis hinzufügen, um das Dialogfeld “Zwischenergebnis definieren” anzuzeigen.
- ▶ Geben Sie im Textfeld “Label” Zwischenergebnis 45+ ein.
- ▶ Klicken Sie dann auf Weiter.

Abbildung 5-11  
Definieren von Zwischenergebnissen



*Wichtiger Hinweis:* Bevor Sie Zwischenergebnisse definieren, sollten Sie die Anzeigeposition für die Gesamtsummen und Zwischenergebnisse (Oberhalb der Kategorien, für die sie gelten oder Unterhalb der Kategorien, für die sie gelten) festlegen. Wenn Sie die Anzeigeposition ändern, hat dies Auswirkungen auf alle Zwischenergebnisse (nicht nur auf das gegenwärtig ausgewählte). Außerdem werden die in den Zwischenergebnissen enthaltenen Kategorien geändert.

- Klicken Sie auf Zuweisen und anschließend in der Tabellenkonstruktion auf OK, um die Tabelle zu erstellen.

Abbildung 5-12  
Zwischenergebnisse für die Alterskategorie

Alterskategorie	Anzahl
Weniger als 25	242
25 bis 34	627
35 bis 44	679
Zwischenergebnis <45	1548
45 bis 54	481
55 bis 64	320
65 oder älter	479
Zwischenergebnis 45+	1280

### **Das Zwischenergebnis entspricht der Summe der angezeigten Werte**

Genau wie Gesamtsummen beruhen Zwischenergebnisse auf den in der Tabelle enthaltenen Kategorien.

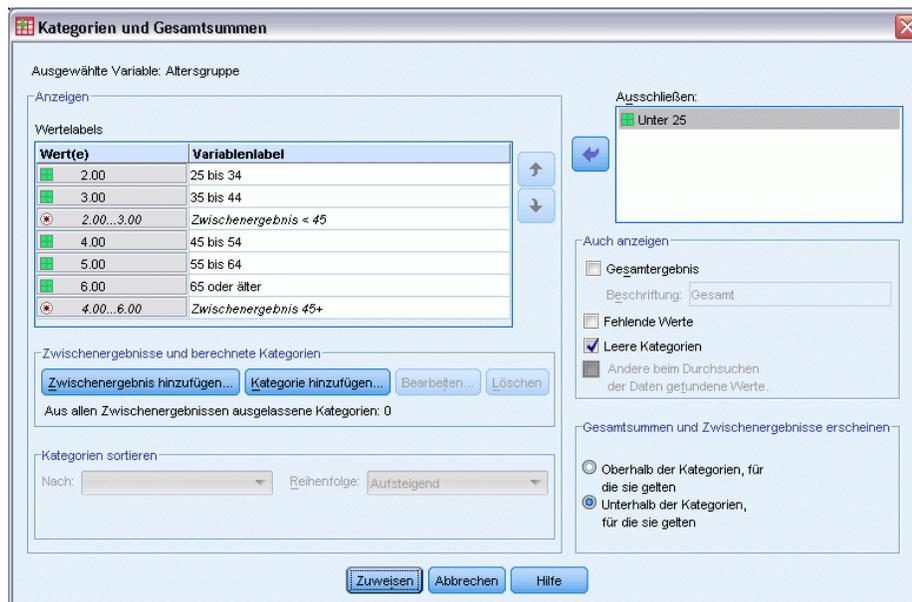
- Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü "Analysieren", Untermenü "Tabellen", Befehl "Benutzerdefinierte Tabellen").

- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option *Kategorien und Gesamtsummen* aus.

*Hinweis:* Der für das erste Zwischenergebnis angezeigte Wert (nicht das Label) beträgt 1,00 ... 3,00. Dies bedeutet, dass das Zwischenergebnis alle Werte der Liste zwischen 1 und 3 einschließt.

- ▶ Wählen Sie in der Liste “Wert(e)” den Eintrag 1,00 aus (oder klicken Sie auf das Label *Unter 25*).
- ▶ Klicken Sie auf die Pfeilschaltfläche links neben der Liste “Ausschließen”.

Abbildung 5-13  
*Zwischenergebnisse bei ausgeschlossenen Kategorien*



Die erste Alterskategorie ist nun ausgeschlossen, und der für das erste Zwischenergebnis angezeigte Wert ändert sich in 2,00 ... 3,00. Dies bedeutet, dass die ausgeschlossene Kategorie nicht im Zwischenergebnis enthalten sein wird, da Zwischenergebnisse auf den in der Tabelle enthaltenen Kategorien beruhen. Durch das Ausschließen einer Kategorie wird diese automatisch aus den Zwischenergebnissen ausgeschlossen, sodass Sie beispielsweise nicht die Zwischenergebnisse ohne die dazugehörigen Kategorien anzeigen lassen können.

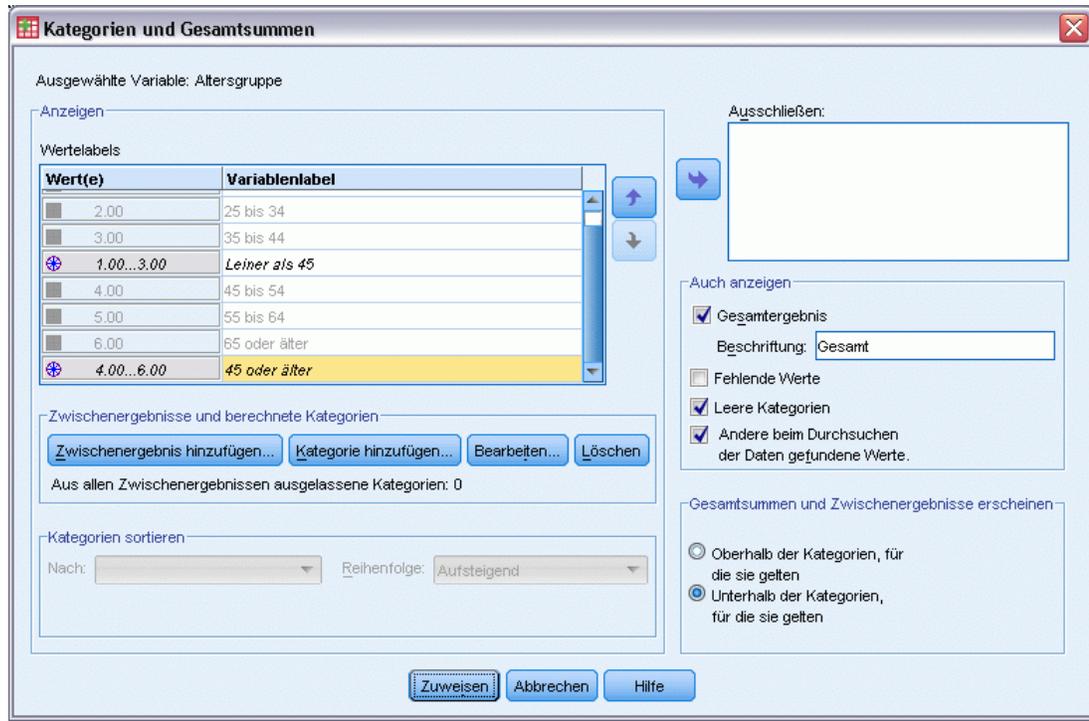
### ***Ausblenden der Kategorien für das Zwischenergebnis***

Sie haben die Möglichkeit, die Anzeige der Kategorien, die zur Definition eines Zwischenergebnisses dienen, zu unterdrücken und nur das Zwischenergebnis selbst anzuzeigen, also die Anzeige von Kategorien ohne Auswirkungen auf die zugrunde liegenden Daten zu reduzieren.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie auf Zurücksetzen, um frühere Einstellungen in der Tabellenkonstruktion zu löschen.

- ▶ Ziehen Sie die Variable *Alterskategorie* aus der Variablenliste der Tabellenkonstruktion in den Bereich “Zeilen” auf der Zeichenfläche.
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.
- ▶ Wählen Sie in der Liste “Wert(e)” den Eintrag 3,00 aus.
- ▶ Klicken Sie auf Zwischenergebnis hinzufügen, um das Dialogfeld “Zwischenergebnis definieren” anzuzeigen.
- ▶ Geben Sie im Textfeld “Label” Weniger als 45 ein.
- ▶ Aktivieren Sie Zwischenergebniskategorien aus Tabelle ausblenden.
- ▶ Klicken Sie dann auf Weiter.  
Hiermit fügen Sie eine Zeile mit dem Zwischenergebnis für die ersten drei Alterskategorien ein.
- ▶ Wählen Sie in der Liste “Wert(e)” den Eintrag 6,00 aus.
- ▶ Klicken Sie auf Zwischenergebnis hinzufügen, um das Dialogfeld “Zwischenergebnis definieren” anzuzeigen.
- ▶ Geben Sie im Textfeld “Label” 45 oder älter ein.
- ▶ Aktivieren Sie Kategorien für Zwischenergebnis ausblenden.
- ▶ Klicken Sie dann auf Weiter.
- ▶ Um eine Gesamtsumme mit den Zwischenergebnissen in die Tabelle aufzunehmen, aktivieren Sie Gesamt im Gruppenfeld “Anzeigen”.

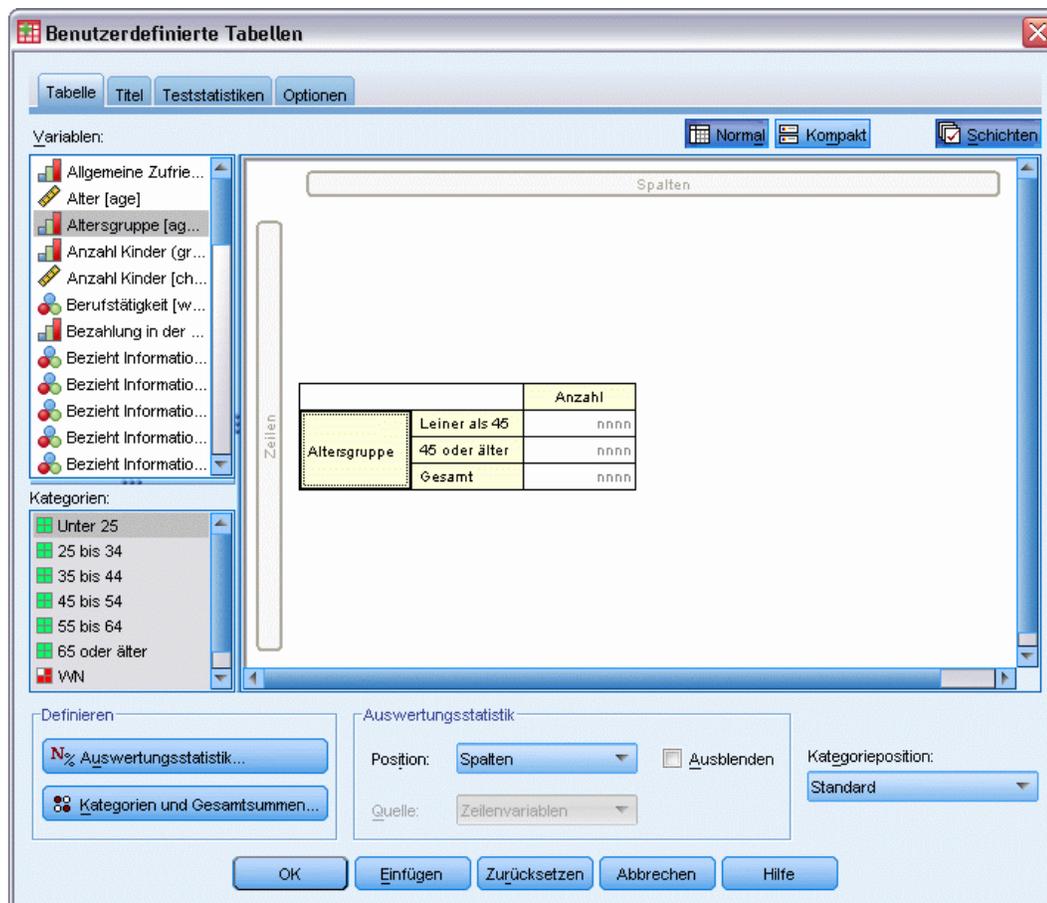
Abbildung 5-14  
Ausblenden der Kategorien für das Zwischenergebnis



- Klicken Sie auf Zuweisen.

Auf der Zeichenfläche wird die Tatsache berücksichtigt, dass die Zwischenergebnisse angezeigt, die Kategorien, die zur Definition der Zwischenergebnisse dienen, jedoch ausgeschlossen werden.

Abbildung 5-15  
Zeichenfläche mit Zwischenergebnissen ohne Kategorien für Zwischenergebnisse



- Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 5-16  
Tabelle nur mit Zwischenergebnissen und Gesamtsummen

		Anzahl
Alterskategorie	Weniger als 45	1548
	45 oder älter	1280
	Gesamt	2828

## Zwischenergebnisse für Schichtvariablen

Genau wie Gesamtsummen werden Zwischenergebnisse für Schichtvariablen als separate Schichten in der Tabelle angezeigt. Im wesentlichen werden die Zwischenergebnisse wie Kategorien behandelt. Jede Kategorie stellt eine eigene Schicht in der Tabelle dar, und die Anzeigereihenfolge der Kategorien in den Schichten wird durch die im Dialogfeld "Kategorien und Gesamtsummen" festgelegte Reihenfolge für die Kategorien bestimmt. Hierbei werden die Anzeigepositionen der Zwischenergebniskategorien berücksichtigt.

# Berechnete Kategorien für kategoriale Variablen

Sie können berechnete Kategorien in benutzerdefinierte Tabellen aufnehmen. Hierbei handelt es sich um neue Kategorien, die aus Kategorien derselben Variable in einer beliebigen Verschachtelungsebene in einer beliebigen Dimension (Zeile, Spalte oder Schicht) berechnet werden. Sie könnten zum Beispiel eine berechnete Kategorie aufnehmen, die die Differenz zwischen zwei Kategorien anzeigt.

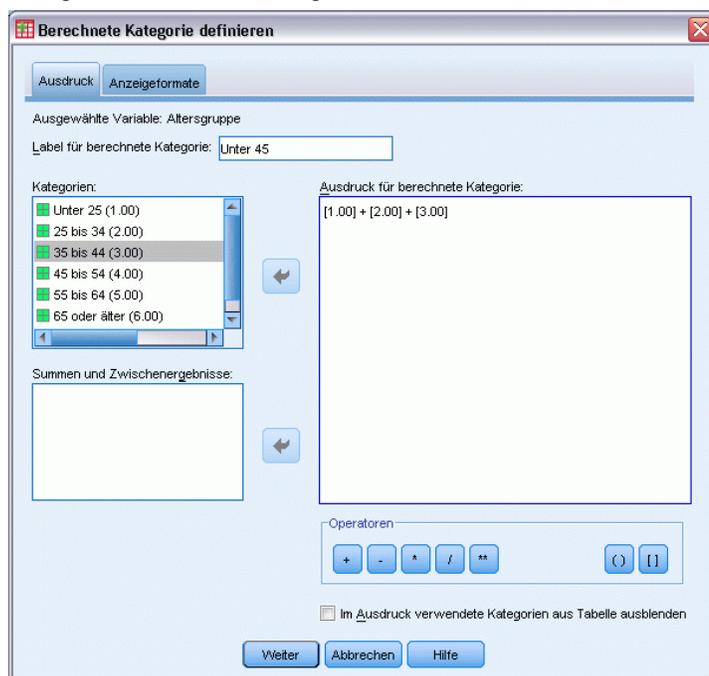
## Beispiel-Datendatei

Für die Beispiele in diesem Kapitel wird die Datendatei *survey\_sample.sav* verwendet. [Für weitere Informationen siehe Thema Beispieldateien in Anhang A auf S. 201.](#)

## Einfache berechnete Kategorie

- Wählen Sie die folgenden Befehle aus den Menüs aus:  
Analysieren > Tabellen > Benutzerdefinierte Tabellen...

Abbildung 6-1  
Dialogfeld "Berechnete Kategorie definieren"

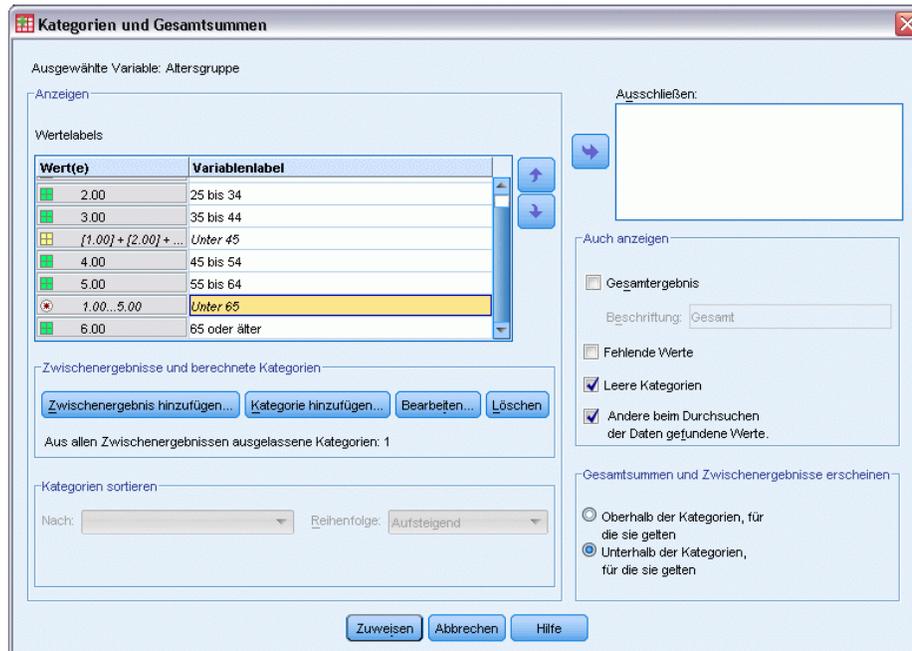


- ▶ Ziehen Sie die Variable *Alterskategorie* aus der Variablenliste der Tabellenkonstruktion in den Bereich “Zeilen” auf der Zeichenfläche.
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.
- ▶ Wählen Sie in der Liste “Wert(e)” den Eintrag 3,00 aus.
- ▶ Klicken Sie auf Kategorie hinzufügen, um das Dialogfeld “Berechnete Kategorie definieren” anzuzeigen.
- ▶ Geben Sie im Textfeld “Label für berechnete Kategorie” Weniger als 45 ein.
- ▶ Wählen Sie in der Liste “Kategorien” Weniger als 25 (1,00) aus und klicken Sie auf die Pfeilschaltfläche, um sie in das Textfeld “Ausdruck für berechnete Kategorie” zu kopieren. [1] wird im Ausdruck angezeigt.
- ▶ Klicken Sie auf die Plus (+)-Operatorschaltfläche im Dialogfeld (oder drücken Sie die +-Taste auf der Tastatur).
- ▶ Wählen Sie in der Liste “Kategorien” 25 bis 34 (2,00) aus und klicken Sie auf die Pfeilschaltfläche, um sie in das Textfeld “Ausdruck für berechnete Kategorie” zu kopieren.
- ▶ Klicken Sie auf die Plus (+)-Operatorschaltfläche im Dialogfeld (oder drücken Sie die +-Taste auf der Tastatur).
- ▶ Wählen Sie in der Liste “Kategorien” 35 bis 44 (3,00) aus und klicken Sie auf die Pfeilschaltfläche, um sie in das Textfeld “Ausdruck für berechnete Kategorie” zu kopieren.
- ▶ Klicken Sie dann auf Weiter.  
Hiermit fügen Sie eine Zeile mit dem Zwischenergebnis für die ersten drei Alterskategorien ein.
- ▶ Wählen Sie in der Liste “Wert(e)” den Eintrag 5,00 aus.
- ▶ Klicken Sie auf Zwischenergebnis hinzufügen, um das Dialogfeld “Zwischenergebnis definieren” anzuzeigen.
- ▶ Geben Sie im Textfeld “Label” Weniger als 65 ein.

- Klicken Sie dann auf Weiter.

Hiermit fügen Sie eine Zeile mit dem Zwischenergebnis für die ersten fünf Kategorien ein.

Abbildung 6-2  
Dialogfeld "Kategorien und Gesamtsummen"



- Klicken Sie auf Zuweisen und anschließend in der Tabellenkonstruktion auf OK, um die Tabelle zu erstellen.

Abbildung 6-3  
Berechnete Kategorie mit Zwischenergebnis

	Anzahl
Altersgruppe	
Unter 25	242
25 bis 34	627
35 bis 44	679
Unter 45	1548
45 bis 54	481
55 bis 64	320
Unter 65	2349
65 oder älter	479

Die Tabelle enthält eine berechnete Kategorie (*Weniger als*) und ein Zwischenergebnis (*Weniger als 65*). Das Zwischenergebnis enthält Kategorien, die auch in der berechneten Kategorie enthalten sind. Sie konnten dieselbe Tabelle nur mit den Zwischenergebnissen nicht erstellen, da Zwischenergebnisse nicht über dieselben Kategorien verfügen.

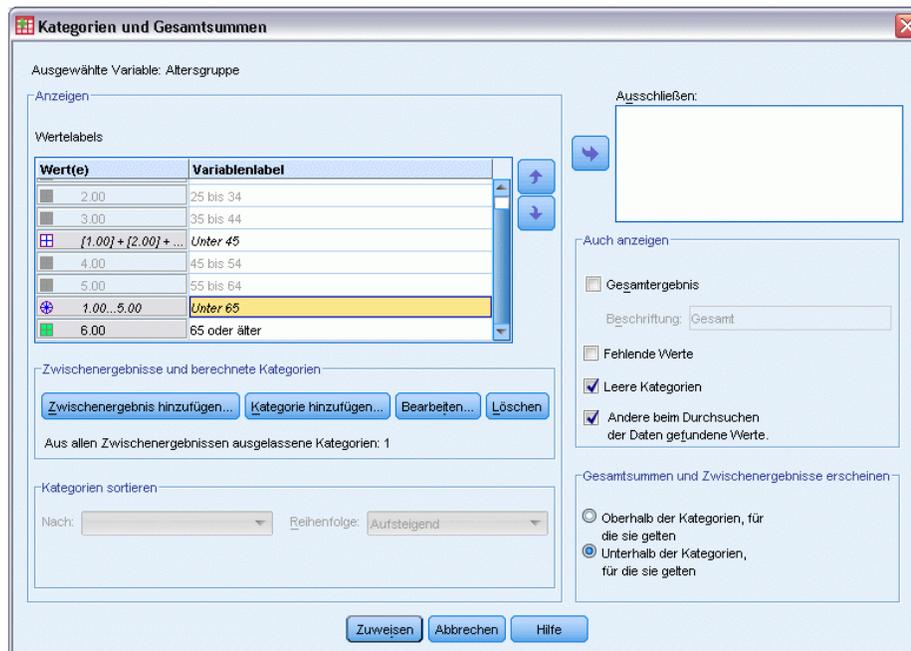
## Ausblenden von Kategorien in einer berechneten Kategorie

Wie bei den Zwischenergebnissen können Sie auch die Anzeige der Kategorien ausblenden, die im Ausdruck einer berechneten Kategorie verwendet werden, und nur die berechnete Kategorie selbst anzeigen. Das folgende Beispiel baut auf dem vorherigen auf.

- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
Analysieren > Tabellen > Benutzerdefinierte Tabellen...
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.
- ▶ Wählen Sie die berechnete Kategorie *Weniger als 45* in der Liste "Wert(e)" aus.
- ▶ Klicken Sie auf Bearbeiten, um das Dialogfeld "Berechnete Kategorie definieren" anzuzeigen.
- ▶ Wählen Sie Im Ausdruck verwendete Kategorien aus Tabelle ausblenden aus.
- ▶ Klicken Sie dann auf Weiter.
- ▶ Wählen Sie das Zwischenergebnis *Weniger als 65* in der Liste "Wert(e)" aus.
- ▶ Klicken Sie auf Bearbeiten, um das Dialogfeld "Zwischenergebnis definieren" anzuzeigen.
- ▶ Wählen Sie Zwischenergebniskategorien aus Tabelle ausblenden aus.
- ▶ Klicken Sie dann auf Weiter.

Abbildung 6-4

Anzeige ausgeblendeter Kategorien im Dialogfeld "Kategorien und Zwischenergebnisse"



- ▶ Klicken Sie auf Zuweisen und anschließend in der Tabellenkonstruktion auf OK, um die Tabelle zu erstellen.

**Abbildung 6-5**  
*Berechnete Kategorie mit Zwischenergebnis und ausgeblendeten Kategorien*

		Anzahl
Altersgruppe	Unter 45	1548
	Unter 65	2349
	65 oder älter	479

Wie im vorherigen Beispiel enthält die Tabelle eine berechnete Kategorie und ein Zwischenergebnis. Doch in diesem Fall sind die Kategorien ausgeblendet und es werden nur die Zwischenergebnisse angezeigt.

## **Verweis auf Zwischenergebnisse in einer berechneten Kategorie**

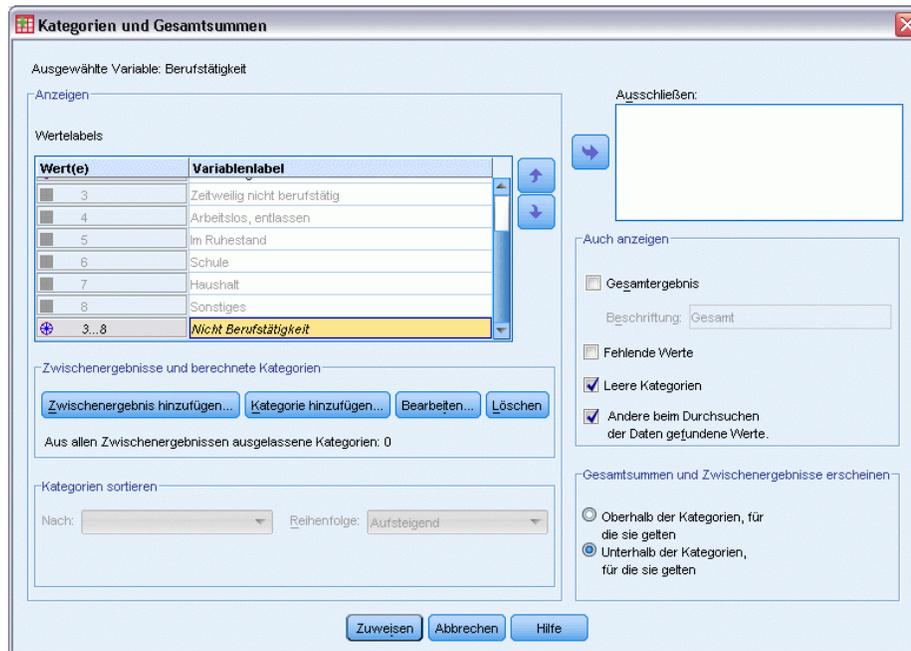
Sie können Zwischenergebnisse in den Ausdruck einer berechneten Kategorie aufnehmen.

- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
 Analysieren > Tabellen > Benutzerdefinierte Tabellen...
  - ▶ Klicken Sie auf Zurücksetzen, um frühere Einstellungen in der Tabellenkonstruktion zu löschen.
  - ▶ Ziehen Sie die Variable *Beschäftigungsstatus* aus der Variablenliste der Tabellenkonstruktion in den Bereich für die Zeilen auf der Zeichenfläche.
  - ▶ Ziehen Sie die Variable *Familienstand* aus der Variablenliste, und legen Sie sie im Spaltenbereich ab.
  - ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Beschäftigungsstatus* und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.
  - ▶ Wählen Sie in der Liste “Wert(e)” den Eintrag 2 aus.
  - ▶ Klicken Sie auf Zwischenergebnis hinzufügen, um das Dialogfeld “Zwischenergebnis definieren” anzuzeigen.
  - ▶ Geben Sie im Textfeld “Label” Beschäftigt ein.
  - ▶ Wählen Sie Zwischenergebniskategorien aus Tabelle ausblenden aus.
  - ▶ Klicken Sie dann auf Weiter.
- Hiermit fügen Sie eine Zeile mit dem Zwischenergebnis für die ersten zwei Kategorien des Beschäftigungsstatus ein.
- ▶ Wählen Sie in der Liste “Wert(e)” den Eintrag 8 aus.
  - ▶ Klicken Sie auf Zwischenergebnis hinzufügen, um das Dialogfeld “Zwischenergebnis definieren” anzuzeigen.
  - ▶ Geben Sie im Textfeld “Label” Nicht beschäftigt ein.
  - ▶ Wählen Sie Kategorien für Zwischenergebnis ausblenden aus.

- ▶ Klicken Sie dann auf Weiter.

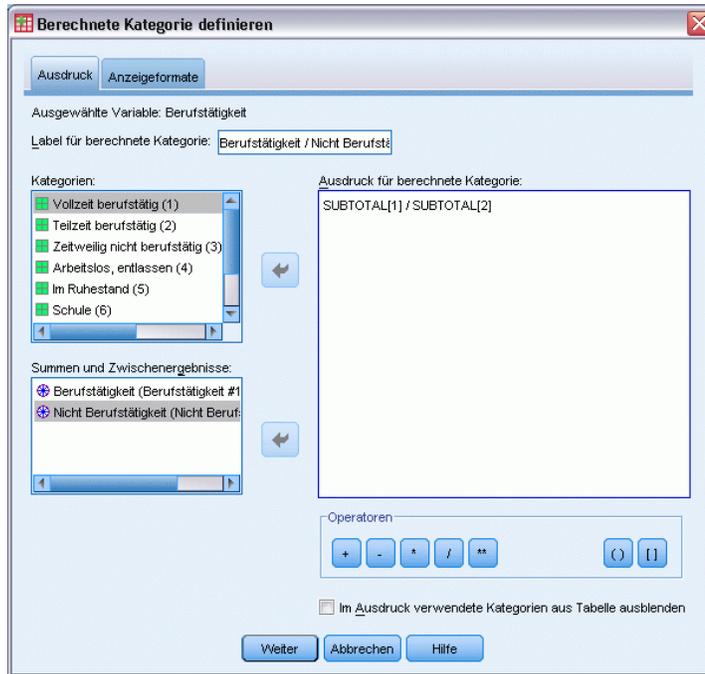
Hiermit fügen Sie eine Zeile mit dem Zwischenergebnis für die restlichen Kategorien des Beschäftigungsstatus ein.

Abbildung 6-6  
Dialogfeld "Kategorien und Gesamtsummen"



- ▶ Wählen Sie das Zwischenergebnis *Nicht beschäftigt* in der Liste "Wert(e)" aus.
- ▶ Klicken Sie auf Kategorie hinzufügen, um das Dialogfeld "Berechnete Kategorie definieren" anzuzeigen.
- ▶ Geben Sie im Textfeld "Label für berechnete Kategorie" Beschäftigt / Nicht beschäftigt ein.
- ▶ Wählen Sie in der Liste "Gesamtsummen und Zwischenergebnisse" Beschäftigt (Beschäftigt #1) aus und klicken Sie auf die Pfeilschaltfläche, um sie in das Textfeld "Ausdruck für berechnete Kategorie" zu kopieren.
- ▶ Klicken Sie auf die Divisions (/)-Operatorschaltfläche im Dialogfeld (oder drücken Sie die /-Taste auf der Tastatur).
- ▶ Wählen Sie in der Liste "Gesamtsummen und Zwischenergebnisse" Nicht beschäftigt (Nicht beschäftigt #2) aus und klicken Sie auf die Pfeilschaltfläche, um sie in das Textfeld "Ausdruck für berechnete Kategorie" zu kopieren.

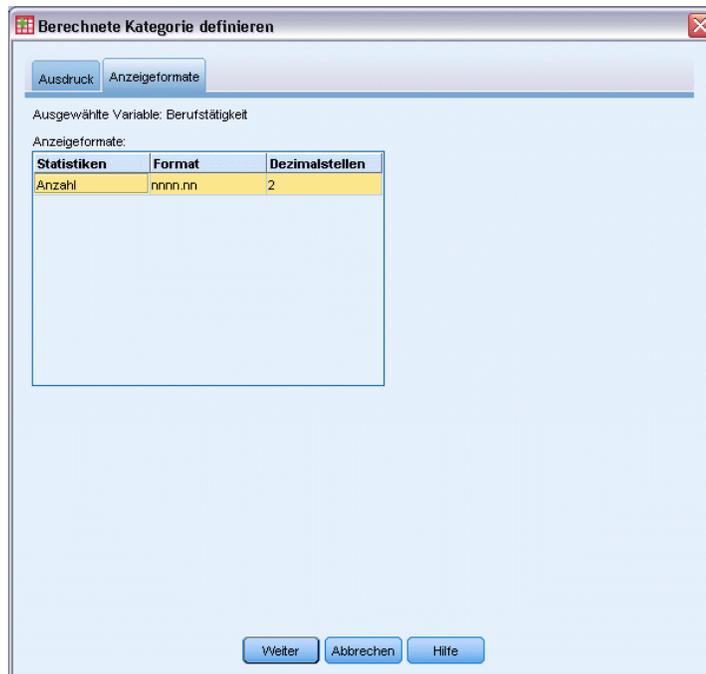
Abbildung 6-7  
Dialogfeld "Berechnete Kategorie definieren"



Die berechnete Kategorie verwendet standardmäßig das gleiche Format wie die Variablenstatistik, in diesem Fall "Anzahl". Da nach der Division im Ausdruck der berechneten Kategorie Dezimalstellen angezeigt werden sollen und das Standardformat für "Anzahl" keine Dezimalstellen enthält, muss das Format geändert werden.

- ▶ Klicken Sie auf die Registerkarte "Anzeigeformate".
- ▶ Ändern Sie die Einstellung der Dezimalstellen für "Anzahl" auf 2.

Abbildung 6-8  
Registerkarte "Anzeigeformate" im Dialogfeld "Berechnete Kategorie definieren"



- ▶ Klicken Sie dann auf Weiter.
- ▶ Klicken Sie auf Zuweisen und anschließend in der Tabellenkonstruktion auf OK, um die Tabelle zu erstellen.

Abbildung 6-9  
Anzeige des Anteils von Zwischenergebnissen in berechneten Kategorien

		Familienstand				
		Verheiratet	Verwitwet	Geschieden	Getrennt	Nie verheiratet
		Anzahl	Anzahl	Anzahl	Anzahl	Anzahl
Berufstätigkeit	Berufstätigkeit	916	64	330	67	494
	Nicht Berufstätigkeit	429	219	116	26	169
	Berufstätigkeit / Nicht Berufstätigkeit	2.14	.29	2.84	2.58	2.92

Die Tabelle enthält zwei Zwischenergebnisse und eine berechnete Kategorie. Die berechnete Kategorie zeigt den Anteil der Zwischenergebnisse an, damit Sie die durch jedes Zwischenergebnis dargestellten Gruppen leicht vergleichen können. Der Anteil von beschäftigten und nicht beschäftigten verwitweten Teilnehmern ist deutlich niedriger als bei den anderen Gruppen. Außerdem herrscht ein etwas niedrigerer Anteil an verheirateten Teilnehmern, möglicherweise aufgrund von Lebensgefährten, die das Beschäftigungsverhältnis kündigen, um sich zu Hause um das Kind zu kümmern.

## ***Einsatz von berechneten Kategorien zur Anzeige nicht abschließender Zwischenergebnisse***

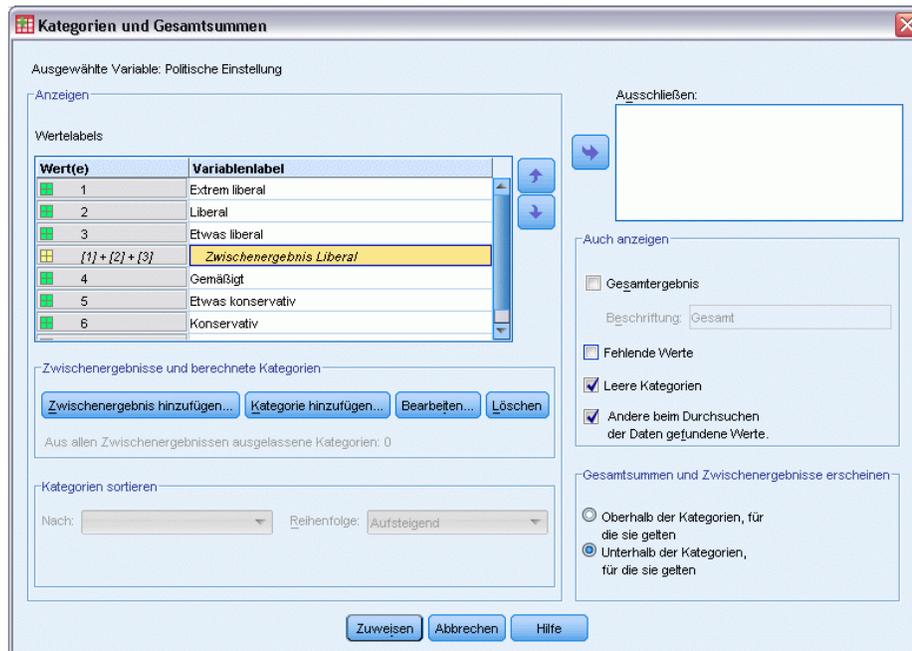
Zwischenergebnisse sind abschließend. Das heißt, alle Zwischenergebnisse in einer Tabelle enthalten alle Werte über und unter ihren Positionen in der Tabelle. Berechnete Kategorien sind dagegen nicht abschließend und ermöglichen Ihnen, eine Summe aus einer Mischung von Kategorien in der Tabelle zu erstellen.

- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
Analysieren > Tabellen > Benutzerdefinierte Tabellen...
- ▶ Klicken Sie auf Zurücksetzen, um frühere Einstellungen in der Tabellenkonstruktion zu löschen.
- ▶ Ziehen Sie die Variable *Betrachtet sich selbst als liberal oder konservativ* aus der Variablenliste der Tabellenkonstruktion in den Bereich für die Zeilen auf der Zeichenfläche.
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Betrachtet sich selbst als liberal oder konservativ* und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.
- ▶ Wählen Sie in der Liste “Wert(e)” den Eintrag 3 aus.
- ▶ Klicken Sie auf Kategorie hinzufügen, um das Dialogfeld “Berechnete Kategorie definieren” anzuzeigen.
- ▶ Geben Sie im Textfeld “Label für berechnete Kategorie” *Liberales Zwischenergebnis* ein. Beachten Sie, dass vor dem Text vier Leerzeichen stehen. Diese vier Leerzeichen werden als Einrückung in der Endtabelle verwendet.
- ▶ Wählen Sie in der Liste “Kategorien” *Extrem liberal (1)* aus und klicken Sie auf die Pfeilschaltfläche, um sie in das Textfeld “Ausdruck für berechnete Kategorie” zu kopieren.
- ▶ Klicken Sie auf die Plus (+)-Operatorschaltfläche im Dialogfeld (oder drücken Sie die +-Taste auf der Tastatur).
- ▶ Wählen Sie in der Liste “Kategorien” *Liberal (2)* aus und klicken Sie auf die Pfeilschaltfläche, um sie in das Textfeld “Ausdruck für berechnete Kategorie” zu kopieren.
- ▶ Klicken Sie auf die Plus (+)-Operatorschaltfläche im Dialogfeld (oder drücken Sie die +-Taste auf der Tastatur).
- ▶ Wählen Sie in der Liste “Kategorien” *Leicht liberal (3)* aus und klicken Sie auf die Pfeilschaltfläche, um sie in das Textfeld “Ausdruck für berechnete Kategorie” zu kopieren.

- Klicken Sie auf Weiter.

Hiermit fügen Sie eine Zeile mit dem Zwischenergebnis für die liberalen Kategorien ein.

Abbildung 6-10  
Dialogfeld "Kategorien und Gesamtsummen"

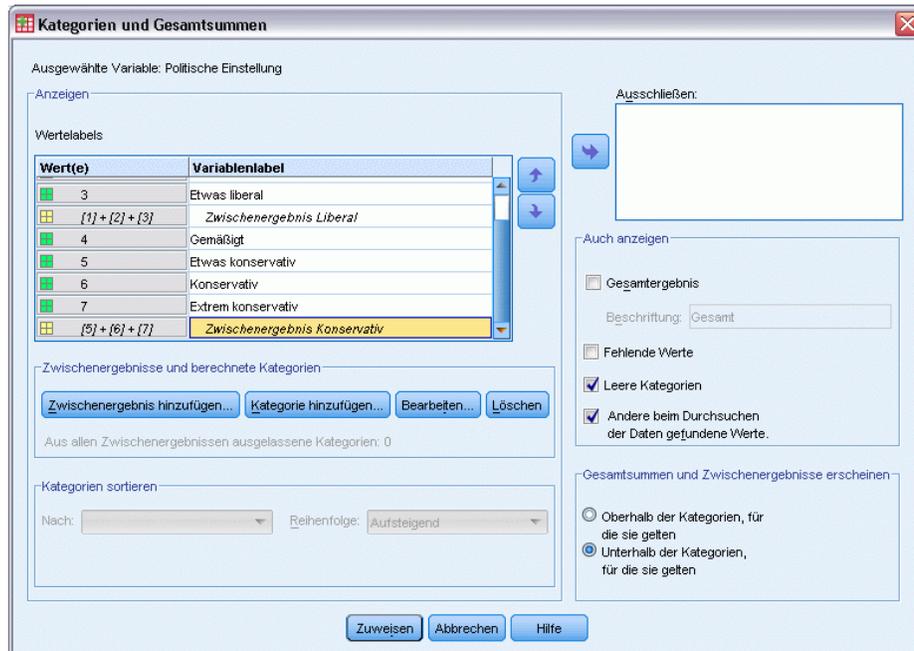


- Wählen Sie in der Liste "Wert(e)" den Eintrag 7 aus.
- Klicken Sie auf Kategorie hinzufügen, um das Dialogfeld "Berechnete Kategorie definieren" anzuzeigen.
- Geben Sie im Textfeld "Label für berechnete Kategorie" Konservatives Zwischenergebnis ein. Beachten Sie, dass vor dem Text vier Leerzeichen stehen. Diese vier Leerzeichen werden als Einrückung in der Endtabelle verwendet.
- Wählen Sie in der Liste "Kategorien" Leicht konservativ (5) aus und klicken Sie auf die Pfeilschaltfläche, um sie in das Textfeld "Ausdruck für berechnete Kategorie" zu kopieren.
- Klicken Sie auf die Plus (+)-Operatorschaltfläche im Dialogfeld (oder drücken Sie die +-Taste auf der Tastatur).
- Wählen Sie in der Liste "Kategorien" Konservativ (6) aus und klicken Sie auf die Pfeilschaltfläche, um sie in das Textfeld "Ausdruck für berechnete Kategorie" zu kopieren.
- Klicken Sie auf die Plus (+)-Operatorschaltfläche im Dialogfeld (oder drücken Sie die +-Taste auf der Tastatur).
- Wählen Sie in der Liste "Kategorien" Extrem konservativ (7) aus und klicken Sie auf die Pfeilschaltfläche, um sie in das Textfeld "Ausdruck für berechnete Kategorie" zu kopieren.

- Klicken Sie auf Weiter.

Hiermit fügen Sie eine Zeile mit dem Zwischenergebnis für die konservativen Kategorien ein.

Abbildung 6-11  
Dialogfeld "Kategorien und Gesamtsummen"



- Klicken Sie auf Zuweisen und anschließend in der Tabellenkonstruktion auf OK, um die Tabelle zu erstellen.

Abbildung 6-12  
Anzeige von nicht abschließenden Zwischenergebnissen für berechnete Kategorien

		Anzahl
Politische Einstellung	Extrem liberal	64
	Liberal	357
	Etwas liberal	351
	Zwischenergebnis Liberal	772
	Gemäßigt	986
	Etwas konservativ	432
	Konservativ	415
	Extrem konservativ	86
	Zwischenergebnis Konservativ	933

Die Tabelle enthält zwei berechnete Kategorien, die nicht alle in der Tabelle angezeigten Kategorien enthalten. Die Kategorie *Mäßig* ist in keiner der berechneten Kategorien enthalten. Sie können die gleiche Tabelle mit Zwischenergebnissen nicht erstellen, da Zwischenergebnisse abschließend sind.

# Tabellen für Variablen mit gemeinsamen Kategorien

Oft enthalten Umfragen viele Fragen mit übereinstimmenden möglichen Antworten. Die Beispielumfrage enthält beispielsweise eine Reihe von Variablen, die sich auf das Vertrauen in verschiedene öffentliche und private Institutionen und Dienstleistungen beziehen. Für diese Variablen ist jeweils dieselbe Gruppe von Antwortkategorien verfügbar: 1 = *Sehr viel*, 2 = *Mäßig* und 3 = *Fast kein*. Sie können diese zusammenhängenden Variablen stapeln, damit sie in einer Tabelle dargestellt werden, und die gemeinsamen Antwortkategorien können als Spalten der Tabelle erscheinen. Diese Funktionen sind auch verfügbar, wenn Sie berechnete Kategorien verwenden, wobei Voraussetzung ist, dass alle Beschriftungen und Ausdrücke berechneter Kategorien in sämtlichen Variablen übereinstimmen.

Abbildung 7-1  
Tabelle für Variablen mit gemeinsamen Kategorien

	Sehr viel	Mäßig	Fast kein
Vertrauen in die Medizin	844	864	167
Vertrauen in Banken & Geldinstitute	490	1068	306
Vertrauen in das Fernsehen	196	936	744
Vertrauen in die Bildung	511	1055	315
Vertrauen in die Presse	176	878	808
Vertrauen in Großkonzerne	500	1078	243

*Hinweis:* In der Vorversion der Funktion für benutzerdefinierte Tabellen wurden diese als "Häufigkeitstabellen" bezeichnet.

## Beispiel-Datendatei

Für die Beispiele in diesem Kapitel wird die Datendatei *survey\_sample.sav* verwendet. [Für weitere Informationen siehe Thema Beispieldateien in Anhang A auf S. 201.](#)

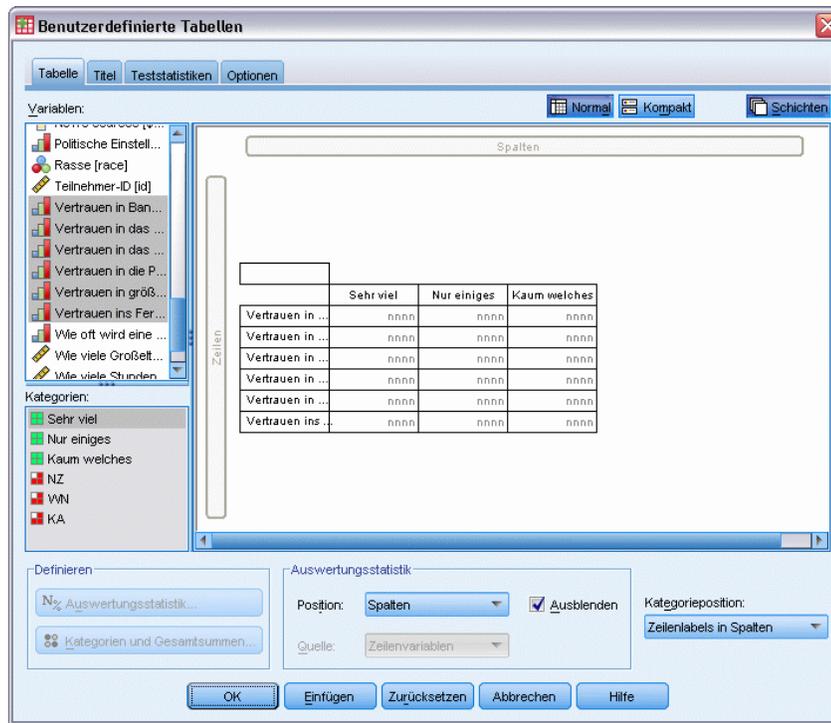
In allen angegebenen Beispielen werden die Variablenlabels in Dialogfeldern in alphabetischer Reihenfolge angezeigt. Die Anzeigeeigenschaften für Variablenlisten werden im Dialogfeld "Optionen" auf der Registerkarte "Allgemein" festgelegt (Menü "Bearbeiten", Befehl "Optionen").

## Häufigkeitstabelle

- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
Analysieren > Tabellen > Benutzerdefinierte Tabellen...
- ▶ Klicken Sie in der Variablenliste des Dialogfelds zum Konstruieren von Tabellen auf *Vertrauen in Banken...* und klicken Sie anschließend mit gedrückt gehaltener Umschalttaste auf *Vertrauen in das Fernsehen*, um alle "Vertrauens"-Variablen auszuwählen. (*Hinweis:* Hierbei wird



Abbildung 7-3  
Kategorienbeschriftungen in Spalten



- Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 7-4  
Tabelle für gestapelte Zeilenvariablen mit gemeinsamen Kategoriebeschriftungen in Spalten

	Sehr viel	Mäßig	Fast kein
Vertrauen in die Medizin	844	864	167
Vertrauen in Banken & Geldinstitute	490	1068	306
Vertrauen in das Fernsehen	196	936	744
Vertrauen in die Bildung	511	1055	315
Vertrauen in die Presse	176	878	808
Vertrauen in Großkonzerne	500	1078	243

Die Variablen müssen nicht notwendigerweise in Zeilen und die Kategorien in Spalten angezeigt werden. Sie können auch eine Tabelle erstellen, in der die Variablen in den Spalten gestapelt und die Kategorien in den Zeilen angezeigt werden. Dies ist vorteilhaft, wenn mehr Kategorien als Variablen vorhanden sind. Im vorliegenden Beispiel gibt es jedoch mehr Variablen als Kategorien.

## Table mit Prozentwerten

Bei Tabellen, in denen die Variablen in Zeilen gestapelt und die Kategorien in Spalten angezeigt werden, stellen Zeilenprozent die aussagekräftigsten (oder zumindest die am leichtesten verständlichen) Prozentwerte dar. (Bei Tabellen, in denen die Variablen in den Spalten gestapelt und die Kategorien in den Zeilen angezeigt werden, würden Sie eher Spaltenprozent wählen.)

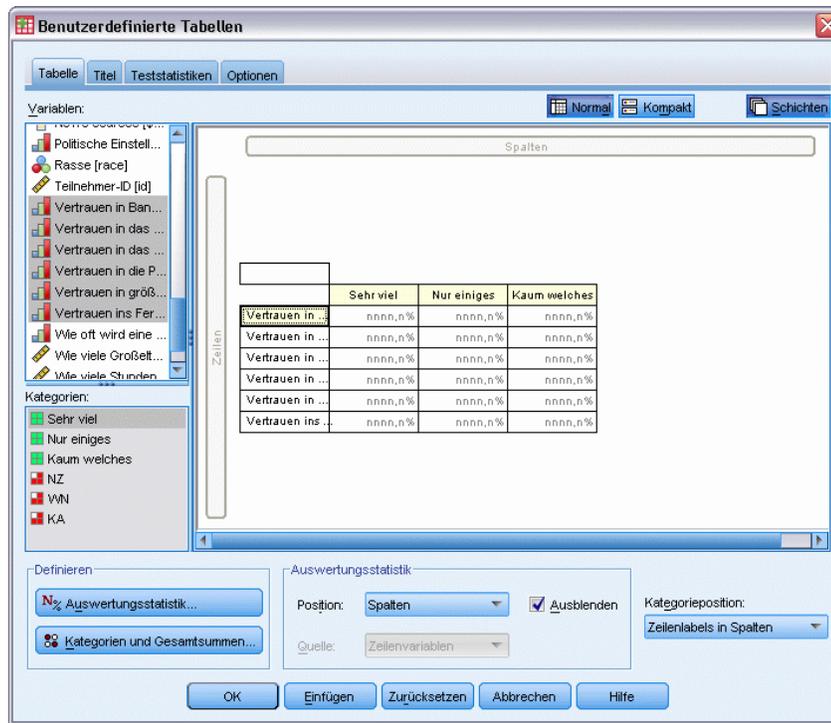
- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen erneut (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf eine der “Vertrauen”-Variablen und wählen Sie im Kontextmenü die Option Auswertungsstatistik aus.
- ▶ Wählen Sie in der Liste “Statistik” die Option Anzahl als Zeilen% aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” zu übertragen.
- ▶ Klicken Sie auf eine beliebige Zelle in der Zeile *Anzahl* und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Statistik” zu übertragen und aus der Liste “Anzeigen” zu entfernen.

Abbildung 7-5  
Ausgewählte Zeilenprozente



- ▶ Klicken Sie auf *Allen zuweisen*, um die Änderung der Auswertungsstatistik allen gestapelten Variablen in der Tabelle zuzuweisen.

Abbildung 7-6  
Zeilenprozente in der Tabellenvorschau auf der Zeichenfläche



*Hinweis:* Wenn die Tabellenvorschau nicht dieser Abbildung entspricht, haben Sie vermutlich nicht auf **Allen** zuweisen, sondern auf **Der Auswahl** zuweisen geklickt. Hierbei wird die neue Auswertungsstatistik nur der ausgewählten Variablen zugewiesen. In diesem Beispiel würden dann zwei Spalten für jede Kategorie entstehen: eine Spalte mit Platzhaltern für die Anzahl, die für alle anderen Variablen gilt, und eine Spalte mit einem Platzhalter für Zeilenprozente, die für die ausgewählte Variable gilt. Diese Tabelle würde dann erzeugt werden, jedoch ist dies im vorliegenden Beispiel *nicht* erwünscht.

- Klicken Sie auf **OK**, um die Tabelle zu erstellen.

Abbildung 7-7  
Tabelle mit Zeilenprozenten für die in Zeilen gestapelten Variablen; die Kategorien werden in Spalten angezeigt.

	Sehr viel	Mäßig	Fast kein
Vertrauen in die Medizin	45,0%	46,1%	8,9%
Vertrauen in Banken & Geldinstitute	26,3%	57,3%	16,4%
Vertrauen in das Fernsehen	10,4%	49,9%	39,7%
Vertrauen in die Bildung	27,2%	56,1%	16,7%
Vertrauen in die Presse	9,5%	47,2%	43,4%
Vertrauen in Großkonzerne	27,5%	59,2%	13,3%

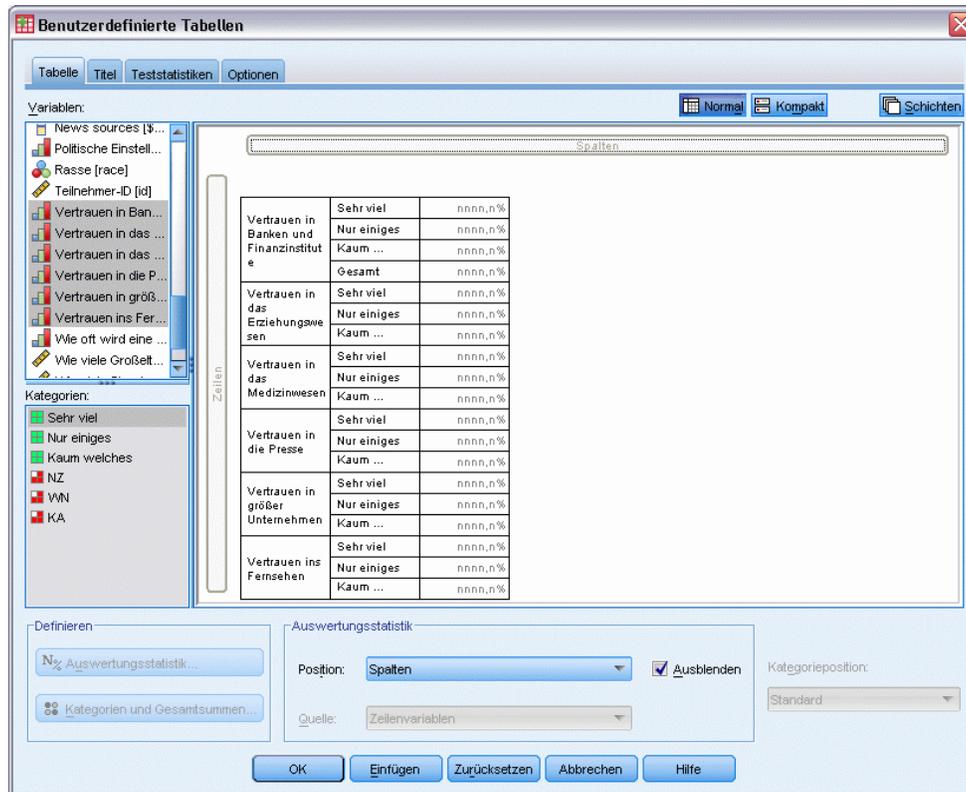
*Hinweis:* Eine Tabelle für Variablen mit gemeinsamen Kategorien kann eine beliebige Anzahl von Auswertungsstatistiken enthalten. In den vorliegenden Beispielen wird der Einfachheit halber jeweils nur eine dargestellt.

## ***Gesamtsummen und Kategorien***

Tabellen, in denen sich die Kategorien in der entgegengesetzten Dimension der Variablen befinden, können nur dann erstellt werden, wenn alle Variablen in der Tabelle dieselben Kategorien aufweisen und diese in derselben Reihenfolge angezeigt werden. Dies betrifft Gesamtsummen, Zwischenergebnisse und alle weiteren Anpassungen an den Kategorien. Dies bedeutet, dass Sie sämtliche Änderungen im Dialogfeld “Kategorien und Gesamtsummen” für alle Variablen der Tabelle vornehmen müssen, die dieselben Kategorien aufweisen.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen erneut (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf die erste “Vertrauen”-Variable und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.
- ▶ Aktivieren Sie im Dialogfeld “Kategorien und Gesamtsummen” die Option Gesamt, indem Sie darauf klicken, und klicken Sie anschließend auf Zuweisen.

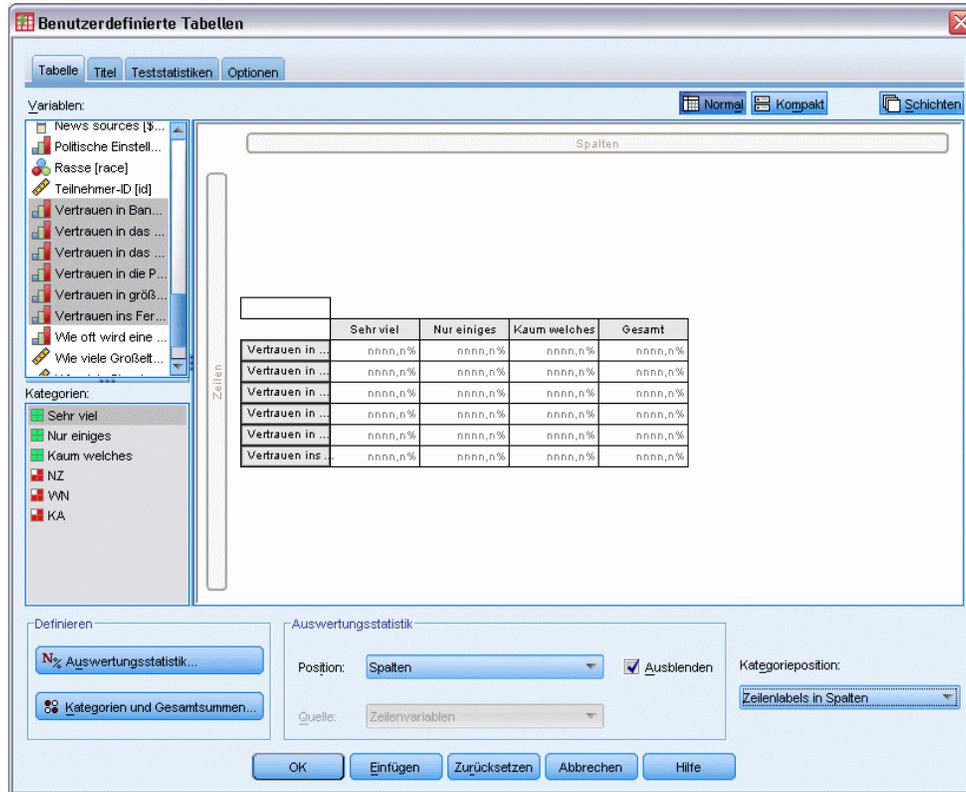
Abbildung 7-8  
Wahrscheinlich nicht die gewünschten Ergebnisse



Als erstes wird Ihnen wahrscheinlich auffallen, dass die Kategorienbeschriftungen aus den Spalten wieder in die Zeilen verschoben wurden. Sie werden vielleicht auch bemerken, dass die Dropdown-Liste "Kategorieposition" nun deaktiviert ist. Die Ursache hierfür ist, dass die Variablen nicht mehr genau die gleiche Gruppe von "Kategorien" aufweisen. Eine der Variablen besitzt nun eine Kategorie für Gesamtsummen.

- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf eine beliebige "Vertrauen"-Variable und wählen Sie im Kontextmenü die Option Alle Zeilenvariablen auswählen. Sie können auch bei gedrückter STRG-Taste auf der Zeichenfläche auf jede gestapelte Variable klicken, bis alle Variablen ausgewählt sind (möglicherweise müssen Sie einen Bildlauf nach unten durchführen oder das Fenster für die Tabellenkonstruktion vergrößern).
- ▶ Klicken Sie im Gruppenfeld "Definieren" auf Kategorien und Gesamtsummen.
- ▶ Wenn im Dialogfeld "Kategorien und Gesamtwerte" die Option Gesamt nicht bereits ausgewählt (aktiviert) ist, wählen Sie diese Option jetzt aus und klicken Sie dann auf Zuweisen.
- ▶ Die Dropdown-Liste "Kategorieposition" ist nun wieder aktiviert, da jetzt alle Variablen über die zusätzliche Kategorie für Gesamtsummen verfügen. Wählen Sie daher Zeilenlabels in Spalten aus.

Abbildung 7-9  
Kategorien und Gesamtsummen in Spalten



- Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 7-10

Tabelle mit Zeilenprozenten für die in Zeilen gestapelten Variablen; die Kategorien und Gesamtsummen werden in Spalten angezeigt.

	Sehr viel	Mäßig	Fast kein	Gesamt
Vertrauen in die Medizin	45,0%	46,1%	8,9%	100%
Vertrauen in Banken & Geldinstitute	26,3%	57,3%	16,4%	100%
Vertrauen in das Fernsehen	10,4%	49,9%	39,7%	100%
Vertrauen in die Bildung	27,2%	56,1%	16,7%	100%
Vertrauen in die Presse	9,5%	47,2%	43,4%	100%
Vertrauen in Großkonzerne	27,5%	59,2%	13,3%	100%

## Verschachteln in Tabellen mit gemeinsamen Kategorien

In verschachtelten Tabellen müssen sich die gestapelten Variablen mit den gemeinsamen Kategorien auf der innersten Verschachtelungsebene der betreffenden Dimension befinden, wenn die Kategoriebezeichnungen in der entgegengesetzten Dimension angezeigt werden sollen.

- Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen erneut (Menü "Analysieren", Untermenü "Tabellen", Befehl "Benutzerdefinierte Tabellen").

- Ziehen Sie die Variable *Geschlecht* aus der Variablenliste, und legen Sie sie an der linken Seite des Zeilenbereichs ab.

Abbildung 7-11

Verschachtelte Variablen mit gemeinsamen Kategorien

Benutzerdefinierte Tabellen

Tabellen Titel Teststatistiken Optionen

Variablen: Country of origin ... Drittes genanntes... Einkommen des B... Einstellung zur To... Eltern in diesem L... Erstes genanntes... Familienstand [ma... Gesamtes Familie... Geschlecht [sex] Glaube an Leben ... Hat eine Waffe z... Herkunftsland der... Herkunft des ...

Kategorien: Männlich Weiblich

Spalten

			Sehr viel	Nur einiges	Kaum welches	Gesamt
Geschlecht	Männlich	Vertrauen in ...	nnnn,n%	nnnn,n%	nnnn,n%	nnnn,n%
		Vertrauen in ...	nnnn,n%	nnnn,n%	nnnn,n%	nnnn,n%
		Vertrauen in ...	nnnn,n%	nnnn,n%	nnnn,n%	nnnn,n%
		Vertrauen in ...	nnnn,n%	nnnn,n%	nnnn,n%	nnnn,n%
	Weiblich	Vertrauen ins ...	nnnn,n%	nnnn,n%	nnnn,n%	nnnn,n%
		Vertrauen in ...	nnnn,n%	nnnn,n%	nnnn,n%	nnnn,n%
		Vertrauen in ...	nnnn,n%	nnnn,n%	nnnn,n%	nnnn,n%
		Vertrauen in ...	nnnn,n%	nnnn,n%	nnnn,n%	nnnn,n%

Definieren

Auswertungsstatistik

Position: Spalten  Ausblenden

Quelle: Zeilenvariablen

Kategorieposition: Zeilenlabels in Spalten

OK Einfügen Zurücksetzen Abbrechen Hilfe

In der Tabellenvorschau sind die gestapelten Variablen mit gemeinsamen Kategorien jetzt innerhalb der Geschlechtskategorien verschachtelt.

- Ziehen Sie nun die Variable *Geschlecht* rechts neben eine der gestapelten “Vertrauen”-Variablen in der Tabellenvorschau, und legen Sie sie dort ab.

Abbildung 7-12

Ein weiteres Beispiel für Ergebnisse, die Sie wahrscheinlich nicht in dieser Form wünschen

The screenshot shows the 'Benutzerdefinierte Tabellen' dialog box in SPSS. The 'Auswertungsstatistik' section is set to 'Spalten' and 'Ausblenden' is checked. The 'Kategorieposition' is set to 'Standard'. The table below shows the resulting output structure.

Zeilen	Spalten			
Vertrauen in Banken und Finanzinstitute	Sehr viel	Geschlecht	Männlich	nnnn
			Weiblich	nnnn
	Nur einiges	Geschlecht	Männlich	nnnn
			Weiblich	nnnn
	Kaum welches	Geschlecht	Männlich	nnnn
		Weiblich	nnnn	
Gesamt	Geschlecht	Männlich	nnnn	
		Weiblich	nnnn	
Vertrauen in das Erziehungswesen	Sehr viel			nnnn,n%
	Nur einiges			nnnn,n%
	Kaum welches			nnnn,n%
	Gesamt			nnnn,n%
Vertrauen in das Medizinwesen	Sehr viel			nnnn,n%
	Nur einiges			nnnn,n%
	Kaum welches			nnnn,n%
	Gesamt			nnnn,n%
Vertrauen in die Presse	Sehr viel			nnnn,n%
	Nur einiges			nnnn,n%
	Kaum welches			nnnn,n%

Die Kategoriebezeichnungen wurden wieder in die Zeilendimension verschoben und die Dropdown-Liste "Kategorieposition" ist deaktiviert. Es ist nun eine gestapelte Variable vorhanden, in der auch die Kategorie *Geschlecht* verschachtelt ist. Die anderen gestapelten Variablen enthalten hingegen keine verschachtelten Variablen. Sie könnten zwar jeder der gestapelten Variablen die Variable *Geschlecht* als verschachtelte Variable hinzufügen. Beim Verschieben der Zeilenbeschriftungen in die Spalten würden dann jedoch nicht die Kategoriebeschriftungen für die gestapelten Variablen mit den gemeinsamen Kategorien, sondern die Kategoriebeschriftungen für *Geschlecht* in den Spalten angezeigt werden. Der Grund hierfür besteht darin, dass *Geschlecht* nun die innerste verschachtelte Variable bilden würde und sich Änderungen an der Kategorieposition immer auf die innerste verschachtelte Variable auswirkt.

# Auswertungsstatistik

Zu den **Auswertungsstatistiken** gehören verschiedene Statistiken, von einfachen Häufigkeiten kategorialer Variablen bis hin zu Streuungsmaßen, z. B. der Standardfehler des Mittelwerts für metrische Variablen. *Nicht* hierzu zählen jedoch Signifikanztests, die im Dialogfeld “Benutzerdefinierte Tabellen” auf der Registerkarte “Teststatistiken” verfügbar sind. [Für weitere Informationen siehe Thema Teststatistiken in Kapitel 10 auf S. 148.](#)

Als Auswertungsstatistiken für kategoriale Variablen und Mehrfachantworten-Sets sind Häufigkeiten und eine breite Auswahl von Prozentrechnungen verfügbar. Dazu gehören:

- Zeilenprozente
- Spaltenprozente
- Untertabellenprozente
- Tabellenprozente
- Gültige N als Prozente

Neben den Auswertungsstatistiken für kategoriale Variablen sind die folgenden Auswertungsstatistiken für metrische Variablen und benutzerdefinierten Gesamtstatistiken für kategoriale Variablen verfügbar:

- Mittelwert
- Median
- Perzentile (Explore)
- Summe
- Standardabweichung
- OK (Opening File Options)
- Minimal- und Maximalwerte

Für Mehrfachantworten-Sets sind weitere Auswertungsstatistiken verfügbar. [Für weitere Informationen siehe Thema Häufigkeiten, Antworten, Prozentwerte und Gesamtwerte in Kapitel 11 auf S. 169.](#) Außerdem ist eine vollständige Liste von Auswertungsstatistiken verfügbar. [Für weitere Informationen siehe Thema Auswertungsstatistik in Kapitel 2 auf S. 35.](#)

## **Beispiel-Datendatei**

Für die Beispiele in diesem Kapitel wird die Datendatei *survey\_sample.sav* verwendet. [Für weitere Informationen siehe Thema Beispieldateien in Anhang A auf S. 201.](#)

In allen angegebenen Beispielen werden die Variablenlabels in Dialogfeldern in alphabetischer Reihenfolge angezeigt. Die Anzeigeeigenschaften für Variablenlisten werden im Dialogfeld “Optionen” auf der Registerkarte “Allgemein” festgelegt (Menü “Bearbeiten”, Befehl “Optionen”).

## **Quellvariable für Auswertungsstatistik**

Es hängt vom Messniveau der jeweiligen Quellvariablen ab, welche Auswertungsstatistiken verfügbar sind. Die Quelle der Auswertungsstatistiken (die Variable, auf der die Auswertungsstatistiken beruhen) wird durch die folgenden Faktoren bestimmt:

- **Messniveau.** Wenn eine Tabelle (oder ein Tabellenabschnitt in einer gestapelten Tabelle) eine metrische Variable enthält, beruhen Auswertungsstatistiken auf der metrischen Variablen.
- **Reihenfolge der Variablenauswahl.** Die Standard-Quelldimension der Statistik (Zeile oder Spalte) für kategoriale Variablen beruht auf der Reihenfolge, in der Sie die Variablen auf die Zeichenfläche ziehen. Wenn Sie z. B. eine Variable zuerst in den Bereich für die Zeilen ziehen, wird die Zeilendimension als Standard-Quelldimension für die Statistik verwendet.
- **Verschachtelung.** Bei kategorialen Variablen beruhen Auswertungsstatistiken auf der innersten Variablen in der Quelldimension der Statistik.

Eine gestapelte Tabelle kann mehrere Quellvariablen für Auswertungsstatistiken aufweisen (sowohl metrisch als auch kategorial); jeder Tabellenabschnitt besitzt jedoch nur eine Quelle für die Auswertungsstatistik.

### **Quelle von Auswertungsstatistiken bei kategorialen Variablen**

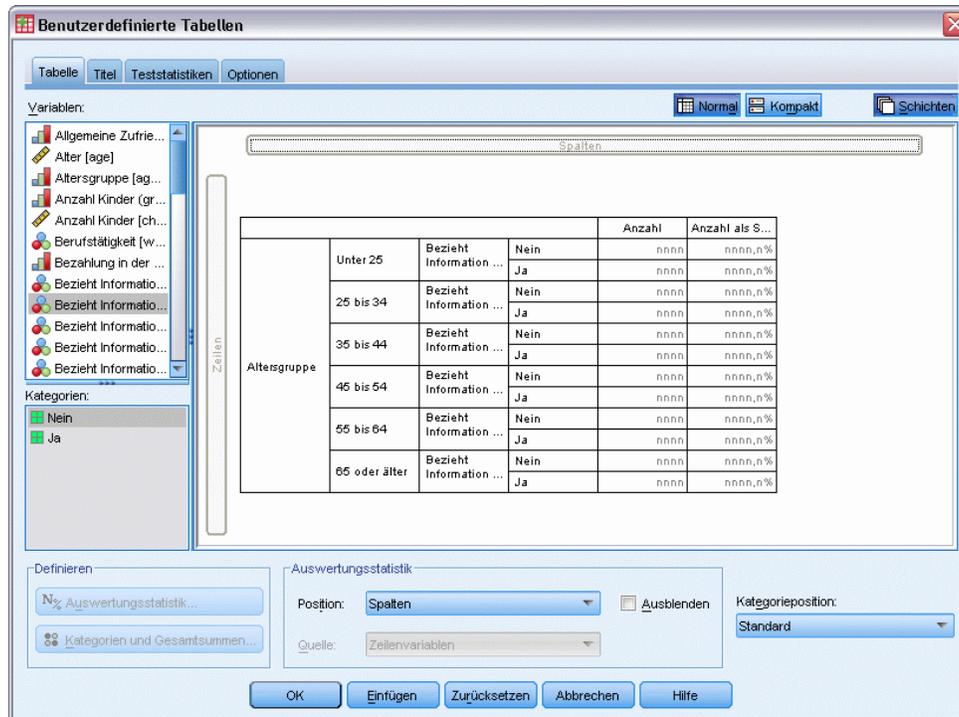
- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
Analysieren > Tabellen > Benutzerdefinierte Tabellen...
- ▶ Ziehen Sie die Variable *Alterskategorie* aus der Variablenliste der Tabellenkonstruktion in den Bereich für die Zeilen auf der Zeichenfläche.
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option Auswertungsstatistik aus. (Da es sich um die einzige Variable in der Tabelle handelt, wird sie als deren Quellvariable verwendet.)
- ▶ Wählen Sie im Dialogfeld “Auswertungsstatistik” in der Liste “Statistik” die Option *Anzahl als Spalten%* aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” zu übertragen.

Abbildung 8-1  
Dialogfeld "Auswertungsstatistik" für kategoriale Variablen



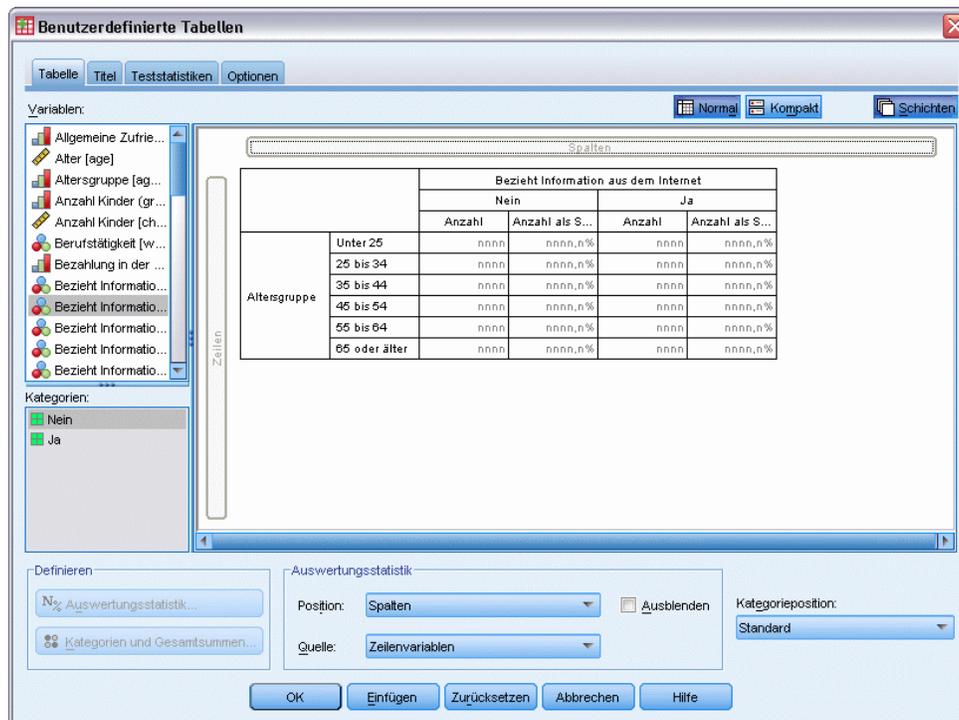
- Klicken Sie auf Der Auswahl zuweisen.
- Ziehen Sie in der Tabellenkonstruktion die Variable *Nachrichten aus dem Internet* auf die Zeichenfläche, und legen Sie sie rechts neben *Alterskategorie* ab.

Abbildung 8-2  
Verschachtelte kategoriale Variablen



- ▶ Klicken Sie mit der rechten Maustaste erneut auf der Zeichenfläche auf *Alterskategorie*. Die Option Auswertungsstatistik im Kontextmenü ist nun deaktiviert, da *Alterskategorie* nicht die innerste Variable in der Quelldimension der Statistik ist.
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Nachrichten aus dem Internet*. Die Option Auswertungsstatistik ist aktiviert, weil diese Variable nun die innerste verschachtelte Variable der Quelldimension der Statistik ist und somit die Quellvariable der Auswertungsstatistik darstellt. (Da die Tabelle nur eine Dimension – Zeilen – besitzt, ist diese die Quelldimension der Statistik.)
- ▶ Ziehen Sie die Variable *Nachrichten aus dem Internet* auf der Zeichenfläche aus dem Bereich für die Zeilen, und legen Sie sie im Bereich für die Spalten ab.

Abbildung 8-3  
Kategoriale Variablen in Kreuztabellen



- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche erneut auf *Nachrichten aus dem Internet*. Die Option Auswertungsstatistik im Kontextmenü ist nun deaktiviert, da sich die Variable nicht mehr in der Quelldimension der Statistik befindet.

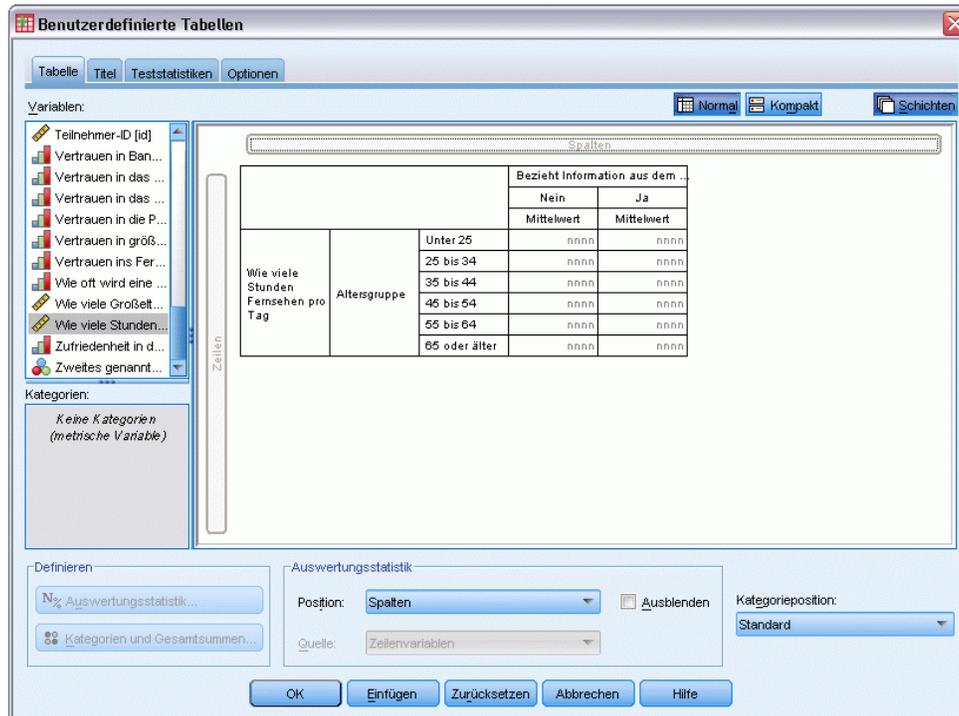
Jetzt ist *Alterskategorie* wieder die Quellvariable der Statistik, da die Standard-Quelldimension der Statistik für kategoriale Variablen die erste Dimension ist, in der Sie beim Erstellen der Tabelle Variablen ablegen. In diesem Beispiel wurden zuerst Variablen in der Zeilendimension abgelegt. Die Zeilendimension ist daher die Standard-Quelldimension der Statistik. Da *Alterskategorie* nun die einzige Variable in dieser Dimension ist, wird sie als Quellvariable der Statistik verwendet.

## Quelle von Auswertungsstatistiken für metrische Variablen

- ▶ Ziehen Sie die metrische Variable *Täglicher Fernsehkonsum in Stunden* aus dem Zeilenbereich der Zeichenfläche, und legen Sie sie links neben der Variablen *Alterskategorie* ab.

Abbildung 8-4

Kreuztabellen mit metrischer Variable für die Auswertungsstatistik



Zunächst ist erkennbar, dass die Auswertungen *Anzahl* und *Anzahl als Spalten%* durch *Mittelwert* ersetzt wurden. Wenn Sie mit der rechten Maustaste auf der Zeichenfläche auf *Täglicher Fernsehkonsum in Stunden* klicken, wird sichtbar, dass diese Variable nun als Quellvariable der Auswertungsstatistik verwendet wird. In Tabellen mit einer metrischen Variablen wird diese stets als Quellvariable der Statistik verwendet. Dies gilt unabhängig von der Verschachtelungsebene oder Dimension, und als Standard-Auswertungsstatistik für metrische Variablen dient der Mittelwert.

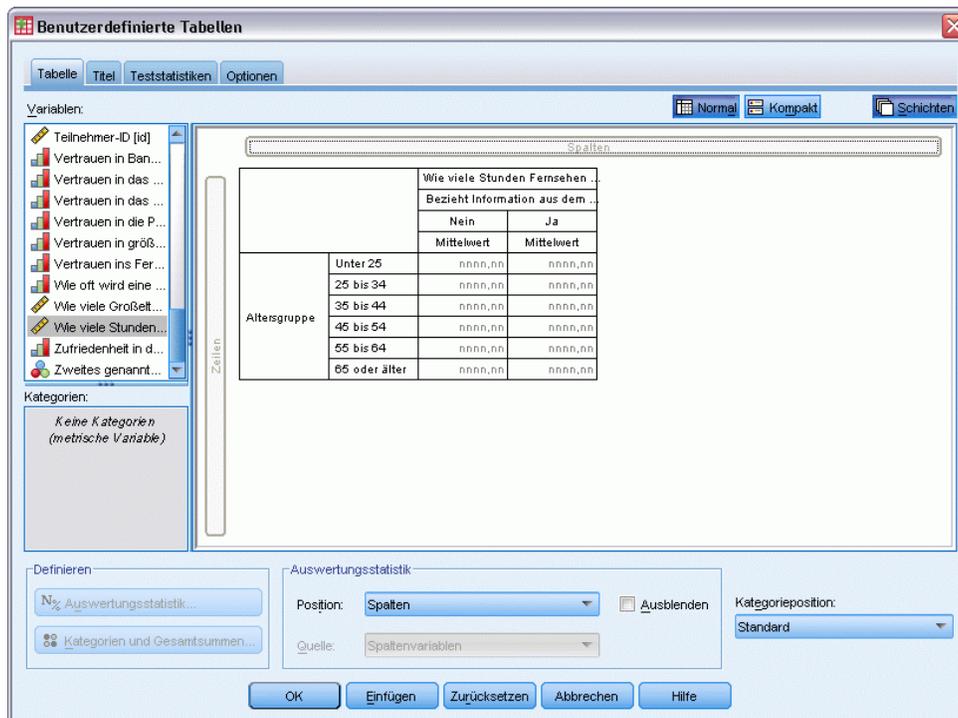
- ▶ Ziehen Sie die Variable *Täglicher Fernsehkonsum in Stunden* aus dem Zeilenbereich, und legen Sie sie im Spaltenbereich über der Variablen *Nachrichten aus dem Internet* ab.
- ▶ Klicken Sie mit der rechten Maustaste auf *Täglicher Fernsehkonsum in Stunden* und wählen Sie im Kontextmenü die Option *Auswertungsstatistik* aus. (Diese bleibt als Quellvariable der Statistik erhalten, auch wenn Sie sie in eine andere Dimension verschieben.)
- ▶ Klicken Sie im Dialogfeld "Auswertungsstatistik" in der Liste "Anzeigen" auf die Zelle *Format* für den Mittelwert und wählen Sie aus der Dropdown-Liste "Format" den Eintrag *nnnn* aus. (Möglicherweise müssen Sie einen Bildlauf durch die Liste durchführen, um diese Option zu finden.)
- ▶ Geben Sie in der Zelle "Dezimalstellen" den Wert 2 ein.

Abbildung 8-5  
Dialogfeld "Auswertungsstatistik" für metrische Variablen



- Klicken Sie auf Der Auswahl zuweisen.

Abbildung 8-6  
Metrische Auswertungsstatistik mit zwei Dezimalstellen



Die Tabellenvorschau auf der Zeichenfläche verdeutlicht nun, dass die Mittelwerte mit zwei Dezimalstellen angezeigt werden.

- Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 8-7

Auswertung einer metrischen Variablen innerhalb von kategorialen Variablen in einer Kreuztabelle

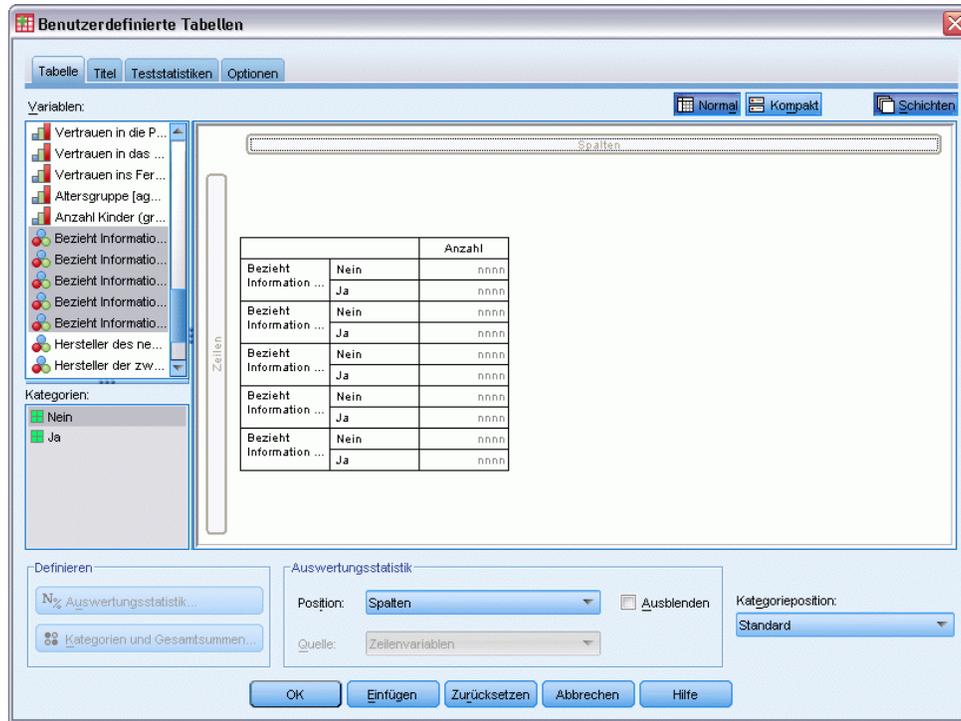
		Täglicher Fernsehkonsum in Stunden	
		Nachrichten aus dem Internet	
		Nein	Ja
		Mittelwert	Mittelwert
Alterskategorie	Weniger als 25	3,54	2,12
	25 bis 34	3,42	2,14
	35 bis 44	3,00	2,01
	45 bis 54	2,83	2,06
	55 bis 64	3,24	2,37
	65 oder älter	3,82	2,33

## Gestapelte Variablen

Da eine gestapelte Tabelle mehrere Quellvariablen für Statistiken enthalten kann und Sie für diese Quellvariablen unterschiedliche Auswertungsstatistiken festlegen können, müssen beim Festlegen von Auswertungsstatistiken in gestapelten Tabellen einige besondere Überlegungen getroffen werden.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie auf Zurücksetzen, um frühere Einstellungen in der Tabellenkonstruktion zu löschen.
- ▶ Klicken Sie in der Variablenliste auf *Nachrichten aus dem Internet* und klicken Sie bei gedrückter Umschalttaste in der Variablenliste auf *Nachrichten aus dem Fernsehen*, um alle “Nachrichten”-Variablen auszuwählen. (*Anmerkung:* Hierbei wird vorausgesetzt, dass die Variablenlabels in der Variablenliste in alphabetischer Reihenfolge und nicht in der Reihenfolge angezeigt werden, in der sich in der Datei befinden.)
- ▶ Ziehen Sie die fünf “Nachrichten”-Variablen in den Bereich für die Zeilen auf der Zeichenfläche, und legen Sie sie dort ab.

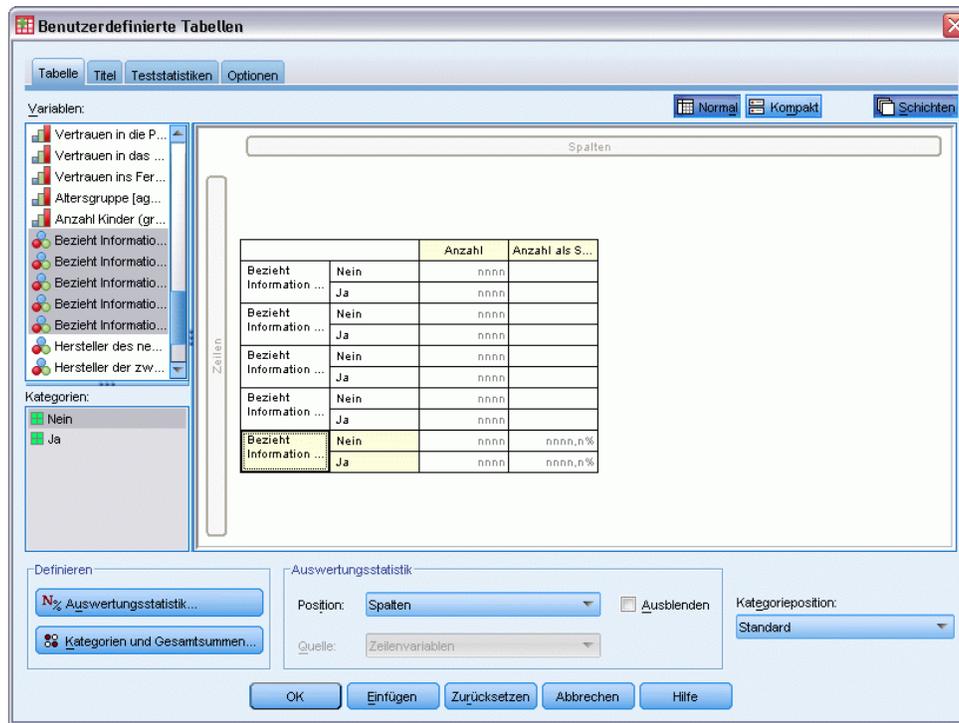
Abbildung 8-8  
Stapeln der "Nachrichten"-Variablen in Zeilen



Die fünf "Nachrichten"-Variablen werden in der Zeilendimension gestapelt.

- ▶ Klicken Sie auf der Zeichenfläche auf *Nachrichten aus dem Internet*, sodass nur diese Variable ausgewählt ist.
- ▶ Klicken Sie jetzt mit der rechten Maustaste auf *Nachrichten aus dem Internet* und wählen Sie im Kontextmenü die Option *Auswertungsstatistik* aus.
- ▶ Wählen Sie im Dialogfeld "Auswertungsstatistik" in der Liste "Statistik" die Option *Anzahl als Spalten%* aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste "Anzeigen" zu übertragen. (Mit der Pfeilschaltfläche können Sie ausgewählte Statistiken aus der Liste "Statistik" in die Liste "Anzeigen" übertragen. Sie können die Statistiken auch aus der Liste "Statistik" in die Liste "Anzeigen" ziehen und dort ablegen.)
- ▶ Klicken Sie dann auf *Der Auswahl zuweisen*.

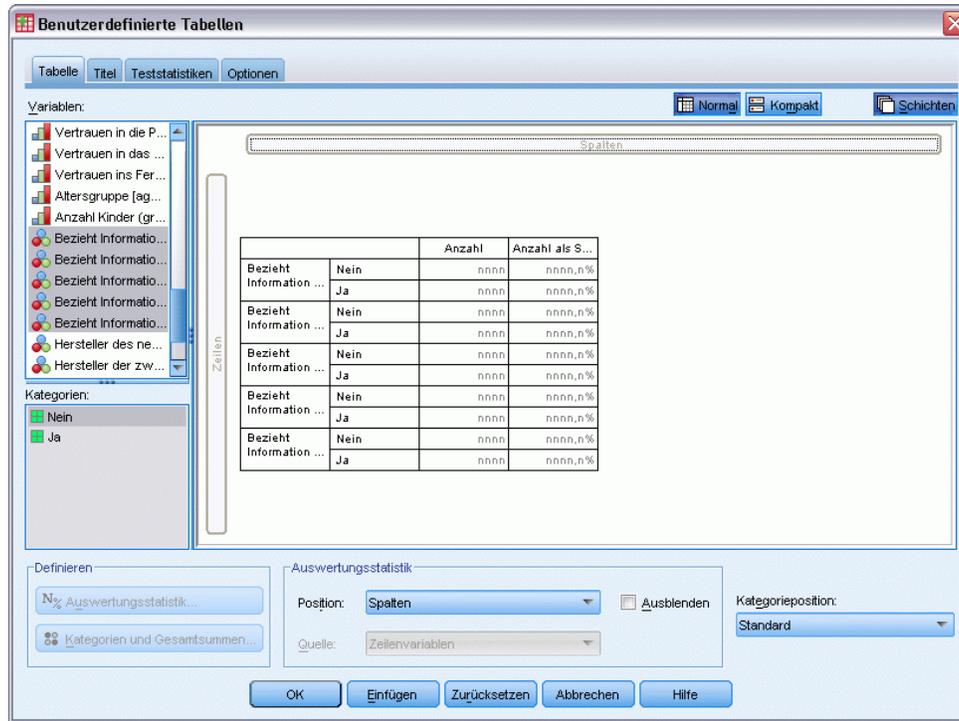
Abbildung 8-9  
Zuweisen von zusätzlichen Statistiken zu einer Variablen in einer gestapelten Tabelle



Eine Spalte für Spaltenprozentage wird hinzugefügt. In der Tabellenvorschau auf der Zeichenfläche wird jedoch angezeigt, dass Spaltenprozentage nur für eine Variable angezeigt werden. Der Grund dafür besteht darin, dass in einer gestapelten Tabelle mehrere Quellvariablen für Statistiken vorhanden sind und für jede dieser Variablen eine andere Auswertungsstatistik vorhanden sein kann. In diesem Beispiel soll jedoch für alle Variablen die gleiche Auswertungsstatistik angezeigt werden.

- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf *Nachrichten aus Zeitungen* und wählen Sie im Kontextmenü die Option *Auswertungsstatistik* aus.
- ▶ Wählen Sie im Dialogfeld "Auswertungsstatistik" in der Liste "Statistik" die Option *Anzahl als Spalten%* aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste "Anzeigen" zu übertragen.
- ▶ Klicken Sie dann auf *Allen zuweisen*.

Abbildung 8-10  
Zuweisen von zusätzlichen Statistiken zu allen Variablen in einer gestapelten Tabelle



In der Tabellenvorschau wird nun dargestellt, dass Spaltenprozent für alle gestapelten Variablen angezeigt werden.

## Angepasste Auswertungsstatistik für Gesamtsummen bei kategorialen Variablen

Bei kategorialen Quellvariablen für Statistiken können Sie angepaßte Auswertungsstatistiken für Gesamtsummen einbeziehen, die sich von den Statistiken unterscheiden, die für die Kategorien der Variablen gezeigt werden. Bei einer ordinalen Variablen können Sie beispielsweise Prozente für jede Kategorie sowie den Mittelwert oder Median als angepaßte Auswertungsstatistik für Gesamtsummen anzeigen.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü "Analysieren", Untermenü "Tabellen", Befehl "Benutzerdefinierte Tabellen").
- ▶ Klicken Sie auf Zurücksetzen, um frühere Einstellungen in der Tabellenkonstruktion zu löschen.
- ▶ Klicken Sie in der Variablenliste auf *Vertrauen in die Presse* und klicken Sie dann bei gedrückter STRG-Taste auf *Vertrauen in das Fernsehen*, um beide Variablen auszuwählen.
- ▶ Ziehen Sie die beiden Variablen in den Bereich für die Zeilen auf der Zeichenfläche, und legen Sie sie dort ab. Hiermit werden die beiden Variablen in der Zeilendimension gestapelt.

- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf eine der Variablen und wählen Sie im Kontextmenü die Option *Alle Zeilenvariablen auswählen* aus. (Dies ist nur als Absicherung gedacht; möglicherweise sind die Variablen bereits ausgewählt.)
- ▶ Klicken Sie mit der rechten Maustaste erneut auf die Variable und wählen Sie im Kontextmenü die Option *Kategorien und Gesamtsummen* aus.
- ▶ Wählen Sie im Dialogfeld “Kategorien und Gesamtsummen” die Option *Gesamt* aus (aktivieren Sie diese) und klicken Sie anschließend auf *Zuweisen*.  

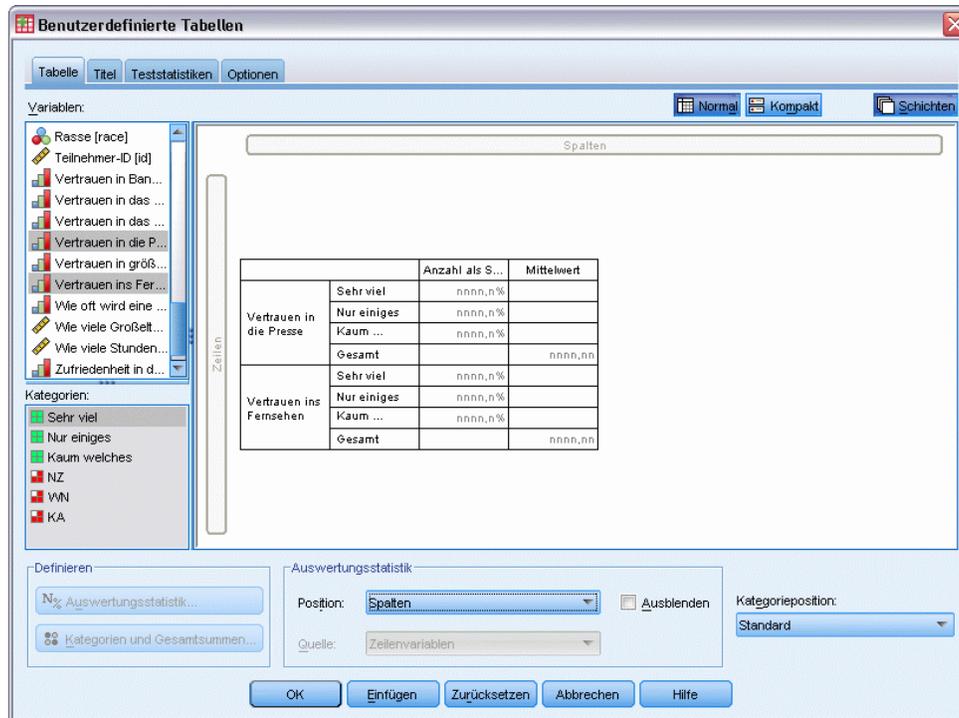
In der Tabellenvorschau auf der Zeichenfläche wird nun eine Zeile für die Gesamtsumme für beide Variablen angezeigt. Um eine angepaßte Auswertungsstatistik für die Gesamtsumme anzuzeigen, müssen für die Tabelle *Gesamtsummen* bzw. *Zwischenergebnisse* festgelegt werden.
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf eine der Variablen und wählen Sie im Kontextmenü die Option *Auswertungsstatistik* aus.
- ▶ Klicken Sie im Dialogfeld “Auswertungsstatistik” in der Liste “Anzeigen” auf die Option *Anzahl* und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Statistik” zu übertragen und aus der Liste “Anzeigen” zu entfernen.
- ▶ Klicken Sie in der Liste “Statistik” auf *Anzahl als Spalten%* und klicken Sie auf die Pfeilschaltfläche, um die Variable in die Liste “Anzeigen” zu verschieben.
- ▶ Klicken Sie auf die Option *Angepaßte Auswertungsstatistik für Gesamtsummen und Zwischenergebnisse* (aktivieren Sie diese).
- ▶ Klicken Sie in der Liste “Anzeigen” der benutzerdefinierten Auswertung auf die Option *Anzahl* und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Statistik” der benutzerdefinierten Auswertung zu übertragen und aus der Liste “Anzeigen” zu entfernen.
- ▶ Klicken Sie in der Liste “Statistik” der benutzerdefinierten Auswertung auf die Option *Mittelwert* und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” der benutzerdefinierten Auswertung zu übertragen.
- ▶ Klicken Sie in der Liste “Anzeigen” auf die Zelle *Format für den Mittelwert* und wählen Sie aus der Dropdown-Liste der Formate den Eintrag *nnn* aus. (Möglicherweise müssen Sie einen Bildlauf durch die Liste durchführen, um diese Option zu finden.)
- ▶ Geben Sie in der Zelle “Dezimalstellen” den Wert 2 ein.

Abbildung 8-11  
Auswählen von angepassten Auswertungsstatistiken für Gesamtsummen



- Klicken Sie auf Allen zuweisen, um diese Einstellungen beiden Variablen in der Tabelle zuzuweisen.

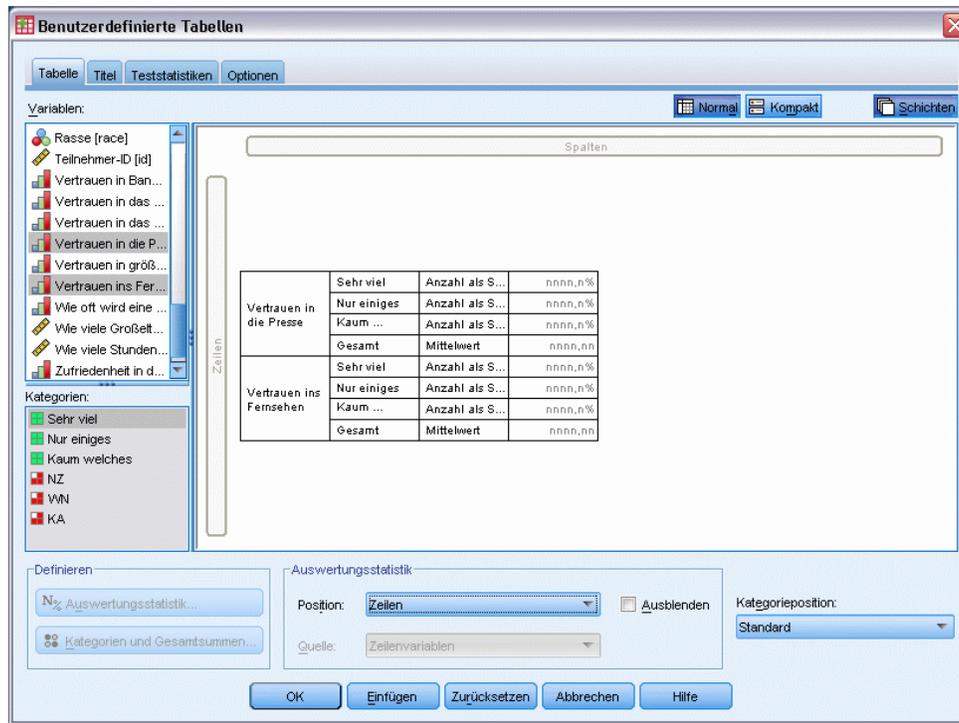
Abbildung 8-12  
 Angepasste Auswertungsstatistik für Gesamtsummen bei Zeilenvariablen, die in Spalten angezeigt werden



Für die angepasste Auswertungsstatistik für Gesamtsummen wurde eine neue Spalte hinzugefügt. Dies ist jedoch möglicherweise unerwünscht, da in der Vorschau auf der Zeichenfläche deutlich gezeigt wird, dass die Tabelle hierbei eine große Anzahl leerer Zellen enthält.

- Wählen Sie im Dialogfeld zum Konstruieren von Tabellen im Gruppenfeld "Auswertungsstatistik" den Eintrag Zeilen aus der Dropdown-Liste "Position" aus.

Abbildung 8-13  
Auswertungsstatistik für Zeilenvariablen, die in Zeilen angezeigt werden



Hiermit verschieben Sie alle Auswertungsstatistiken in die Zeilendimension, wodurch diese in einer einzigen Spalte der Tabelle angezeigt werden.

- Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 8-14  
Kategoriale Variablen mit angepasster Auswertungsstatistik für Gesamtsummen

Vertrauen in das Fernsehen	Sehr viel	Spalten%	10,4%
	Mäßig	Spalten%	49,9%
	Fast kein	Spalten%	39,7%
	Gesamt	Mittelwert	2,29
Vertrauen in die Presse	Sehr viel	Spalten%	9,5%
	Mäßig	Spalten%	47,2%
	Fast kein	Spalten%	43,4%
	Gesamt	Mittelwert	2,34

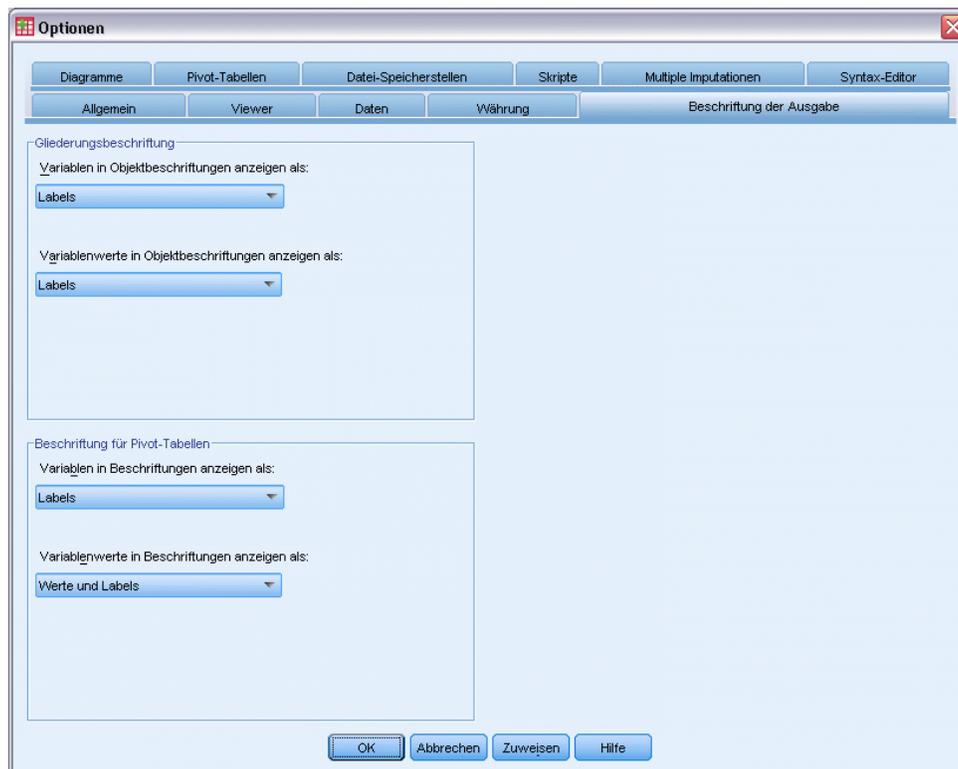
## Anzeigen von Kategoriewerten

Im Zusammenhang mit der oben beschriebenen Tabelle besteht nur ein kleines Problem: Es kann schwierig sein, den Mittelwert zu interpretieren, ohne die zugrunde liegenden Kategoriewerte zu kennen. Befindet sich ein Mittelwert von 2,34 zwischen *Sehr viel* und *Mäßig* oder zwischen *Mäßig* und *Fast kein*?

Dieses Problem kann zwar nicht direkt für die benutzerdefinierte Tabelle, aber in allgemeinerer Form gelöst werden.

- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
Bearbeiten > Optionen...
- ▶ Klicken Sie im Dialogfeld “Optionen” auf die Registerkarte Beschriftung der Ausgabe.
- ▶ Wählen Sie im Gruppenfeld “Beschriftung für Pivot-Tabellen” aus der Dropdown-Liste Variablenwerte in Beschriftungen anzeigen als die Option Werte und Labels aus.

Abbildung 8-15  
Optionen für die Beschriftung der Ausgabe



- ▶ Klicken Sie auf OK, um diese Einstellung zu speichern.
- ▶ Öffnen Sie erneut das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”) und klicken Sie auf OK, um die Tabelle erneut zu erstellen.

Abbildung 8-16  
Werte und Labels, die für Variablenkategorien angezeigt werden

Vertrauen in das Fernsehen	1	Sehr viel	Spalten%	10,4%
	2	Mäßig	Spalten%	49,9%
	3	Fast kein	Spalten%	39,7%
		Gesamt	Mittelwert	2,29
Vertrauen in die Presse	1	Sehr viel	Spalten%	9,5%
	2	Mäßig	Spalten%	47,2%
	3	Fast kein	Spalten%	43,4%
		Gesamt	Mittelwert	2,34

Die Kategoriewerte zeigen eindeutig an, dass ein Mittelwert von 2,34 zwischen *Mäßig* und *Fast kein* liegt. Durch Anzeigen der Kategoriewerte in der Tabelle wird die Interpretation der angepassten Auswertungsstatistik für Gesamtsummen (z. B. des Mittelwerts) erleichtert.

Diese Anzeigeeinstellung ist eine globale Einstellung, die sich auf die ausgegebenen Pivot-Tabellen aller Prozeduren auswirkt. Die Einstellung wird in allen künftigen Sitzungen beibehalten, bis Sie sie ausdrücklich ändern. So setzen Sie die Einstellung zurück auf die ausschließliche Anzeige von Wertelabels:

- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
Bearbeiten > Optionen...
- ▶ Klicken Sie im Dialogfeld “Optionen” auf die Registerkarte Beschriftung der Ausgabe.
- ▶ Wählen Sie im Gruppenfeld “Beschriftung für Pivot-Tabellen” aus der Dropdown-Liste Variablenwerte in Beschriftungen anzeigen als die Option Labels aus.
- ▶ Klicken Sie auf OK, um diese Einstellung zu speichern.

# Auswerten von metrischen Variablen

Es stehen eine Reihe von Auswertungsstatistiken für metrische Variablen zur Verfügung. Neben Häufigkeiten und Prozentwerten, die für kategoriale Variablen zur Verfügung stehen, sind die folgenden Auswertungsstatistiken für metrische Variablen verfügbar:

- Mittelwert
- Median
- Perzentile (Explore)
- Summe
- Standardabweichung
- OK (Opening File Options)
- Minimal- und Maximalwerte

Für weitere Informationen siehe [Thema Auswertungsstatistiken für metrische Variablen und angepasste Gesamtsummen für kategoriale Variablen in Kapitel 2 auf S. 39](#).

## **Beispiel-Datendatei**

Für die Beispiele in diesem Kapitel wird die Datendatei *survey\_sample.sav* verwendet. [Für weitere Informationen siehe Thema Beispieldateien in Anhang A auf S. 201](#).

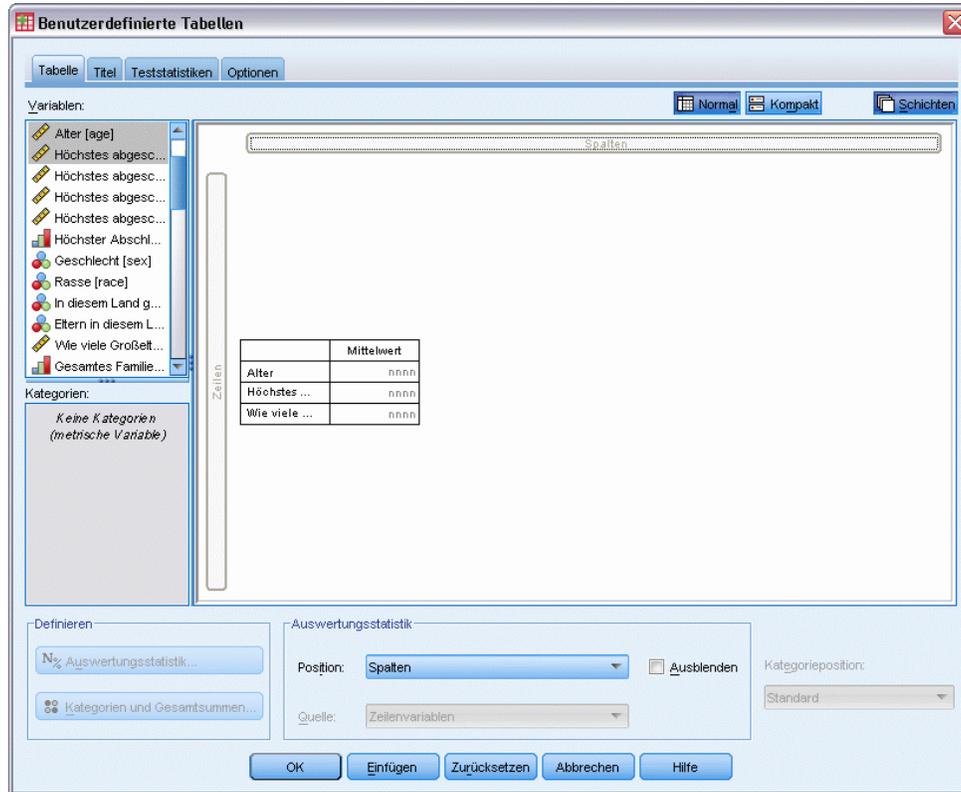
In allen angegebenen Beispielen werden die Variablenlabels in Dialogfeldern in alphabetischer Reihenfolge angezeigt. Die Anzeigeeigenschaften für Variablenlisten werden im Dialogfeld “Optionen” auf der Registerkarte “Allgemein” festgelegt (Menü “Bearbeiten”, Befehl “Optionen”).

## **Gestapelte metrische Variablen**

Durch Stapeln können mehrere metrische Variablen in derselben Tabelle gleichzeitig ausgewertet werden.

- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
Analysieren > Tabellen > Benutzerdefinierte Tabellen...
- ▶ Klicken Sie im Dialogfeld zum Konstruieren von Tabellen in der Variablenliste auf *Alter der befragten Person*, klicken Sie anschließend mit gedrückter STRG-Taste auf *Letztes abgeschlossenes Schuljahr* und klicken Sie dann mit gedrückter STRG-Taste auf *Täglicher Fernsehkonsum in Stunden*, um alle drei Variablen auszuwählen.
- ▶ Ziehen Sie die drei ausgewählten Variablen auf die Zeichenfläche, und legen Sie sie im Zeilenbereich ab.

Abbildung 9-1  
Gestapelte metrische Variablen im Dialogfeld zum Konstruieren von Tabellen



Die drei Variablen werden in der Zeilendimension gestapelt. Da es sich bei allen drei Variablen um metrische Variablen handelt, werden keine Kategorien angezeigt. Die Standard-Auswertungsstatistik ist der Mittelwert.

- Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 9-2  
Tabelle mit Mittelwerten von gestapelten Variablen

	Mittelwert
Alter der befragten Person	46
Letztes abgeschlossenes Schuljahr	13
Täglicher Fernsehkonsum in Stunden	3

## Mehrere Auswertungsstatistiken

In der Standardeinstellung wird für metrische Variablen der Mittelwert angezeigt. Es können aber auch andere Statistiken für metrische Variablen ausgewählt werden, und es können mehrere Auswertungsstatistiken gleichzeitig angezeigt werden.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf eine der drei metrischen Variablen und wählen Sie im Kontextmenü die Option Auswertungsstatistik.
- ▶ Wählen Sie im Dialogfeld “Auswertungsstatistik” in der Liste “Statistik” die Option *Median* aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” zu übertragen. (Mit der Pfeilschaltfläche können Sie ausgewählte Statistiken aus der Liste “Statistik” in die Liste “Anzeigen” übertragen. Sie können die Statistiken auch aus der Liste “Statistik” in die Liste “Anzeigen” ziehen und dort ablegen.)
- ▶ Klicken Sie in der Liste “Anzeigen” auf die Zelle Format für den Median und wählen Sie aus der Dropdown-Liste der Formate den Eintrag nnnn.n aus.
- ▶ Geben Sie in der Zelle “Dezimalstellen” den Wert 1 ein.
- ▶ Nehmen Sie die gleichen Änderungen für den Mittelwert in der Liste “Anzeigen” vor.

Abbildung 9-3

Im Dialogfeld “Auswertungsstatistik” ist der Median und der Mittelwert ausgewählt.



- ▶ Klicken Sie auf *Allen zuweisen*, um diese Änderung allen drei metrischen Variablen zuzuweisen.
- ▶ Klicken Sie im Dialogfeld zum Konstruieren von Tabellen auf *OK*, um die Tabelle zu erstellen.

Abbildung 9-4

In der Tabelle mit gestapelten metrischen Variablen werden der Mittelwert und der Median angezeigt.

	Mittelwert	Median
Alter der befragten Person	45,6	42,0
Letztes abgeschlossenes Schuljahr	13,3	13,0
Täglicher Fernsehkonsum in Stunden	2,9	2,0

## Anzahl, gültige N und fehlende Werte

Es ist häufig sinnvoll, die bei der Berechnung von Auswertungsstatistiken wie dem Mittelwert verwendete Anzahl der Fälle anzuzeigen. Vermutlich gehen Sie (nicht ganz ohne Grund) davon aus, dass die Auswertungsstatistik *Anzahl* diese Informationen bereitstellt. Wenn Werte fehlen,

ergibt dies jedoch keine korrekte Fallbasis. Um eine korrekte Fallbasis zu erhalten, verwenden Sie *Gültige N*.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf eine der drei metrischen Variablen und wählen Sie im Kontextmenü die Option Auswertungsstatistik.
- ▶ Wählen Sie im Dialogfeld “Auswertungsstatistik” in der Liste “Statistik” die Option Anzahl aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” zu übertragen.
- ▶ Wählen Sie anschließend in der Liste “Statistik” die Option Gültige N aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” zu übertragen.
- ▶ Klicken Sie auf Allen zuweisen, um diese Änderung allen drei metrischen Variablen zuzuweisen.
- ▶ Klicken Sie im Dialogfeld zum Konstruieren von Tabellen auf OK, um die Tabelle zu erstellen.

Abbildung 9-5  
*Anzahl und gültige N*

	Mittelwert	Median	Anzahl	Gültige N
Age of respondent	45.6	42.0	2832	2828
Highest year of school completed	13.3	13.0	2832	2820
Hours per day watching TV	2.9	2.0	2832	2337

*Anzahl* liefert für alle drei Variablen den gleichen Wert: 2,832. Es ist kein Zufall, dass dieser Wert mit der Gesamtanzahl der Fälle in der Datendatei übereinstimmt. Da die metrischen Variablen in keiner kategorialen Variablen verschachtelt sind, stellt *Anzahl* einfach die Gesamtanzahl der Fälle in der Datendatei dar.

*Gültige N* ergibt dagegen für jede Variable einen anderen Wert, der sich bei *Täglicher Fernsehkonsum in Stunden* erheblich von *Anzahl* unterscheidet. Dieser Unterschied resultiert aus der großen Anzahl von **fehlenden Werten** für diese Variable. Diese sind auf Fälle zurückzuführen, für die in dieser Variablen kein Wert aufgezeichnet wurde oder für die Werte definiert wurden, die fehlende Daten darstellen (z. B. der Code “99”, der für die Schwangerschaft bei Männern *nicht zutreffend* angibt).

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf eine der drei metrischen Variablen und wählen Sie im Kontextmenü die Option Auswertungsstatistik.
- ▶ Wählen Sie im Dialogfeld “Auswertungsstatistik” in der Liste “Anzeigen” die Option Gültige N aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Statistik” zu übertragen und aus der Liste “Anzeigen” zu entfernen.
- ▶ Wählen Sie in der Liste “Anzeigen” die Option Anzahl aus und klicken Sie auf die Pfeilschaltfläche, um sie zurück in die Liste “Statistik” zu verschieben und aus der Liste “Anzeigen” zu entfernen.
- ▶ Wählen Sie in der Liste “Statistik” die Option Fehlend aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” zu übertragen.

- ▶ Klicken Sie auf **Allen zuweisen**, um diese Änderung allen drei metrischen Variablen zuzuweisen.
- ▶ Klicken Sie im Dialogfeld zum Konstruieren von Tabellen auf **OK**, um die Tabelle zu erstellen.

**Abbildung 9-6**

*In der Tabelle mit der Auswertungsstatistik für metrische Variablen wird die Anzahl der fehlenden Werte angezeigt.*

	Mittelwert	Median	Fehlend
Alter der befragten Person	45,6	42,0	4
Letztes abgeschlossenes Schuljahr	13,3	13,0	12
Täglicher Fernsehkonsum in Stunden	2,9	2,0	495

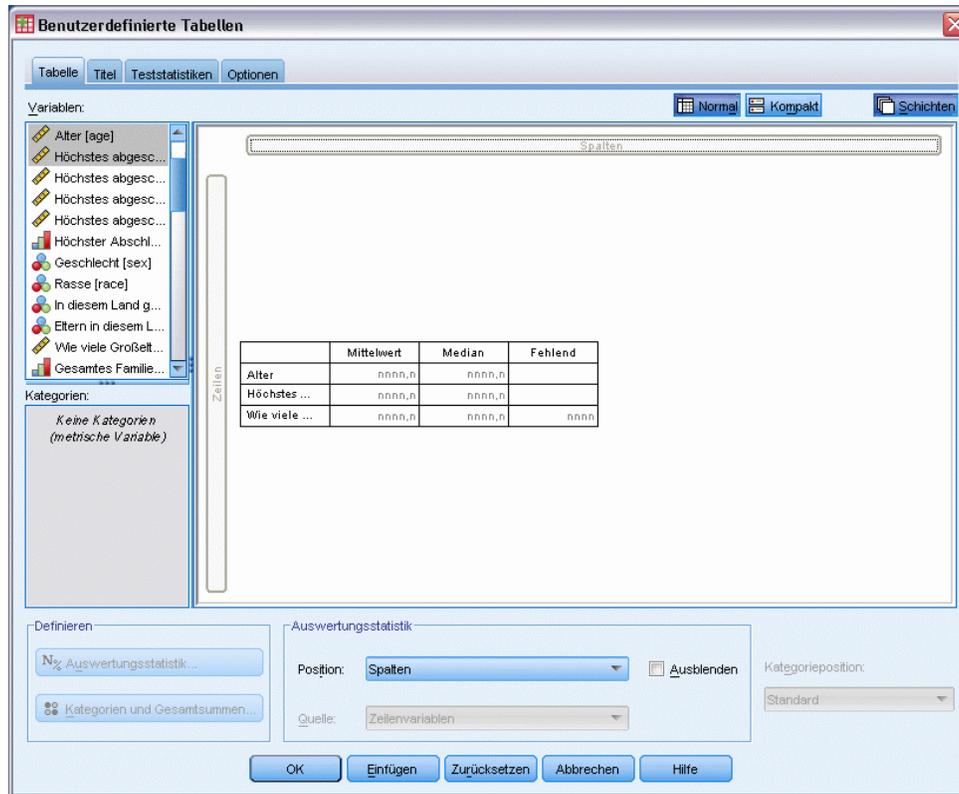
In der Tabelle wird nun für jede metrische Variable die Anzahl der fehlenden Werte angezeigt. Daraus wird leicht ersichtlich, dass bei *Täglicher Fernsehkonsum in Stunden* eine große Anzahl von fehlenden Werten vorhanden ist, während bei den anderen beiden Variablen nur sehr wenige zu verzeichnen sind. Dies sollte bei der Beurteilung der Vertrauenswürdigkeit dieser Auswertung für die Variable berücksichtigt werden.

## **Unterschiedliche Auswertungen für unterschiedliche Variablen**

In einer gestapelten Tabelle können nicht nur mehrere Auswertungsstatistiken, sondern auch unterschiedliche Auswertungsstatistiken für unterschiedliche metrische Variablen angezeigt werden. Aus der vorangegangenen Tabelle geht beispielsweise hervor, dass nur eine der drei Variablen über eine große Anzahl fehlender Werte verfügt. Möglicherweise empfiehlt es sich daher, nur für diese Variable die Anzahl der fehlenden Werte in die Tabelle aufzunehmen.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie in der Tabellenvorschau auf der Zeichenfläche auf *Alter der befragten Person* und klicken Sie anschließend mit gedrückter STRG-Taste auf *Letztes abgeschlossenes Schuljahr*, um beide Variablen auszuwählen.
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf eine der beiden Variablen und wählen Sie im Kontextmenü die Option **Auswertungsstatistik** aus.
- ▶ Wählen Sie im Dialogfeld “Auswertungsstatistik” in der Liste “Anzeigen” die Option **Fehlend** aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Statistik” zu übertragen und aus der Liste “Anzeigen” zu entfernen.
- ▶ Klicken Sie auf **Der Auswahl zuweisen**, um die Änderung ausschließlich den beiden ausgewählten Variablen zuzuweisen.

Abbildung 9-7  
Tabellenvorschau mit unterschiedlichen Auswertungsstatistiken für unterschiedliche Variablen



Die Platzhalter in den Datenzellen der Tabelle zeigen an, dass die Anzahl der fehlenden Werte nur für *Täglicher Fernsehkonsum in Stunden* angezeigt wird.

- Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 9-8  
Tabelle mit unterschiedlichen Auswertungsstatistiken für unterschiedliche Variablen

	Mittelwert	Median	Fehlend
Alter der befragten Person	45,6	42,0	
Letztes abgeschlossenes Schuljahr	13,3	13,0	
Täglicher Fernsehkonsum in Stunden	2,9	2,0	495

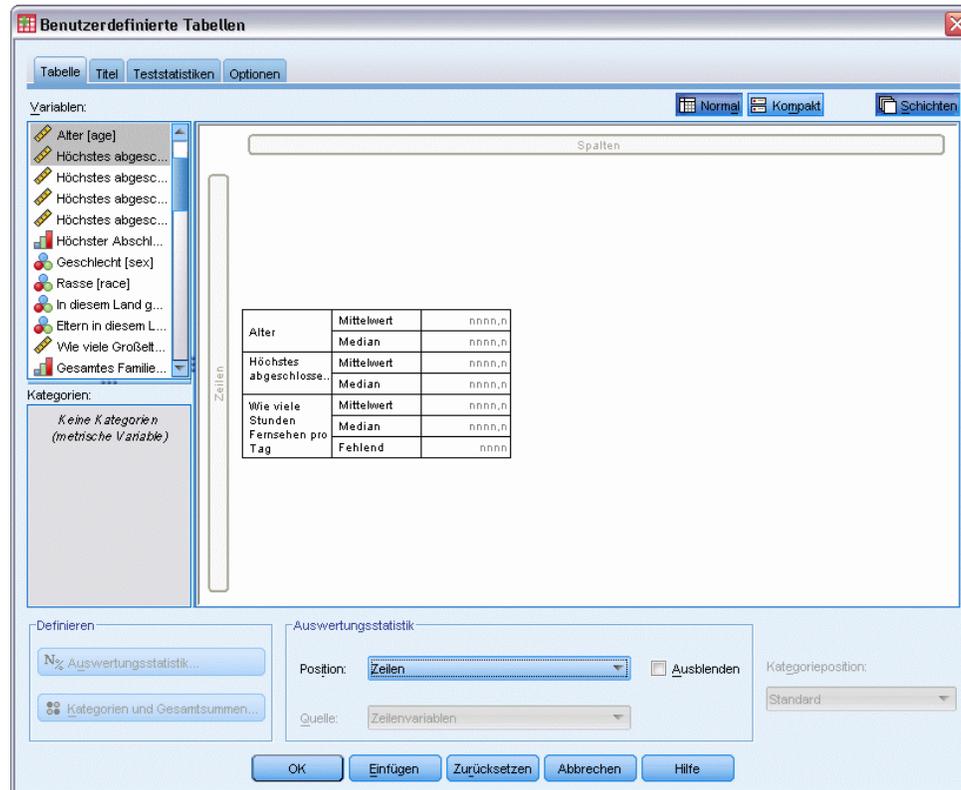
Diese Tabelle liefert zwar die gewünschten Informationen, das Layout erschwert jedoch die Interpretation der Tabelle. Beim Lesen der Tabelle könnte den Eindruck entstehen, dass die leeren Zellen in der Spalte *Fehlend* bedeuten, dass 0 fehlende Werte für die entsprechenden Variablen vorliegen.

- Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).

- Wählen Sie im Dialogfeld zum Konstruieren von Tabellen im Gruppenfeld “Auswertungsstatistik” den Eintrag Zeilen aus der Dropdown-Liste “Position “ aus.

Abbildung 9-9

Verschieben von Auswertungsstatistiken aus der Spaltendimension in die Zeilendimension



- Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 9-10

Die Auswertungsstatistiken und Variablen werden in der Zeilendimension angezeigt.

Alter der befragten Person	Mittelwert	45,6
	Median	42,0
Letztes abgeschlossenes Schuljahr	Mittelwert	13,3
	Median	13,0
Täglicher Fernsehkonsum in Stunden	Mittelwert	2,9
	Median	2,0
	Fehlend	495

Jetzt geht aus der Tabelle eindeutig hervor, dass nur für eine Variable die Anzahl der fehlenden Werte aufgeführt wird.

## Gruppierte Auswertungen in Kategorien

Kategoriale Variablen können als Gruppenvariablen verwendet werden, um Auswertungen für metrische Variablen in den durch die Kategorien der kategorialen Variablen definierten Gruppen anzuzeigen.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Ziehen Sie die Variable *Geschlecht* aus der Variablenliste, und legen Sie sie im Spaltenbereich der Zeichenfläche ab.

Wenn Sie mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf *Geschlecht* klicken, werden Sie feststellen, dass der Befehl Auswertungsstatistik im Kontextmenü deaktiviert ist. Die Ursache hierfür ist, dass in einer Tabelle mit metrischen Variablen diese immer die Quellvariablen für die Statistiken darstellen.

- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 9-11

*Gruppierete Auswertungen von metrischen Variablen anhand einer kategorialen Spaltenvariablen*

		Geschlecht	
		Männlich	Weiblich
Alter der befragten Person	Mittelwert	44,6	46,3
	Median	42,0	43,0
Letztes abgeschlossenes Schuljahr	Mittelwert	13,4	13,2
	Median	13,0	13,0
Täglicher Fernsehkonsum in Stunden	Mittelwert	2,8	2,9
	Median	2,0	2,0
	Fehlend	213	282

Anhand dieser Tabelle lassen sich die Durchschnittswerte (Mittelwert und Median) für Frauen und Männer leicht vergleichen. Sie zeigt eindeutig, dass die Unterschiede relativ gering sind. Dies ist zwar möglicherweise nicht sehr interessant, kann aber eine nützliche Information darstellen.

## **Mehrere Gruppenvariablen**

Die Gruppen können durch Verschachtelung und/oder gleichzeitiger Verwendung von kategorialen Zeilen- und Spaltenvariablen als Gruppenvariablen weiter unterteilt werden.

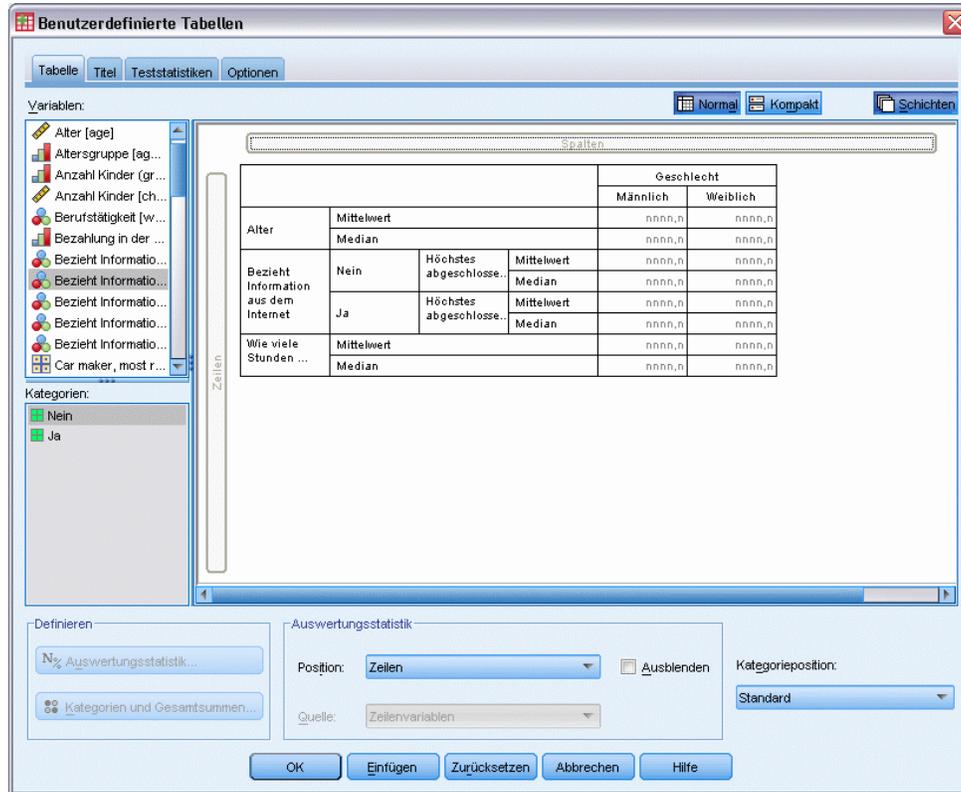
- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Ziehen Sie *Nachrichten aus dem Internet* aus der Variablenliste, und legen Sie sie im äußersten linken Rand des Zeilenbereichs der Zeichenfläche ab. Stellen Sie sicher, dass Sie die Variable so ablegen, dass wirklich alle drei metrischen Variablen darin verschachtelt sind, nicht nur eine der Variablen.

Abbildung 9-12  
 Richtig: Alle drei metrischen Variablen sind in der kategorialen Variablen verschachtelt.

The screenshot shows the 'Benutzerdefinierte Tabellen' (Custom Tables) dialog box in SPSS. The 'Variablen:' list on the left contains several variables, including 'Alter [age]', 'Altersgruppe [ag...]', 'Anzahl Kinder (gr...', 'Anzahl Kinder [ch...', 'Berufstätigkeit [w...', 'Bezahlung in der ...', 'Bezieht Informatio...', 'Bezieht Informatio...', 'Bezieht Informatio...', 'Bezieht Informatio...', 'Bezieht Informatio...', and 'Car maker, most r...'. The 'Kategorien:' list shows 'Nein' and 'Ja'. The main table area displays a nested structure with 'Bezieht Information aus dem Internet' as the primary row variable, split into 'Nein' and 'Ja'. Each category is further split into 'Alter', 'Höchstes abgeschlosse...', and 'Wie viele Stunden ...'. For each of these, 'Mittelwert' and 'Median' are calculated. The columns represent 'Geschlecht' (Männlich and Weiblich), with numerical values like 'nnnn,n' shown. The 'Auswertungsstatistik' section at the bottom is set to 'Position: Zeilen' and 'Quelle: Zeilenvariablen'. The 'Kategorieposition' is set to 'Standard'. Buttons for 'OK', 'Einfügen', 'Zurücksetzen', 'Abbrechen', and 'Hilfe' are visible at the bottom.

				Geschlecht	
				Männlich	Weiblich
Bezieht Information aus dem Internet	Nein	Alter	Mittelwert	nnnn,n	nnnn,n
			Median	nnnn,n	nnnn,n
		Höchstes abgeschlosse...	Mittelwert	nnnn,n	nnnn,n
		Median	nnnn,n	nnnn,n	
	Ja	Alter	Mittelwert	nnnn,n	nnnn,n
			Median	nnnn,n	nnnn,n
Höchstes abgeschlosse...		Mittelwert	nnnn,n	nnnn,n	
	Median	nnnn,n	nnnn,n		
	Wie viele Stunden ...	Mittelwert	nnnn,n	nnnn,n	
		Median	nnnn,n	nnnn,n	

Abbildung 9-13  
 Falsch: Nur eine metrische Variable ist in der kategorialen Variablen verschachtelt.



Es mag Situationen geben, in denen eine Verschachtelung wie im zweiten Beispiel oben erwünscht ist. Im vorliegenden Beispiel trifft dies jedoch nicht zu.

- Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 9-14  
 Auswertungen von metrischen Variablen gruppiert nach kategorialen Zeilen- und Spaltenvariablen

				Geschlecht	
				Männlich	Weiblich
Nachrichten aus dem Internet	Nein	Alter der befragten Person	Mittelwert	47,0	48,8
			Median	45,0	46,0
		Letztes abgeschlossenes Schuljahr	Mittelwert	13,4	13,1
			Median	13,0	12,0
		Täglicher Fernsehkonsum in Stunden	Mittelwert	3,2	3,4
			Median	2,0	3,0
	Ja	Alter der befragten Person	Fehlend	213	282
			Mittelwert	38,7	41,1
		Median	35,0	38,0	
		Letztes abgeschlossenes Schuljahr	Mittelwert	13,2	13,3
			Median	13,0	13,0
		Täglicher Fernsehkonsum in Stunden	Mittelwert	2,1	2,1
Median	2,0		2,0		
		Fehlend	0	0	

## Verschachteln von kategorialen in metrischen Variablen

Die vorangegangene Tabelle liefert zwar die gewünschten Informationen, bietet sie jedoch in einem Format an, das die Interpretation nicht gerade erleichtert. So kann beispielsweise anhand der Tabelle das durchschnittliche Alter der Personen verglichen werden, die Nachrichten aus dem Internet bzw. nicht aus dem Internet erfahren. Der Vergleich ließe sich jedoch leichter durchführen, wenn die Werte nebeneinander angezeigt würden. Durch Vertauschen der Positionen der beiden Zeilenvariablen und Verschachteln der kategorialen Gruppenvariablen in den drei metrischen Variablen kann die Tabellendarstellung verbessert werden. Bei metrischen Variablen hat die Verschachtelungsebene keinen Einfluß auf die Quellvariable der Statistik. Die metrische Variable ist unabhängig von der Verschachtelungsebene immer die Quellvariable für die Statistik.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie in der Tabellenvorschau auf der Zeichenfläche auf *Alter der befragten Person*, klicken Sie anschließend mit gedrückter STRG-Taste auf *Letztes abgeschlossenes Schuljahr* und klicken Sie dann mit gedrückter STRG-Taste auf *Täglicher Fernsehkonsum in Stunden*, um alle drei metrischen Variablen auszuwählen.
- ▶ Ziehen Sie die drei metrischen Variablen, und legen Sie sie so am äußersten linken Rand des Zeilenbereichs ab, dass die kategoriale Variable *Nachrichten aus dem Internet* in jeder der drei metrischen Variablen verschachtelt ist.
- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 9-15

Die kategoriale Zeilenvariable ist in gestapelten metrischen Variablen verschachtelt.

				Geschlecht	
				Männlich	Weiblich
Alter der befragten Person	Nachrichten aus dem Internet	Nein	Mittelwert	47,0	48,8
			Median	45,0	46,0
	Ja	Mittelwert	38,7	41,1	
		Median	35,0	38,0	
Letztes abgeschlossenes Schuljahr	Nachrichten aus dem Internet	Nein	Mittelwert	13,4	13,1
			Median	13,0	12,0
	Ja	Mittelwert	13,2	13,3	
		Median	13,0	13,0	
Täglicher Fernsehkonsum in Stunden	Nachrichten aus dem Internet	Nein	Mittelwert	3,2	3,4
			Median	2,0	3,0
		Fehlend	213	282	
	Ja	Mittelwert	2,1	2,1	
		Median	2,0	2,0	
		Fehlend	0	0	

Die Wahl der Verschachtelungsreihenfolge hängt von den Beziehungen oder Vergleichen ab, die in der Tabelle hervorgehoben werden sollen. Ein Ändern der Verschachtelungsreihenfolge der metrischen Variablen zieht keine Änderung der Werte in den Auswertungsstatistiken nach sich. Lediglich die relative Position in der Tabelle wird geändert.

# Teststatistiken

Es stehen drei verschiedene Tests auf Signifikanz zum Untersuchen der Beziehung zwischen Zeilen- und Spaltenvariablen zur Verfügung. In diesem Kapitel wird die Ausgabe dieser Tests erläutert, wobei besonderes Augenmerk auf die Auswirkungen von verschachtelten und gestapelten Variablen gerichtet wird. [Für weitere Informationen siehe Thema Stapeln, Verschachteln und Schichten von kategorialen Variablen in Kapitel 4 auf S. 70.](#)

## **Beispiel-Datendatei**

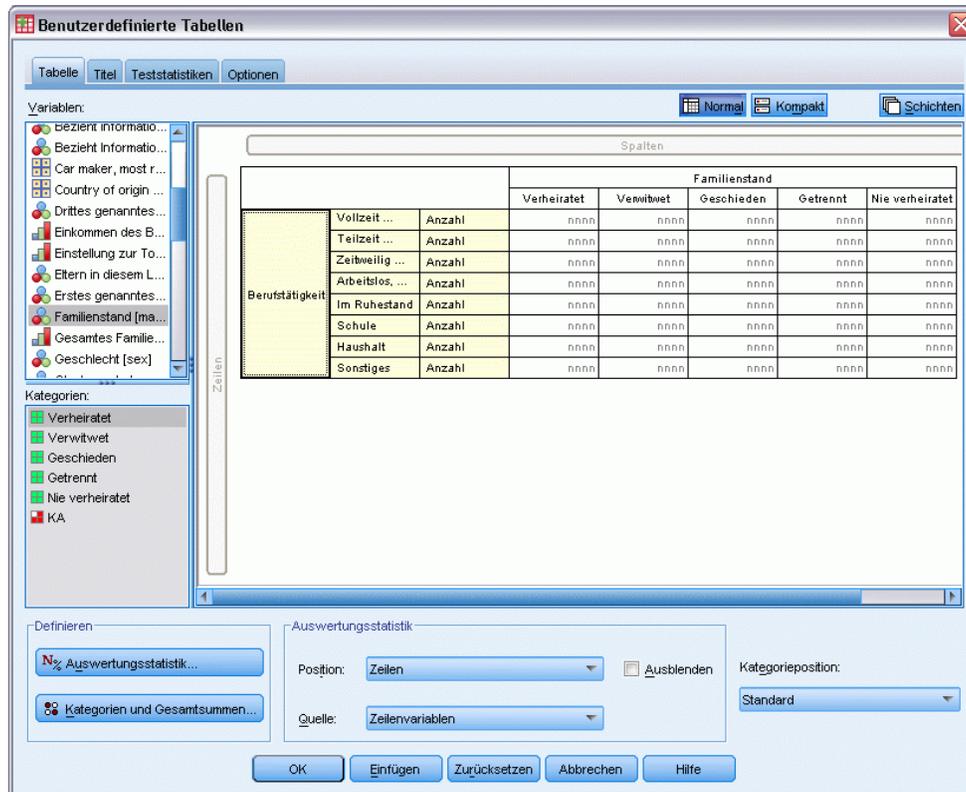
Für die Beispiele dieses Kapitels wird die Datendatei *survey\_sample.sav* verwendet. [Für weitere Informationen siehe Thema Beispieldateien in Anhang A auf S. 201.](#)

## **Tests auf Unabhängigkeit (Chi-Quadrat)**

Mit dem Chi-Quadrat-Test auf Unabhängigkeit kann ermittelt werden, ob eine Beziehung zwischen zwei kategorialen Variablen vorliegt. So kann beispielsweise ermittelt werden, ob es einen Zusammenhang zwischen *Beschäftigungsstatus* und *Familienstand* gibt.

- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
Analysieren > Tabellen > Benutzerdefinierte Tabellen...
- ▶ Ziehen Sie die Variable *Beschäftigungsstatus* aus der Variablenliste der Tabellenkonstruktion in den Bereich für die Zeilen auf der Zeichenfläche.
- ▶ Ziehen Sie die Variable *Familienstand* aus der Variablenliste, und legen Sie sie im Spaltenbereich ab.

Abbildung 10-1  
Variablen auf der Zeichenfläche



- ▶ Wählen Sie Zeilen als Position für die Auswertungsstatistiken.
- ▶ Wählen Sie *Beschäftigungsstatus* aus, und klicken Sie im Gruppenfeld "Definieren" auf Auswertungsstatistik.

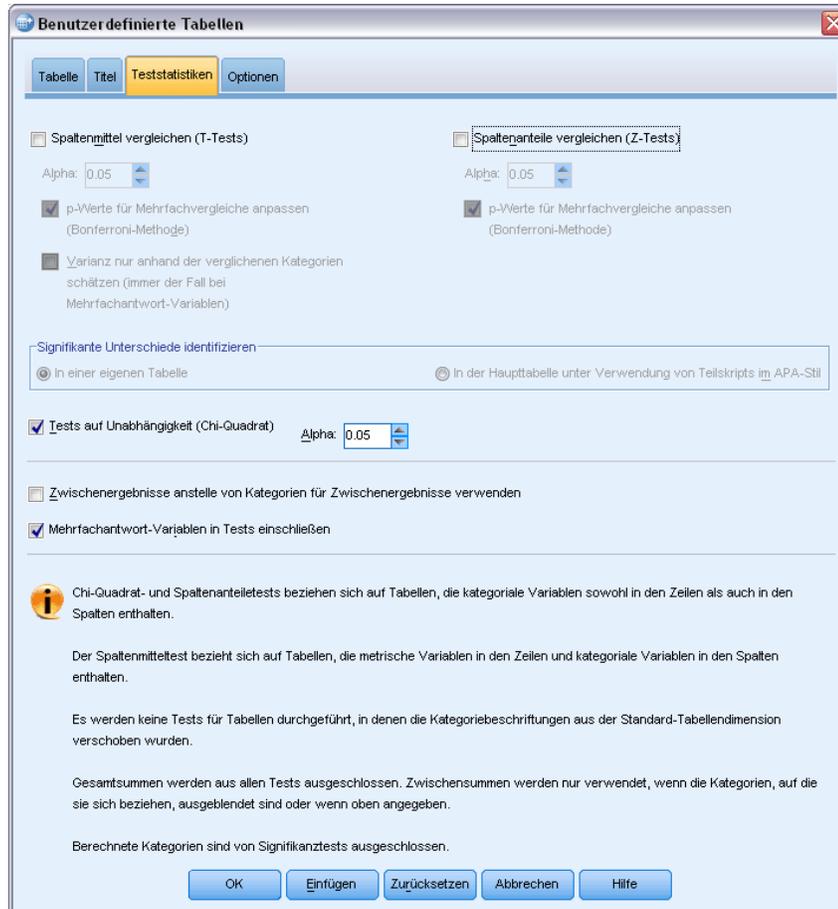
Abbildung 10-2  
Dialogfeld "Auswertungsstatistik"



- ▶ Wählen Sie in der Liste “Statistik” die Option Anzahl als Spalten% aus, und übertragen Sie sie in die Liste “Anzeigen”.
- ▶ Klicken Sie auf Der Auswahl zuweisen.
- ▶ Klicken Sie im Dialogfeld “Benutzerdefinierte Tabellen” auf die Registerkarte Teststatistiken.

Abbildung 10-3

Auf der Registerkarte “Teststatistiken” ist die Option “Tests auf Unabhängigkeit (Chi-Quadrat)” ausgewählt.



- ▶ Wählen Sie Tests auf Unabhängigkeit (Chi-Quadrat) aus.
- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen und den Chi-Quadrat-Test durchzuführen.

Abbildung 10-4  
Beschäftigungsstatus nach Familienstand

			Familienstand				
			Verheiratet	Verwitwet	Geschieden	Getrennt lebend	Ledig
Beschäftigungsstatus	Vollzeit	Anzahl	778	44	295	58	392
		Spalten%	57,8%	15,5%	66,1%	62,4%	59,1%
	Teilzeit	Anzahl	138	20	35	9	102
		Spalten%	10,3%	7,1%	7,8%	9,7%	15,4%
	Zeitweise unbeschäftigt	Anzahl	23	2	9	1	11
		Spalten%	1,7%	,7%	2,0%	1,1%	1,7%
	Arbeitslos, entlassen	Anzahl	13	3	10	0	32
		Spalten%	1,0%	1,1%	2,2%	,0%	4,8%
	Ruhestand	Anzahl	168	150	53	6	17
		Spalten%	12,5%	53,0%	11,9%	6,5%	2,8%
	Ausbildung	Anzahl	9	1	7	2	60
		Spalten%	,7%	,4%	1,6%	2,2%	9,0%
	Arbeit im Haushalt	Anzahl	200	55	25	13	35
		Spalten%	14,9%	19,4%	5,6%	14,0%	5,3%
	Andere	Anzahl	16	8	12	4	14
		Spalten%	1,2%	2,8%	2,7%	4,3%	2,1%

Bei dieser Tabelle handelt es sich um eine Kreuztabelle von *Beschäftigungsstatus* nach *Familienstand*, in der die Häufigkeiten und Spaltenanteile als Auswertungsstatistiken angezeigt werden. Die Spaltenanteile werden so berechnet, daß die Summe jeder Spalte 100 % ergibt. Wenn zwischen den beiden Variablen kein Zusammenhang besteht, sollten sich in jeder Zeile ähnliche Anteile in allen Spalten befinden. Bei den Anteilen scheint es Differenzen zu geben. Sie können jedoch den Chi-Quadrat-Test überprüfen, um sich zu vergewissern.

Abbildung 10-5  
Chi-Quadrat-Test nach Pearson

		Familiens tand
Beschäftigungsstatus	Chi-Quadrat	729,242
	df	28
	Sig.	,000*

\*. Die Chi-Quadrat-Statistik ist auf dem Niveau 0.05 signifikant.

Beim Test auf Unabhängigkeit wird davon ausgegangen, daß zwischen *Beschäftigungsstatus* und *Familienstand* keine Beziehung besteht, also die Spaltenanteile in allen Spalten gleich sind und beobachtete Abweichungen auf zufällige Schwankungen zurückzuführen sind. Bei der Chi-Quadrat-Statistik wird die Gesamtabweichung zwischen den beobachteten Zellenhäufigkeiten und den bei Gleichheit der Spaltenanteile über alle Spalten erwarteten Häufigkeiten gemessen. Eine größere Chi-Quadrat-Statistik zeigt eine größere Abweichung zwischen den beobachteten und den erwarteten Zellenhäufigkeiten an. Dies deutet darauf hin, daß die Spaltenanteile nicht gleich sind, daß die Hypothese der Unabhängigkeit somit falsch ist und daß deshalb ein Zusammenhang zwischen *Beschäftigungsstatus* und *Familienstand* besteht.

Die berechnete Chi-Quadrat-Statistik weist den Wert 729,242 auf. Um festzustellen, ob dies ausreicht, um die Hypothese der Unabhängigkeit zu widerlegen, wird der Signifikanzwert der Statistik berechnet. Der Signifikanzwert ist die Wahrscheinlichkeit, daß eine Zufallsvariante aus einer Chi-Quadrat-Verteilung mit 28 Freiheitsgraden größer als 729,242 ist. Da dieser Wert kleiner als das auf der Registerkarte "Teststatistiken" angegebene Alpha-Niveau ist, kann die Hypothese der Unabhängigkeit auf dem Niveau 0,05 widerlegt werden. Somit besteht tatsächlich ein Zusammenhang zwischen *Beschäftigungsstatus* und *Familienstand*.

## Auswirkungen von verschachtelten und gestapelten Variablen auf Tests auf Unabhängigkeit

Die Regel für Tests auf Unabhängigkeit lautet folgendermaßen: für jede innerste Untertabelle wird ein separater Test durchgeführt. Um die Auswirkung von verschachtelten Variablen auf die Tests zu verdeutlichen, betrachten Sie das vorhergehende Beispiel, wobei jedoch *Familienstand* in Ebenen von *Geschlecht* verschachtelt wird.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen erneut (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Ziehen Sie die Variable *Geschlecht* aus der Variablenliste auf die Zeichenfläche, und legen Sie sie im Spaltenbereich über *Familienstand* ab.
- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 10-6  
Chi-Quadrat-Test nach Pearson

		Geschlecht	
		Männlich	Weiblich
		Familiens- stand	Familiens- stand
Beschäftigungsstatus	Chi-Quadrat	246,637	542,589
	df	28	28
	Sig.	,000 <sup>a,b</sup>	,000 <sup>a,b</sup>

\*. Die Chi-Quadrat-Statistik ist auf dem Niveau 0.05 signifikant.

a. In dieser Untertabelle weisen mehr als 20% der Zellen erwartete Zellenhäufigkeiten von weniger als 5 auf. Daher sind die Ergebnisse von Chi-Quadrat möglicherweise ungültig.

b. In dieser Untertabelle ist die kleinste erwartete Zellenhäufigkeit kleiner als 1. Daher sind die Ergebnisse von Chi-Quadrat möglicherweise ungültig.

Wenn *Familienstand* in Ebenen von *Geschlecht* verschachtelt wird, werden zwei Tests durchgeführt: einer für jede Ebene von *Geschlecht*. Der Signifikanzwert für die einzelnen Tests zeigt an, daß die Hypothese der Unabhängigkeit zwischen *Familienstand* und *Beschäftigungsstatus* sowohl für Männer als auch für Frauen widerlegt werden kann. Den Anmerkungen der Tabelle kann jedoch entnommen werden, dass über 20 % der Zellen jeder Tabelle erwartete Häufigkeiten von weniger als 5 aufweisen und die niedrigste erwartete Zellenhäufigkeit unter 1 liegt. Die Anmerkungen weisen darauf hin, dass die Annahmen des Chi-Quadrat-Tests möglicherweise nicht auf diese Tabellen zutreffen und die Ergebnisse der Tests somit nicht zuverlässig sind.

*Hinweis:* Die Fußnoten sind möglicherweise aufgrund der Zellenbegrenzungen nicht sichtbar. Sie können die Fußnoten anzeigen, indem Sie die Ausrichtung der Zellen im Dialogfeld “Zelleneigenschaften” ändern.

Gehen Sie folgendermaßen vor, um die Auswirkung von gestapelten Variablen auf die Tests zu verdeutlichen:

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen erneut (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).

- ▶ Ziehen Sie die Variable *Höchster Schulabschluß* aus der Variablenliste, und legen Sie sie im Zeilenbereich unter *Beschäftigungsstatus* ab.
- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 10-7  
Chi-Quadrat-Test nach Pearson

		Geschlecht	
		Männlich	Weiblich
		Familiens- stand	Familiens- stand
Beschäftigungsstatus	Chi-Quadrat	246,637	542,589
	df	28	28
	Sig.	,000 <sup>a,b</sup>	,000 <sup>a,b</sup>
Höchster Schulabschluß	Chi-Quadrat	43,844	105,506
	df	16	16
	Sig.	,000 <sup>a</sup>	,000 <sup>a</sup>

\*. Die Chi-Quadrat-Statistik ist auf dem Niveau 0.05 signifikant.

- a. In dieser Untertabelle weisen mehr als 20% der Zellen erwartete Zellenhäufigkeiten von weniger als 5 auf. Daher sind die Ergebnisse von Chi-Quadrat möglicherweise ungültig.
- b. In dieser Untertabelle ist die kleinste erwartete Zellenhäufigkeit kleiner als 1. Daher sind die Ergebnisse von Chi-Quadrat möglicherweise ungültig.

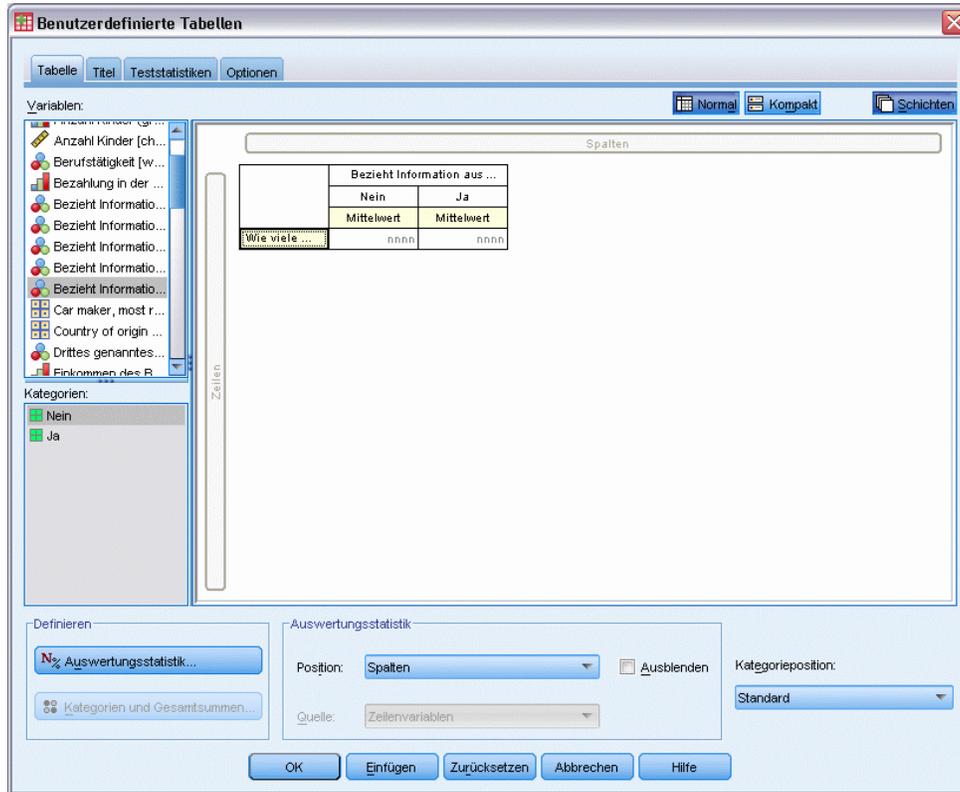
Wenn *Höchster Schulabschluß* mit *Beschäftigungsstatus* gestapelt wird, werden vier Tests durchgeführt: für jede Ebene von *Geschlecht* ein Test auf Unabhängigkeit von *Familienstand* und *Beschäftigungsstatus* und ein Test auf Unabhängigkeit von *Familienstand* und *Höchster Schulabschluß*. Die Testergebnisse für *Familienstand* und *Beschäftigungsstatus* sind die gleichen wie vorher. Die Testergebnisse für *Familienstand* und *Höchster Schulabschluß* weisen darauf hin, daß diese Variablen nicht unabhängig sind.

## Vergleichen der Spaltenmittel

Mit den Spaltenmitteltests wird ermittelt, ob eine Beziehung zwischen einer kategorialen Variablen in den Spalten und einer stetigen Variablen in den Zeilen vorliegt. Außerdem können Sie mit den Testergebnissen die relative Ordnung der Kategorien der kategorialen Variablen hinsichtlich des Mittelwerts der stetigen Variablen ermitteln. So können Sie beispielsweise ermitteln, ob zwischen *Täglicher Fernsehkonsum in Stunden* und *Nachrichten aus Zeitungen* ein Zusammenhang besteht.

- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
Analysieren > Tabellen > Benutzerdefinierte Tabellen...
- ▶ Klicken Sie auf Zurücksetzen, um die Standardeinstellungen aller Registerkarten wiederherzustellen.
- ▶ Ziehen Sie im Dialogfeld zur Tabellenkonstruktion die Variable *Täglicher Fernsehkonsum in Stunden* aus der Variablenliste, und legen Sie sie im Zeilenbereich der Zeichenfläche ab.
- ▶ Ziehen Sie die Variable *Nachrichten aus Zeitungen* aus der Variablenliste, und legen Sie sie im Spaltenbereich ab.

Abbildung 10-8  
Variablen auf der Zeichenfläche



- Wählen Sie *Täglicher Fernsehkonsum in Stunden* aus, und klicken Sie im Gruppenfeld “Definieren” auf Auswertungsstatistik.

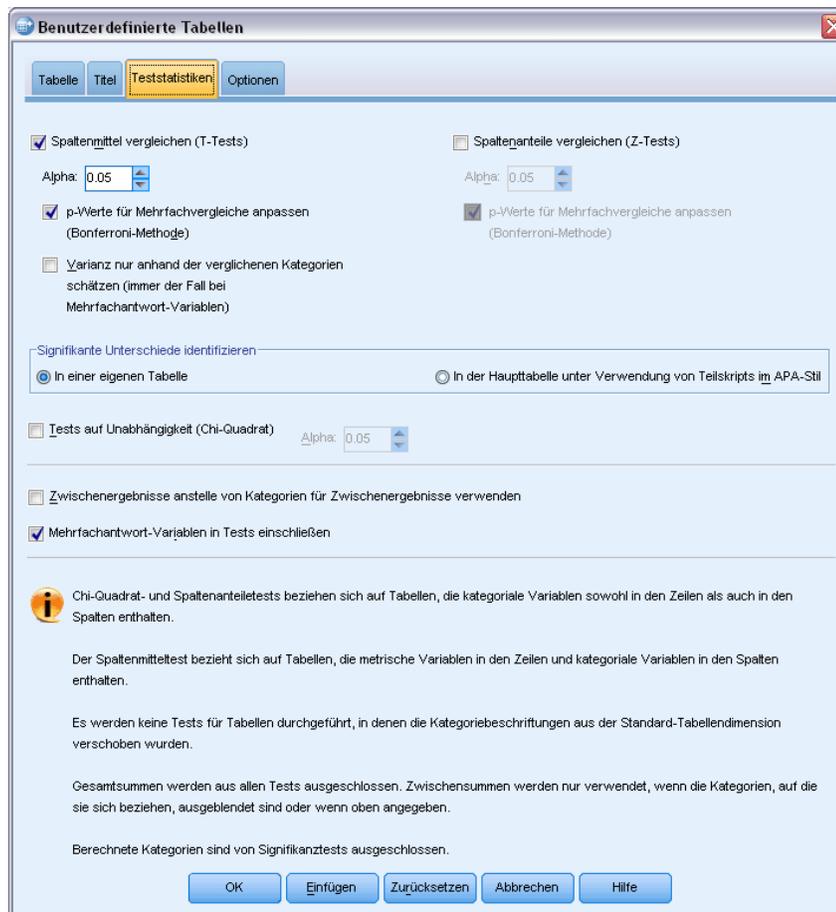
Abbildung 10-9  
Dialogfeld “Auswertungsstatistik”



- Wählen Sie nnnn als Format aus.
- Wählen Sie 2 als Anzahl der angezeigten Dezimalstellen aus. Beachten Sie, daß das Format nun als nnnn,nn dargestellt wird.
- Klicken Sie auf Der Auswahl zuweisen.
- Klicken Sie im Dialogfeld “Benutzerdefinierte Tabellen” auf die Registerkarte Teststatistiken.

Abbildung 10-10

Auf der Registerkarte "Teststatistiken" ist die Option "Spaltenmittel vergleichen (T-Tests)" ausgewählt.



- ▶ Wählen Sie Spaltenmittel vergleichen (T-Tests) aus.
- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen und die Spaltenmitteltests durchzuführen.

Abbildung 10-11

Nachrichten aus Zeitungen nach täglichem Fernsehkonsum in Stunden

	Bezieht Information aus Zeitungen	
	Nein	Ja
	Mittelwert	Mittelwert
Wie viele Stunden Fernsehen pro Tag	2,92	2,74

Diese Tabelle zeigt den Mittelwert von *Täglicher Fernsehkonsum in Stunden* für Personen, die Nachrichten aus Zeitungen bzw. nicht aus Zeitungen erfahren. Die beobachtete Differenz in diesen Mittelwerten deutet darauf hin, daß Personen, die Nachrichten nicht aus Zeitungen erfahren, zirka 0,18 Stunden länger fernsehen als Personen, die Nachrichten aus Zeitungen erfahren. Um herauszufinden, ob diese Differenz auf zufällige Schwankungen zurückzuführen ist, überprüfen Sie die Spaltenmitteltests.

Abbildung 10-12  
Vergleich der Spaltenmittel

	Nachrichten aus Zeitungen	
	Nein	Ja
	(A)	(B)
Täglicher Fernsehkonsum in Stunden		

In der Tabelle für den Spaltenmitteltest wird jeder Kategorie der Spaltenvariablen ein Buchstabenschlüssel zugewiesen. Bei *Nachrichten aus Zeitungen* wird der Kategorie *Nein* der Buchstabe A zugewiesen, und *Ja* erhält den Buchstaben B. Für jedes Spaltenpaar werden die Spaltenmittel mit Hilfe des *T*-Tests verglichen. Da nur zwei Spalten vorliegen, wird nur ein Test durchgeführt. Für jedes signifikante Paar wird der Schlüssel der Kategorie mit dem kleineren Mittelwert unterhalb der Kategorie mit dem größeren Mittelwert ausgegeben. Da in den Zellen der Tabelle keine Schlüssel ausgegeben werden, bedeutet dies, daß die Spaltenmittel statistisch gesehen nicht differieren.

#### **Signifikanzergebnisse durch Notation im APA-Stil**

Wenn Sie nicht möchten, dass die Signifikanzergebnisse in einer separaten Tabelle angezeigt werden, können Sie diese auch in der Haupttabelle anzeigen. Signifikanzergebnisse werden anhand einer Notation im APA-Stil mit tiefgestellten Buchstaben identifiziert. Führen Sie die obigen Schritte zum Vergleich von Spaltenmitteln aus, nehmen Sie jedoch auf der Registerkarte "Teststatistiken" folgende Änderung vor:

- ▶ Wählen Sie im Bereich "Signifikante Unterschiede identifizieren" die Option In der Haupttabelle unter Verwendung von Teilskripts im APA-Stil aus.

Abbildung 10-13

Registerkarte "Teststatistiken" mit der Auswahl "Spaltenmittel vergleichen (T-Tests)" und "In der Haupttabelle, anhand von tiefgestellten Buchstaben"



- Klicken Sie auf OK, um die Tabelle zu erstellen und die Spaltenmitteltests anhand von Notation im APA-Stil durchzuführen.

Abbildung 10-14

Vergleiche von Spaltenmitteln anhand von Notation im APA-Stil

	Bezieht Information aus Zeitungen	
	Nein	Ja
	Mittelwert	Mittelwert
Wie viele Stunden Fernsehen pro Tag	2.92 <sub>a</sub>	2.74 <sub>a</sub>

In der Tabelle für den Spaltenmitteltest wird den Kategorien der Spaltenvariablen ein tiefgestellter Buchstabe zugewiesen. Die Spaltenmittel werden für jedes Spaltenpaar mithilfe eines *T*-Tests verglichen. Wenn sich ein Wertepaar stark unterscheidet, werden den Werten *unterschiedliche* tiefgestellte Buchstaben zugewiesen. Da nur zwei Spalten vorliegen, wird nur ein Test durchgeführt. Da die Spaltenmittel in diesem Beispiel über den gleichen tiefgestellten Buchstaben verfügen, unterscheiden sich die Spaltenmittel statistisch gesehen nicht.

### **Auswirkungen von verschachtelten und gestapelten Variablen auf Spaltenmitteltests**

Die Regel für Spaltenmitteltests lautet folgendermaßen: für jede innerste Untertabelle wird eine separate Reihe paarweiser Tests durchgeführt. Um die Auswirkung von verschachtelten Variablen auf die Tests zu verdeutlichen, betrachten Sie das vorhergehende Beispiel, wobei jedoch *Täglicher Fernsehkonsum in Stunden* in Ebenen von *Beschäftigungsstatus* verschachtelt wird.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen erneut (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Ziehen Sie die Variable *Beschäftigungsstatus* aus der Variablenliste, und legen Sie sie im Zeilenbereich auf der Zeichenfläche ab.
- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 10-15  
Vergleich der Spaltenmittel

			Nachrichten aus Zeitungen	
			Nein	Ja
			(A)	(B)
Beschäftigungsstatus	Vollzeit	Täglicher Fernsehkonsum in Stunden	B	
	Teilzeit	Täglicher Fernsehkonsum in Stunden		
	Zeitweise unbeschäftigt	Täglicher Fernsehkonsum in Stunden		
	Arbeitslos, entlassen	Täglicher Fernsehkonsum in Stunden		
	Ruhestand	Täglicher Fernsehkonsum in Stunden		
	Ausbildung	Täglicher Fernsehkonsum in Stunden		
	Arbeit im Haushalt	Täglicher Fernsehkonsum in Stunden		
	Andere	Täglicher Fernsehkonsum in Stunden		

Wenn *Täglicher Fernsehkonsum in Stunden* in Ebenen von *Beschäftigungsstatus* verschachtelt ist, werden sieben Reihen von Spaltenmitteltests durchgeführt: eine für jede Ebene von *Beschäftigungsstatus*. Den Kategorien von *Nachrichten aus Zeitungen* werden die gleichen Buchstabenschlüssel zugewiesen. Für Befragte mit Beschäftigung in *Vollzeit* wird der Schlüssel B in der Spalte A angezeigt. Das bedeutet, daß bei Vollzeitbeschäftigten der Mittelwert von *Täglicher Fernsehkonsum in Stunden* bei den Personen niedriger ist, die Nachrichten aus Zeitungen erfahren. Da in den Spalten keine anderen Schlüssel angezeigt werden, können Sie daraus schließen, daß bei den Spaltenmitteln keine weiteren statistisch signifikanten Differenzen vorliegen.

**Bonferroni-Korrekturen.** Wenn mehrere Tests durchgeführt werden, wird die Bonferroni-Korrektur auf die Spaltenmitteltests angewendet, um sicherzustellen, daß das auf der Registerkarte “Teststatistiken” angegebene Alpha-Niveau (Rate der fälschlicherweise positiven Ergebnisse) für jede *Testreihe* gilt. In dieser Tabelle wurden keine Bonferroni-Korrekturen angewendet, denn obwohl sieben Testreihen durchgeführt wurden, wurde in jeder Reihe nur ein Spaltenpaar verglichen.

Gehen Sie folgendermaßen vor, um die Auswirkung von gestapelten Variablen auf die Tests zu verdeutlichen:

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen erneut (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Ziehen Sie die Variable *Nachrichten aus dem Internet* aus der Variablenliste, und legen Sie sie im Spaltenbereich links neben *Nachrichten aus Zeitungen* ab.
- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 10-16  
Vergleich der Spaltenmittel

			Nachrichten aus dem Internet		Nachrichten aus Zeitungen	
			Nein	Ja	Nein	Ja
			(A)	(B)	(A)	(B)
Beschäftigungsstatus	Vollzeit	Täglicher Fernsehkonsum in Stunden	B		B	
	Teilzeit	Täglicher Fernsehkonsum in Stunden	B			
	Zeitweise unbeschäftigt	Täglicher Fernsehkonsum in Stunden				
	Arbeitslos, entlassen	Täglicher Fernsehkonsum in Stunden	B			
	Ruhestand	Täglicher Fernsehkonsum in Stunden	B			
	Ausbildung	Täglicher Fernsehkonsum in Stunden	B			
	Arbeit im Haushalt	Täglicher Fernsehkonsum in Stunden	B			
	Andere	Täglicher Fernsehkonsum in Stunden	B			

Wenn *Nachrichten aus dem Internet* mit *Nachrichten aus Zeitungen* gestapelt ist, werden 14 Reihen von Spaltenmitteltests durchgeführt: eine für jede Ebene von *Beschäftigungsstatus* sowohl für *Nachrichten aus dem Internet* als auch für *Nachrichten aus Zeitungen*. Wiederum werden keine Bonferroni-Korrekturen vorgenommen, da in jeder Reihe nur ein Spaltenpaar verglichen wird. Die Tests für *Nachrichten aus Zeitungen* sind die gleichen wie vorher. Bei *Nachrichten aus dem Internet* wird der Kategorie *Nein* der Buchstabe A zugewiesen, und *Ja* erhält den Buchstaben B. Außer für die Befragten, die derzeit “zeitweise unbeschäftigt” sind, wird in jeder Reihe von Spaltenmitteltests der Schlüssel B in der Spalte A ausgegeben. Dies bedeutet, daß der Mittelwert

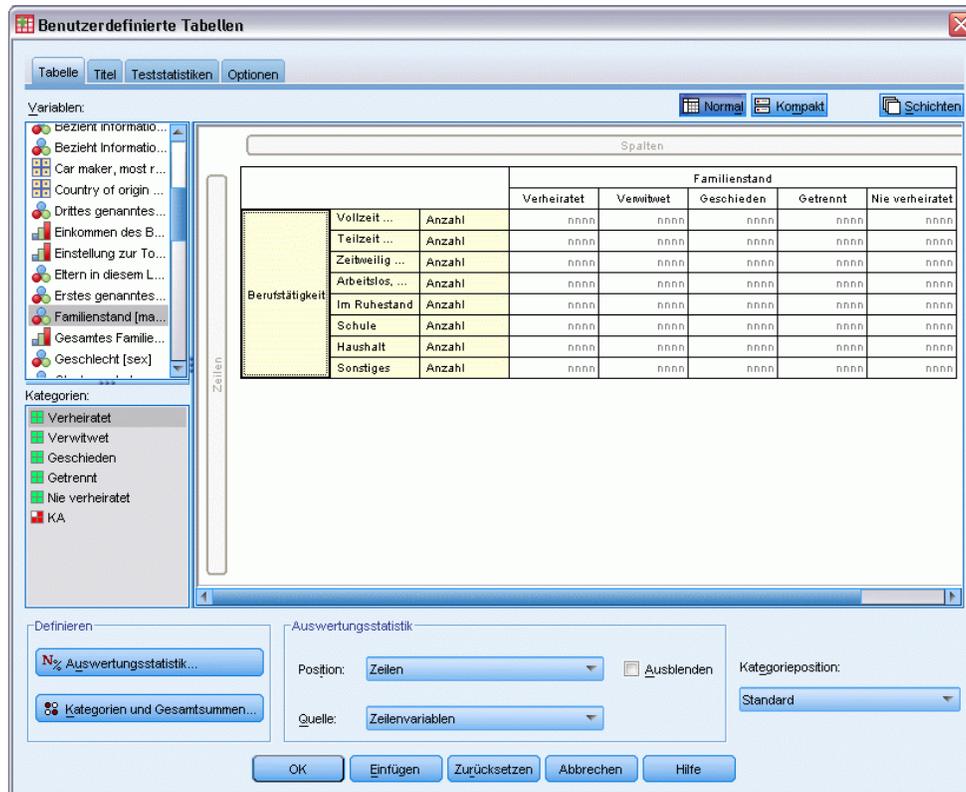
von *Täglicher Fernsehkonsum in Stunden* bei Personen, die Nachrichten aus dem Internet erfahren, niedriger ist als bei Personen, die Nachrichten nicht aus der Zeitung erfahren. Für die Testreihe für *Zeitweise unbeschäftigt* werden keine Schlüssel ausgegeben; somit differieren die Spaltenmittel für diese Befragten statistisch gesehen nicht.

## **Vergleichen der Spaltenanteile**

Mit den Spaltenanteiletests wird die relative Ordnung der Kategorien der kategorialen Spaltenvariablen hinsichtlich der Kategorienanteile der kategorialen Zeilenvariablen ermittelt. Nachdem Sie beispielsweise mit einem Chi-Quadrat-Test ermittelt haben, daß *Beschäftigungsstatus* und *Familienstand* nicht unabhängig sind, möchten Sie möglicherweise herausfinden, welche Zeilen und Spalten für diese Beziehung verantwortlich sind.

- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
Analysieren > Tabellen > Benutzerdefinierte Tabellen...
- ▶ Klicken Sie auf Zurücksetzen, um die Standardeinstellungen aller Registerkarten wiederherzustellen.
- ▶ Ziehen Sie die Variable *Beschäftigungsstatus* aus der Variablenliste der Tabellenkonstruktion in den Bereich für die Zeilen auf der Zeichenfläche.
- ▶ Ziehen Sie die Variable *Familienstand* aus der Variablenliste, und legen Sie sie im Spaltenbereich ab.

Abbildung 10-17  
Variablen auf der Zeichenfläche



- Wählen Sie *Beschäftigungsstatus* aus, und klicken Sie im Gruppenfeld “Definieren” auf Auswertungsstatistik.

Abbildung 10-18  
Dialogfeld “Auswertungsstatistik”

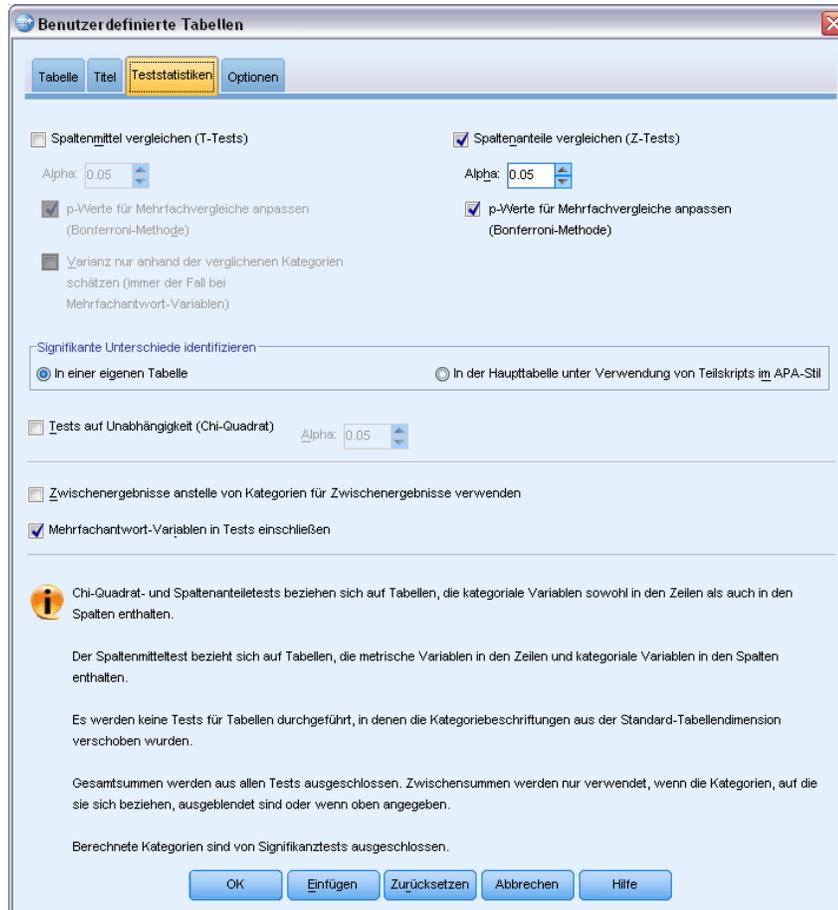


- Wählen Sie in der Liste “Statistik” die Option Anzahl als Spalten% aus, und übertragen Sie sie in die Liste “Anzeigen”.

- ▶ Deaktivieren Sie in der Liste “Anzeigen” die Option Anzahl.
- ▶ Klicken Sie auf Der Auswahl zuweisen.
- ▶ Klicken Sie im Dialogfeld “Benutzerdefinierte Tabellen” auf die Registerkarte Teststatistiken.

Abbildung 10-19

Auf der Registerkarte “Teststatistiken” ist die Option “Spaltenanteile vergleichen (Z-Tests)” ausgewählt.



- ▶ Wählen Sie Spaltenanteile vergleichen (Z-Tests) aus.
- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen und die Spaltenanteiletests durchzuführen.

Abbildung 10-20  
Beschäftigungsstatus nach Familienstand

		Familienstand				
		Verheiratet	Verwitwet	Geschieden	Getrennt lebend	Ledig
		Spalten%	Spalten%	Spalten%	Spalten%	Spalten%
Beschäftigungsstatus	Vollzeit	57,8%	15,5%	66,1%	62,4%	59,1%
	Teilzeit	10,3%	7,1%	7,8%	9,7%	15,4%
	Zeitweise unbeschäftigt	1,7%	,7%	2,0%	1,1%	1,7%
	Arbeitslos, entlassen	1,0%	1,1%	2,2%	,0%	4,8%
	Ruhestand	12,5%	53,0%	11,9%	6,5%	2,6%
	Ausbildung	,7%	,4%	1,6%	2,2%	9,0%
	Arbeit im Haushalt	14,9%	19,4%	5,6%	14,0%	5,3%
	Andere	1,2%	2,8%	2,7%	4,3%	2,1%

Bei dieser Tabelle handelt es sich um eine Kreuztabelle von *Beschäftigungsstatus* nach *Familienstand*, in der die Spaltenanteile als Auswertungsstatistik angezeigt werden.

Abbildung 10-21  
Vergleich der Spaltenanteile

		Familienstand				
		Verheiratet	Verwitwet	Geschieden	Getrennt lebend	Ledig
		(A)	(B)	(C)	(D)	(E)
Beschäftigungsstatus	Vollzeit	B		A B	B	B
	Teilzeit					A B C
	Zeitweise unbeschäftigt				<sup>a</sup>	
	Arbeitslos, entlassen					A B
	Ruhestand	E	A C D E	E		
	Ausbildung					A B C
	Arbeit im Haushalt	C E	C E		C E	
	Andere					

In der Tabelle für den Spaltenanteiletest wird jeder Kategorie der Spaltenvariablen ein Buchstabenschlüssel zugewiesen. Bei *Familienstand* wird der Kategorie *verheiratet* der Buchstabe A zugewiesen, *verwitwet* wird der Buchstabe B zugewiesen usw., bis zur Kategorie *ledig*, der der Buchstabe E zugewiesen wird. Für jedes Spaltenpaar werden die Spaltenanteile mit einem Z-Test verglichen. Es werden sieben Reihen von Spaltenanteiletests durchgeführt, eine für jede Ebene von *Beschäftigungsstatus*. Da bei *Familienstand* fünf Ebenen vorliegen, werden in jeder Testreihe  $(5 \cdot 4) / 2 = 10$  Spaltenpaare verglichen, und zum Korrigieren der Signifikanzwerte werden Bonferroni-Korrekturen verwendet. Für jedes signifikante Paar wird der Schlüssel der kleineren Kategorie unterhalb der Kategorie mit dem größeren Anteil ausgegeben.

Für die Testreihen mit *Vollzeit* wird der Schlüssel B in jeder der anderen Spalten angezeigt. Außerdem wird der Schlüssel A in der Spalte C angezeigt. In anderen Spalten werden keine weiteren Schlüssel gezeigt. Somit können Sie schließen, daß der Anteil der geschiedenen Personen, die einer Vollzeitbeschäftigung nachgehen, größer ist als der Anteil der verheirateten Personen, die Vollzeit arbeiten. Dieser Anteil ist wiederum größer als der Anteil der verwitweten Personen, die in Vollzeit tätig sind. Der Anteil der Personen, die getrennt lebend oder ledig sind und Vollzeit arbeiten, kann nicht vom Anteil der Personen unterschieden werden, die geschieden oder verheiratet sind und eine Vollzeitbeschäftigung haben. Diese Anteile sind jedoch größer als der Anteil der vollzeitbeschäftigten verwitweten Personen.

Für die Tests im Zusammenhang mit *Teilzeit* oder *Ausbildung* werden die Schlüssel A, B und C in der Spalte E angezeigt. In anderen Spalten werden keine weiteren Schlüssel gezeigt. Somit sind die Anteile von Personen, die ledig sind und zur Schule gehen oder Teilzeit arbeiten, größer als die Anteile von verheirateten, verwitweten oder geschiedenen Personen, die zur Schule gehen oder teilzeitbeschäftigt sind.

Für die Tests im Zusammenhang mit *Zeitweise unbeschäftigt* oder *Andere* werden in keiner Spalten weitere Schlüssel ausgegeben. Somit gibt es keine erkennbare Differenz bei den Anteilen der verheirateten, verwitweten, geschiedenen, getrennt lebenden oder ledigen Personen, die zeitweise unbeschäftigt sind oder sich in einer anderen nicht kategorisierten Beschäftigungssituation befinden.

Die Tests im Zusammenhang mit *Ruhestand* zeigen, daß der Anteil der verwitweten Personen im Ruhestand größer ist als die Anteile aller anderen Kategorien von Personen im Ruhestand. Außerdem sind die Anteile von verheirateten und geschiedenen Personen im Ruhestand größer als die Anteile von ledigen Personen im Ruhestand.

Es gibt größere Anteile von verheirateten, verwitweten und getrennt lebenden Personen, die den Haushalt führen, als Anteile von geschiedenen oder ledigen Personen, die den Haushalt führen.

Der Anteil der ledigen Personen, die *arbeitslos, entlassen* sind, ist höher als der Anteil der verheirateten bzw. verwitweten arbeitslosen Personen. Beachten Sie auch, daß die Spalte *getrennt lebend* mit einem “.” gekennzeichnet ist, was darauf hinweist, daß der beobachtete Anteil von getrennt lebenden Personen in der Zeile *arbeitslos, entlassen* 0 oder 1 beträgt und deshalb für arbeitslose Befragte kein Vergleich mit dieser Spalte erfolgen kann.

### **Signifikanzergebnisse durch Notation im APA-Stil**

Wenn Sie nicht möchten, dass die Signifikanzergebnisse in einer separaten Tabelle angezeigt werden, können Sie diese auch in der Haupttabelle anzeigen. Signifikanzergebnisse werden anhand einer Notation im APA-Stil mit tiefgestellten Buchstaben identifiziert. Führen Sie die obigen Schritte zum Vergleich von Spaltenanteilen aus, nehmen Sie jedoch auf der Registerkarte “Teststatistiken” folgende Änderung vor:

- ▶ Wählen Sie im Bereich “Signifikante Unterschiede identifizieren” die Option In der Haupttabelle unter Verwendung von Teilskripts im APA-Stil aus.

Abbildung 10-22  
Registrierkarte "Teststatistiken" mit der Auswahl "Spaltenanteile vergleichen (Z-Tests)" und "In der Haupttabelle, anhand von tiefgestellten Buchstaben"

- Klicken Sie auf OK, um die Tabelle zu erstellen und die Spaltenmitteltests anhand von Notation im APA-Stil durchzuführen.

Abbildung 10-23  
Vergleiche von Spaltenanteilen anhand von Notation im APA-Stil

		Familienstand				
		Verheiratet	Verwitwet	Geschieden	Getrennt	Nie verheiratet
		Anzahl der Spalten (%)				
Berufstätigkeit	Vollzeit berufstätig	57.8% <sub>a</sub>	15.5% <sub>b</sub>	66.1% <sub>c</sub>	62.4% <sub>a,c</sub>	59.1% <sub>a,c</sub>
	Teilzeit berufstätig	10.3% <sub>a</sub>	7.1% <sub>a</sub>	7.8% <sub>a</sub>	9.7% <sub>a,b</sub>	15.4% <sub>b</sub>
	Zeitweilig nicht berufstätig	1.7% <sub>a</sub>	.7% <sub>a</sub>	2.0% <sub>a</sub>	1.1% <sub>a</sub>	1.7% <sub>a</sub>
	Arbeitslos, entlassen	1.0% <sub>a</sub>	1.1% <sub>a</sub>	2.2% <sub>b</sub>	.0%	4.8% <sub>b</sub>
	Im Ruhestand	12.5% <sub>a</sub>	53.0% <sub>b</sub>	11.9% <sub>a</sub>	6.5% <sub>a,c</sub>	2.6% <sub>c</sub>
	Schule	.7% <sub>a</sub>	.4% <sub>a</sub>	1.6% <sub>a</sub>	2.2% <sub>a,b</sub>	9.0% <sub>b</sub>
	Haushalt	14.9% <sub>a</sub>	19.4% <sub>a</sub>	5.6% <sub>b</sub>	14.0% <sub>a</sub>	5.3% <sub>b</sub>
Sonstiges	1.2% <sub>a</sub>	2.8% <sub>a</sub>	2.7% <sub>a</sub>	4.3% <sub>a</sub>	2.1% <sub>a</sub>	

In der Tabelle für den Spaltenanteiletest wird den Kategorien der Spaltenvariablen ein tiefgestellter Buchstabe zugewiesen. Die Spaltenanteile werden für jedes Spaltenpaar mithilfe eines Z-Tests verglichen. Wenn sich ein Wertepaar stark unterscheidet, werden den Werten *unterschiedliche* tiefgestellte Buchstaben zugewiesen.

Für die Testreihen mit *Vollzeit* verfügt die Kategorie “Verwitwet” über einen tiefgestellten Buchstaben, der nicht in den anderen Spalten verwendet wird. Außerdem verfügen die Kategorien “Verheiratet” und “Geschieden” genau wie die Kategorien “Getrennt lebend” und “Ledig” über dieselben tiefgestellten Buchstaben. Anhand dieser tiefgestellten Buchstaben und der tatsächlichen in der Tabelle angezeigten Anteile können Sie die gleiche Analyse durchführen, wie im vorherigen Beispiel mit getrennten Tabellen demonstriert wurde. Somit können Sie schließen, daß der Anteil der geschiedenen Personen, die einer Vollzeitbeschäftigung nachgehen, größer ist als der Anteil der verheirateten Personen, die Vollzeit arbeiten. Dieser Anteil ist wiederum größer als der Anteil der verwitweten Personen, die in Vollzeit tätig sind. Der Anteil der Personen, die getrennt lebend oder ledig sind und Vollzeit arbeiten, kann nicht vom Anteil der Personen unterschieden werden, die geschieden oder verheiratet sind und eine Vollzeitbeschäftigung haben. Diese Anteile sind jedoch größer als der Anteil der vollzeitbeschäftigten verwitweten Personen. Es gilt der Rest der Analyse aus dem vorherigen Beispiel.

### ***Auswirkungen von verschachtelten und gestapelten Variablen auf Spaltenanteiletests***

Die Regel für Spaltenanteiletests lautet folgendermaßen: für jede innerste Untertabelle wird eine separate Reihe paarweiser Tests durchgeführt. Um die Auswirkung von verschachtelten Variablen auf die Tests zu verdeutlichen, betrachten Sie das vorhergehende Beispiel, wobei jedoch *Beschäftigungsstatus* in Ebenen von *Geschlecht* verschachtelt wird.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen erneut (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Ziehen Sie die Variable *Geschlecht* aus der Variablenliste, und legen Sie sie im Spaltenbereich auf der Zeichenfläche ab.
- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 10-24  
Vergleich der Spaltenanteile

				Familienstand				
				Verheiratet (A)	Verwitwet (B)	Geschieden (C)	Getrennt lebend (D)	Ledig (E)
Geschlecht	Männlich	Beschäftigungsstatus	Vollzeit	B		B	B	B
			Teilzeit					A
		Arbeitslos, entlassen	Zeitweise unbeschäftigt		.a			
			Arbeitslos, entlassen		.a		.a	A
			Ruhestand	E	A C D E	E		
		Ausbildung	Arbeit im Haushalt		.a			A C
			Andere				A	
	Weiblich	Beschäftigungsstatus	Vollzeit	B		AB	B	B
			Teilzeit	B			B	B
		Arbeitslos, entlassen	Zeitweise unbeschäftigt				.a	
			Arbeitslos, entlassen				.a	A
			Ruhestand	E	A C D E	E		
		Ausbildung	Arbeit im Haushalt	C E	C E			A B C
			Andere				C	

Wenn *Beschäftigungsstatus* in Ebenen von *Geschlecht* verschachtelt ist, werden 14 Reihen von Spaltenanteiletests durchgeführt: eine für jede Ebene von *Beschäftigungsstatus* für jede Ebene von *Geschlecht*. Den Kategorien von *Familienstand* werden die gleichen Buchstabenschlüssel zugewiesen.

Beachten Sie die folgenden Punkte bei den Tabellenergebnissen:

- Wenn mehr Tests durchgeführt werden, sind mehr Spalten mit einem Spaltenanteil von null vorhanden. Diese kommen bei getrennt lebenden Befragten und verwitweten Männern am häufigsten vor.
- Die Spaltendifferenzen bei den Befragten, die den *Haushalt* führen, können scheinbar gänzlich auf Frauen zurückgeführt werden.

Gehen Sie folgendermaßen vor, um die Auswirkung von gestapelten Variablen auf die Tests zu verdeutlichen:

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen erneut (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Ziehen Sie die Variable *Höchster Schulabschluß* aus der Variablenliste, und legen Sie sie im Zeilenbereich unter *Geschlecht* ab.
- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 10-25  
Vergleich der Spaltenanteile

				Familienstand				
				Verheiratet (A)	Verwitwet (B)	Geschieden (C)	Getrennt lebend (D)	Ledig (E)
Geschlecht	Männlich	Beschäftigungsstatus	Vollzeit	B		B	B	B
			Teilzeit					A
			Zeitweise unbeschäftigt		. <sup>a</sup>			
			Arbeitslos, entlassen		. <sup>a</sup>		. <sup>a</sup>	A
			Ruhestand	E	A C D E	E		
			Ausbildung		. <sup>a</sup>			A C
			Arbeit im Haushalt					
	Weiblich	Beschäftigungsstatus	Andere			A B	A	B
			Vollzeit	B			B	B
			Teilzeit	B				
			Zeitweise unbeschäftigt				. <sup>a</sup>	
			Arbeitslos, entlassen				. <sup>a</sup>	A
			Ruhestand	E	A C D E	E		
			Ausbildung					A B C
Höchster Schulabschluß	Hauptschule Realschule Gymnasium Hochschulabschluß Promotion	Arbeit im Haushalt	C E	C E		C		
		Andere						
				A C E				
			B		B		B	
			B				B	

Wenn *Höchster Schulabschluß* mit *Geschlecht* gestapelt ist, werden 19 Reihen von Spaltenmitteltests durchgeführt: die 14 bereits beschriebenen sowie eine für jede Ebene von *Höchster Schulabschluß*. Den Kategorien von *Familienstand* werden die gleichen Buchstabenschlüssel zugewiesen.

Beachten Sie die folgenden Punkte bei den Tabellenergebnissen:

- Die Testergebnisse für die 14 bereits durchgeführten Testreihen sind dieselben.
- Personen, die nur einen Hauptschulabschluß besitzen, sind unter den verwitweten Befragten häufiger anzutreffen als unter den verheirateten, geschiedenen oder ledigen.
- Personen mit einem Hochschulabschluß finden sich tendenziell häufiger unter den verheirateten, geschiedenen und ledigen Personen als unter verwitweten Personen.

### **Anmerkung zu Gewichtungen und Mehrfachantworten-Sets**

Fallgewichtungen basieren immer auf Häufigkeiten, nicht auf Antworten, auch wenn es sich bei einer der Variablen um eine Mehrfachantworten-Variable handelt.

# Mehrfachantworten-Sets

In benutzerdefinierten Tabellen und der Diagrammerstellung kann eine besondere Art von “Variable” verwendet werden, die als **Mehrfachantworten-Set** bezeichnet wird. Bei Mehrfachantworten-Sets handelt es sich nicht um “Variablen” im üblichen Sinn. Mehrfachantworten-Sets können nicht im Daten-Editor angezeigt werden, sie werden von anderen Prozeduren nicht erkannt. Mehrfachantworten-Sets verwenden mehrere Variablen, um Antworten auf Fragen aufzuzeichnen, auf welche der Befragte mehr als eine Antwort geben kann. Sie werden wie kategoriale Variablen behandelt und bieten weitestgehend dieselben Möglichkeiten wie kategoriale Variablen.

Mehrfachantworten-Sets werden aus mehreren Variablen in der Datendatei gebildet. Beim Mehrfachantworten-Set handelt es sich um ein spezielles Konstrukt innerhalb einer Datendatei. Mehrfachantworten-Sets können in einer Datendatei im IBM® SPSS® Statistics-Format definiert und gespeichert werden. Das Importieren oder Exportieren von Mehrfachantworten-Sets in bzw. aus anderen Dateiformaten ist jedoch nicht möglich. Mit der Option “Dateneigenschaften kopieren” können Sie Mehrfachantworten-Sets aus anderen SPSS Statistics-Datendateien kopieren. Sie finden diese Option im Menü “Daten” im Fenster “Daten-Editor”.

## **Beispiel-Datendatei**

Für die Beispiele dieses Kapitels wird die Datendatei *survey\_sample.sav* verwendet. [Für weitere Informationen siehe Thema Beispieldateien in Anhang A auf S. 201.](#)

In allen angegebenen Beispielen werden die Variablenlabels in Dialogfeldern in alphabetischer Reihenfolge angezeigt. Die Anzeigeeigenschaften für Variablenlisten werden im Dialogfeld “Optionen” auf der Registerkarte “Allgemein” festgelegt (Menü “Bearbeiten”, Befehl “Optionen”).

## **Häufigkeiten, Antworten, Prozentwerte und Gesamtwerte**

Für Mehrfachantworten-Sets sind dieselben Auswertungsstatistiken verfügbar wie für kategoriale Variablen. Darüber hinaus stehen für Mehrfachantworten-Sets einige zusätzliche Statistiken zur Verfügung.

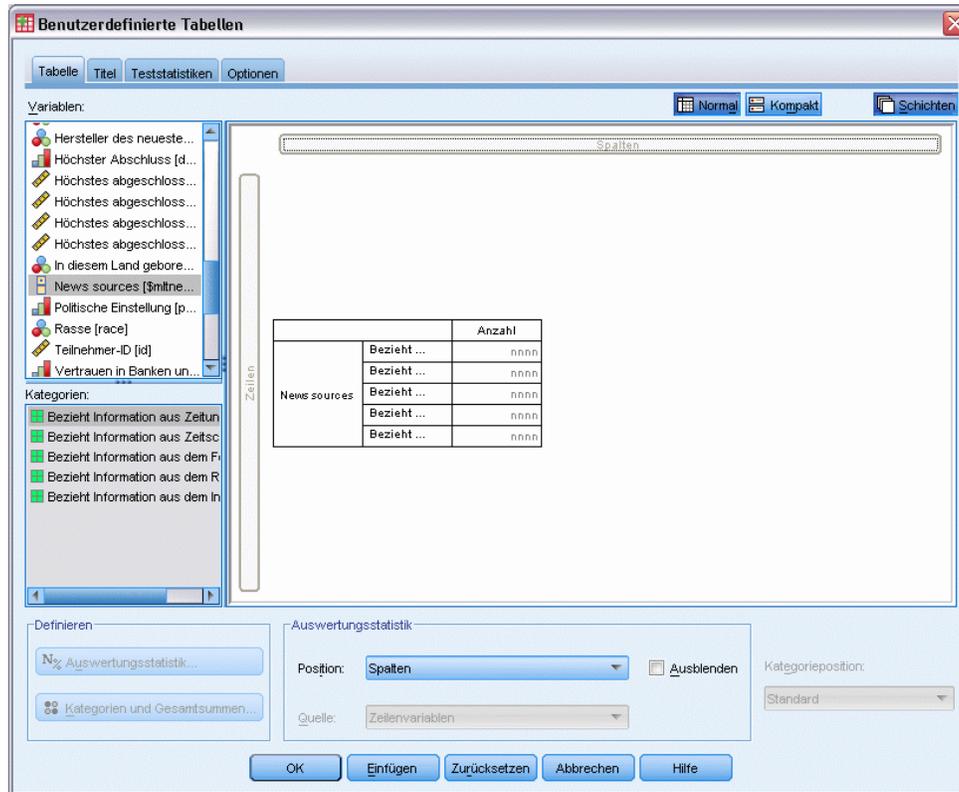
- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
Analysieren > Tabellen > Benutzerdefinierte Tabellen...
- ▶ Ziehen Sie *Nachrichtenquellen*, die aussagekräftige Beschriftung für das Mehrfachantworten-Set *\$mltnews*, aus der Variablenliste auf die Zeichenfläche, und legen Sie es im Bereich für die Zeilen ab.

Das Symbol neben der Variablen in der Variablenliste kennzeichnet die Variable als Set aus dichotomen Variablen.

Abbildung 11-1  
Symbol "Set aus dichotomen Variablen"



Abbildung 11-2  
Set aus dichotomen Variablen in der Tabellenvorschau



Bei einem Set aus dichotomen Variablen ist jede "Kategorie" tatsächlich eine separate Variable, und die Kategoriebeschriftungen sind die Variablenlabels (bzw. die Variablennamen für Variablen, für die keine Variablenlabels definiert wurden). In diesem Beispiel repräsentieren die angezeigten Häufigkeiten die Anzahl der Fälle, in denen die Antwort für eine Variable im Set *Ja* lautete.

- ▶ Klicken Sie in der Tabellenvorschau auf der Zeichenfläche mit der rechten Maustaste auf *Nachrichtenquellen* und wählen Sie im Kontextmenü die Option *Kategorien und Gesamtsummen* aus.
- ▶ Wählen Sie im Dialogfeld "Kategorien und Gesamtsummen" die Option *Gesamt* aus, indem Sie darauf klicken, und klicken Sie anschließend auf *Zuweisen*.
- ▶ Klicken Sie mit der rechten Maustaste erneut auf *Nachrichtenquellen* und wählen Sie im Kontextmenü *Auswertungsstatistik* aus.
- ▶ Wählen Sie im Dialogfeld "Auswertungsstatistik" in der Liste "Statistik" die Option *Anzahl* als *Spalten%* aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste "Anzeigen" zu übertragen.

- Klicken Sie auf Der Auswahl zuweisen und anschließend auf OK, um die Tabelle zu erstellen.

Abbildung 11-3

Häufigkeiten und Spaltenprozentage in einem Set aus dichotomen Variablen

		Anzahl	Spalten%
Nachrichtenquellen	Nachrichten aus dem Fernsehen	1077	51,8%
	Nachrichten aus dem Internet	867	41,7%
	Nachrichten aus dem Radio	551	26,5%
	Nachrichten aus Zeitschriften	294	14,1%
	Nachrichten aus Zeitungen	805	38,7%
	Gesamt	2081	100,0%

### Unstimmige Summen

Wenn Sie die Zahlen in der Tabelle betrachten, fällt Ihnen möglicherweise auf, dass ein relativ großer Unterschied zwischen dem Wert für "Gesamtsumme" und der angenommenen Summe aus den einzelnen Werten besteht. Dies gilt insbesondere, wenn die Gesamtsumme wesentlich niedriger ausfällt, als erwartet. Der Grund dafür besteht darin, dass die Anzahl für die einzelnen Kategorien in der Tabelle der Anzahl der Fälle mit dem Wert 1 (einer *Ja*-Antwort) für diese Variable entspricht und die Gesamtzahl der *Ja*-Antworten für alle fünf Variablen im Set aus dichotomen Variablen leicht die Gesamtzahl der Fälle in der Datendatei übersteigen kann.

Demgegenüber wird die Gesamthäufigkeit aus der Summe der Fälle mit einer *Ja*-Antwort für mindestens eine Variable im Set gebildet und kann daher in keinem Fall die Gesamtzahl der Fälle in der Datendatei überschreiten. In diesem Beispiel ist die Gesamthäufigkeit von 2.081 fast 800 niedriger als die Gesamtzahl der Fälle in der Datendatei. Wenn keine Variable fehlende Werte enthält, bedeutet das, dass fast 800 Befragte angegeben haben, dass Sie aus keiner der angegebenen Quellen Nachrichten beziehen. Die Spaltenprozentage werden auf der Grundlage der Gesamthäufigkeit ermittelt. Daher werden die Spaltenprozentage in diesem Beispiel auf mehr als die 100 % summiert, die als Wert für die Gesamtspaltenprozentage angezeigt werden.

### Stimmige Summen

Obwohl "Häufigkeiten" normalerweise ein relativ eindeutiger Begriff ist, zeigt das obige Beispiel, dass dennoch Verwirrung im Zusammenhang mit Gesamtsummen für Mehrfachantworten-Sets auftreten kann. Für Mehrfachantworten-Sets bezieht sich die tatsächlich gewünschte Auswertungsstatistik oftmals auf *Antworten*.

- Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü "Analysieren", Untermenü "Tabellen", Befehl "Benutzerdefinierte Tabellen").
- Klicken Sie mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf die Variable *Nachrichtenquellen* und wählen Sie im Kontextmenü die Option Auswertungsstatistik aus.
- Wählen Sie im Dialogfeld "Auswertungsstatistik" in der Liste "Statistik" die Option Antworten aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste "Anzeigen" zu übertragen.

- ▶ Wählen Sie in der Liste “Statistik” die Option Antworten als Spalten% aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” zu übertragen.
- ▶ Klicken Sie auf Der Auswahl zuweisen und anschließend auf OK, um die Tabelle zu erstellen.

Abbildung 11-4

Antworten und Antworten als Spaltenprozentwerte in einem Set aus dichotomen Variablen

		Anzahl	Spalten%	Antworten	Antworten als Spalten%
Nachrichtenquellen	Nachrichten aus dem Fernsehen	1077	51,8%	1077	30,0%
	Nachrichten aus dem Internet	867	41,7%	867	24,1%
	Nachrichten aus dem Radio	551	26,5%	551	15,3%
	Nachrichten aus Zeitschriften	294	14,1%	294	8,2%
	Nachrichten aus Zeitungen	805	38,7%	805	22,4%
	Gesamt	2081	100,0%	3594	100,0%

Für jede “Kategorie” in einem Set aus dichotomen Variablen ist der Wert für *Antworten* identisch mit dem Wert für *Anzahl*. Dies ist immer der Fall bei Sets aus dichotomen Variablen. Die Gesamtsummen können sich jedoch deutlich unterscheiden. Die Gesamtzahl der Antworten beträgt 3.594. Dieser Wert übersteigt die Gesamthäufigkeit um mehr als 1.500 und die Gesamtzahl der Fälle in der Datendatei um mehr als 700.

Bei den Prozentwerten betragen die Gesamtwerte für *Anzahl als Spalten%* und *Antworten als Spalten%* beide 100 %. Für die einzelnen Kategorien im Set für dichotome Variablen sind die Prozentwerte bei den Antworten als Spaltenprozente jedoch wesentlich niedriger. Der Grund dafür liegt in der Tatsache, dass als Grundlage zur Berechnung der Prozentwerte für die Antworten als Spaltenprozente die Gesamtzahl der Antworten verwendet wird, die im gegebenen Fall 3.594 beträgt. Dies führt zu wesentlich niedrigeren Prozentwerten, als eine Berechnung auf der Grundlage des Werts 2.081, anhand dessen die Spaltenprozentwerte berechnet werden.

### **Die Gesamtsummen für die Prozentwerte sind größer als 100 %**

Sowohl die Spaltenprozentwerte als auch die Werte für Antworten als Spaltenprozente erzielen in diesem Beispiel Gesamtprozentwerte von 100 %, obgleich sich die einzelnen Werte in der Spalte *Anzahl als Spalten%* auf deutlich mehr als 100 % summieren lassen. Wie können Sie also Prozentwerte auf der Grundlage der Gesamthäufigkeit und nicht anhand der Gesamtanzahl der Antworten darstellen und gleichzeitig dafür sorgen, dass der Prozentwert für die “Gesamtsumme” präzise die Summe der Prozentwerte für die einzelnen Kategorien widerspiegelt?

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf die Variable *Nachrichtenquellen* und wählen Sie im Kontextmenü die Option Auswertungsstatistik aus.
- ▶ Wählen Sie im Dialogfeld “Auswertungsstatistik” in der Liste “Statistik” die Option Antworten als Spalten% (Basis: Anzahl) aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” zu übertragen.

- Klicken Sie auf Der Auswahl zuweisen und anschließend auf OK, um die Tabelle zu erstellen.

Abbildung 11-5

Berechnung der Werte für Antworten als Spaltenprozente auf der Grundlage der Prozentwerte für die Anzahl

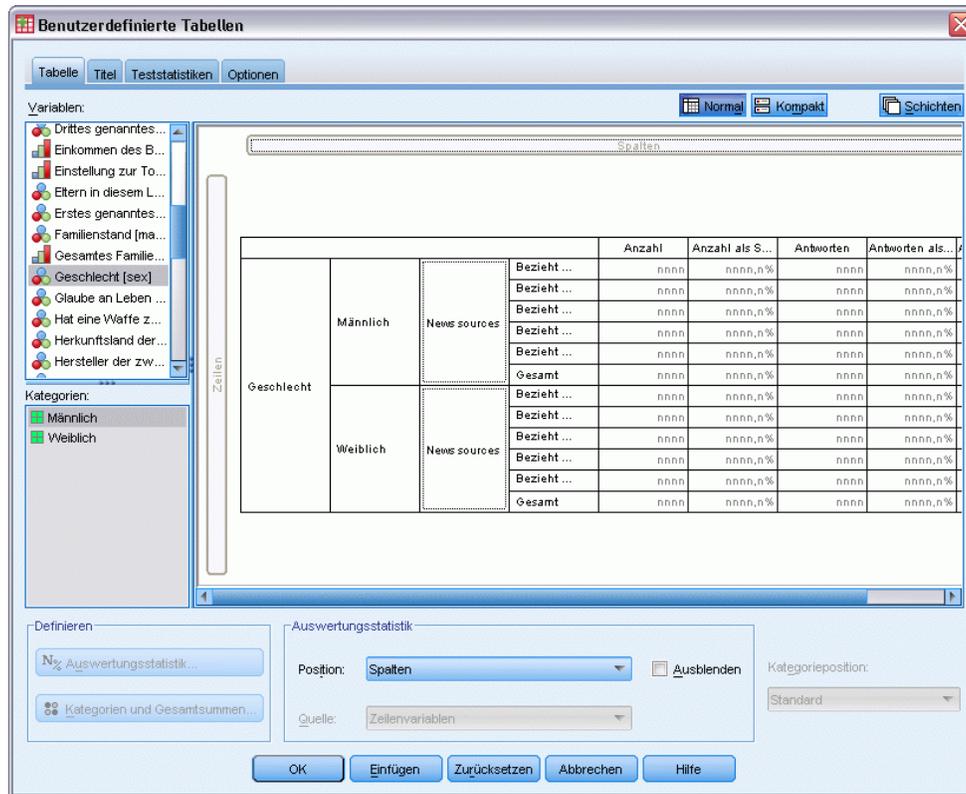
		Anzahl	Spalten%	Antworten	Antworten als Spalten%	Antworten als Spalten% (Basis: Anzahl)
Nachrichtenquellen	Nachrichten aus dem Fernsehen	1077	51,8%	1077	30,0%	51,8%
	Nachrichten aus dem Internet	867	41,7%	867	24,1%	41,7%
	Nachrichten aus dem Radio	551	26,5%	551	15,3%	26,5%
	Nachrichten aus Zeitschriften	294	14,1%	294	8,2%	14,1%
	Nachrichten aus Zeitungen	805	38,7%	805	22,4%	38,7%
	Gesamt	2081	100,0%	3594	100,0%	172,7%

## Verwenden von Mehrfachantworten-Sets mit anderen Variablen

Generell werden Mehrfachantworten-Sets genauso verwendet wie kategoriale Variablen. Sie können beispielsweise ein Mehrfachantworten-Set zusammen mit einer kategorialen Variablen als Kreuztabelle darstellen lassen oder es innerhalb einer kategorialen Variablen verschachteln.

- Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü "Analysieren", Untermenü "Tabellen", Befehl "Benutzerdefinierte Tabellen").
- Ziehen Sie die Variable *Geschlecht* aus der Variablenliste in das Vorschaufenster, und legen Sie sie auf der linken Seite im Bereich für die Zeilen ab. Dadurch wird das Mehrfachantworten-Set *Nachrichtenquellen* innerhalb der Geschlechtskategorien verschachtelt.

Abbildung 11-6  
Tabellenvorschau für verschachtelte Mehrfachantworten-Sets



- ▶ Klicken Sie mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf *Geschlecht*, und deaktivieren Sie im Kontextmenü die Option Variablenlabel anzeigen.
- ▶ Führen Sie denselben Schritt für *Nachrichtenquellen* aus.

Daraufhin werden die Spalten mit den Variablenlabels aus der Tabelle entfernt, da sie in diesem Fall nicht erforderlich sind.

- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 11-7  
Mehrfachantworten-Sets in einer kategorialen Variablen verschachtelt

		Anzahl	Anzahl als ...	Reaktionen	Reaktionen in %	Reaktionen in % (Base: Anzahl)
männlich	Informationen vom Internet	359	40.1%	359	23.3%	40.1%
	Informationen vom Radio	233	26.0%	233	15.1%	26.0%
	Informationen vom Fernsehen	451	50.3%	451	29.3%	50.3%
	Informationen vom Magazin	121	13.5%	121	7.9%	13.5%
	Informationen von Zeitungen	375	41.9%	375	24.4%	41.9%
	Gesamt	896	100.0%	1539	100.0%	171.8%
weiblich	Informationen vom Internet	508	42.9%	508	24.7%	42.9%
	Informationen vom Radio	318	26.8%	318	15.5%	26.8%
	Informationen vom Fernsehen	626	52.8%	626	30.5%	52.8%
	Informationen vom Magazin	173	14.6%	173	8.4%	14.6%
	Informationen von Zeitungen	430	36.3%	430	20.9%	36.3%
	Gesamt	1185	100.0%	2055	100.0%	173.4%

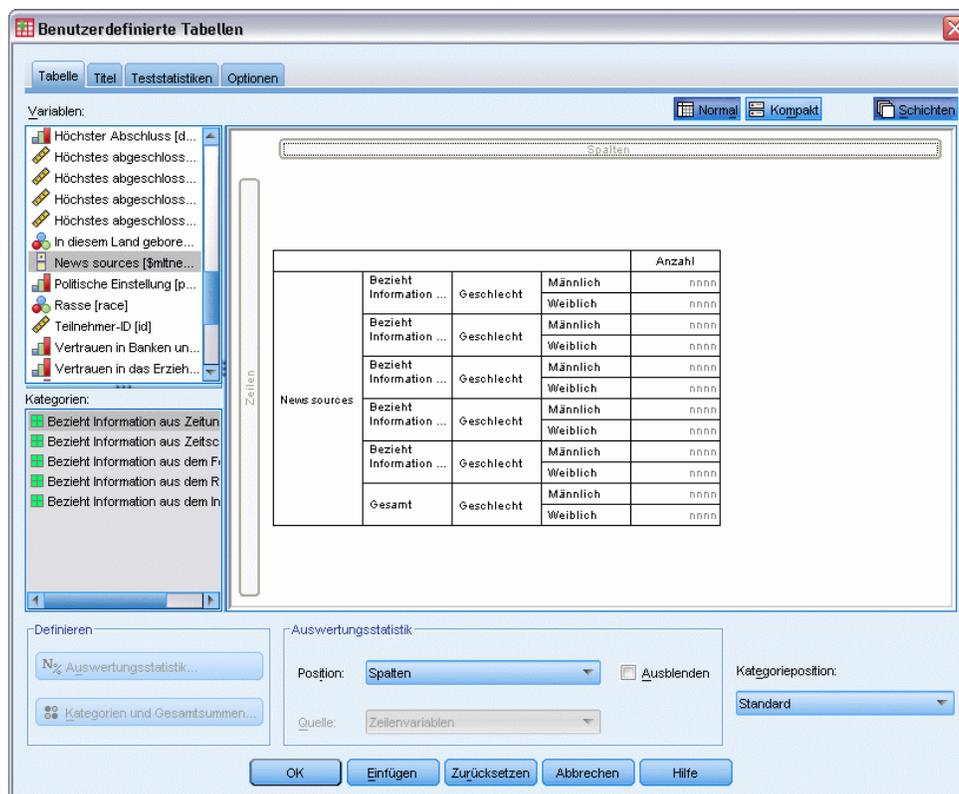
## Statistikquellvariablen und verfügbare Auswertungsstatistiken

Wenn eine Tabelle keine metrische Variable enthält, werden kategoriale Variablen und Mehrfachantworten-Sets hinsichtlich der Statistikquellvariablen identisch behandelt. Die Statistikquellvariable ist die innerste verschachtelte Variable in der Quelldimension der Statistik. Da einige Auswertungsstatistiken ausschließlich nur Mehrfachantworten-Sets zugewiesen werden können, muss das Mehrfachantworten-Set die innerste verschachtelte Variable in der Statistikquelldimension bilden, wenn Sie eine spezielle Auswertungsstatistik für Mehrfachantworten erhalten möchten.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Ändern Sie die Verschachtelungsreihenfolge, indem Sie die Variable *Nachrichtenquellen* aus der Tabellenvorschau auf der Zeichenfläche ziehen und sie links von der Variablen *Geschlecht* ablegen.

Abbildung 11-8

Kategoriale Variable in Mehrfachantworten-Set verschachtelt



Alle speziellen Auswertungsstatistiken für Mehrfachantworten, d. h. Antworten und Antworten als Spaltenprozent, werden aus der Tabellenvorschau entfernt, weil die kategoriale Variable *Geschlecht* nun die innerste verschachtelte Variable und damit die Quellvariable der Statistik darstellt.

Glücklicherweise werden diese Einstellungen vorübergehend im Dialogfeld zum Konstruieren von Tabellen gespeichert. Wenn Sie die Variable *Nachrichtenquellen* zur ursprünglichen Verschachtelung in der Variablen *Geschlecht* zurückverschieben, werden alle Auswertungsstatistiken im Zusammenhang mit Antworten in der Tabellenvorschau wiederhergestellt.

## **Sets aus kategorialen Variablen und doppelte Antworten**

Sets aus kategorialen Variablen verfügen über eine Funktion, die bei Sets aus dichotomen Variablen nicht zur Verfügung steht, und zwar über die Möglichkeit, doppelte Antworten zu zählen. In vielen Fällen handelt es sich bei doppelten Antworten in Sets aus kategorialen Variablen vermutlich um Kodierungsfehler. Wenn die Frage in einer Umfrage beispielsweise “Welche drei Nationen produzieren Ihrer Meinung nach die besten Autos?” lautet, ist die Antwort *Schweden, Deutschland und Schweden* wahrscheinlich ungültig.

In anderen Fällen können Doppeltantworten jedoch durchaus gültig sein. Auf die Frage “In welchen Ländern wurden Ihre letzten drei Autos produziert?” wäre die Antwort *Schweden, Deutschland und Schweden* beispielsweise völlig plausibel.

Benutzerdefinierte Tabellen bieten die Möglichkeit, doppelte Antworten in Sets aus kategorialen Variablen zu berücksichtigen. In der Standardeinstellung werden doppelte Antworten nicht gezählt, Sie können jedoch angeben, dass Doppeltantworten in die Zählung aufgenommen werden.

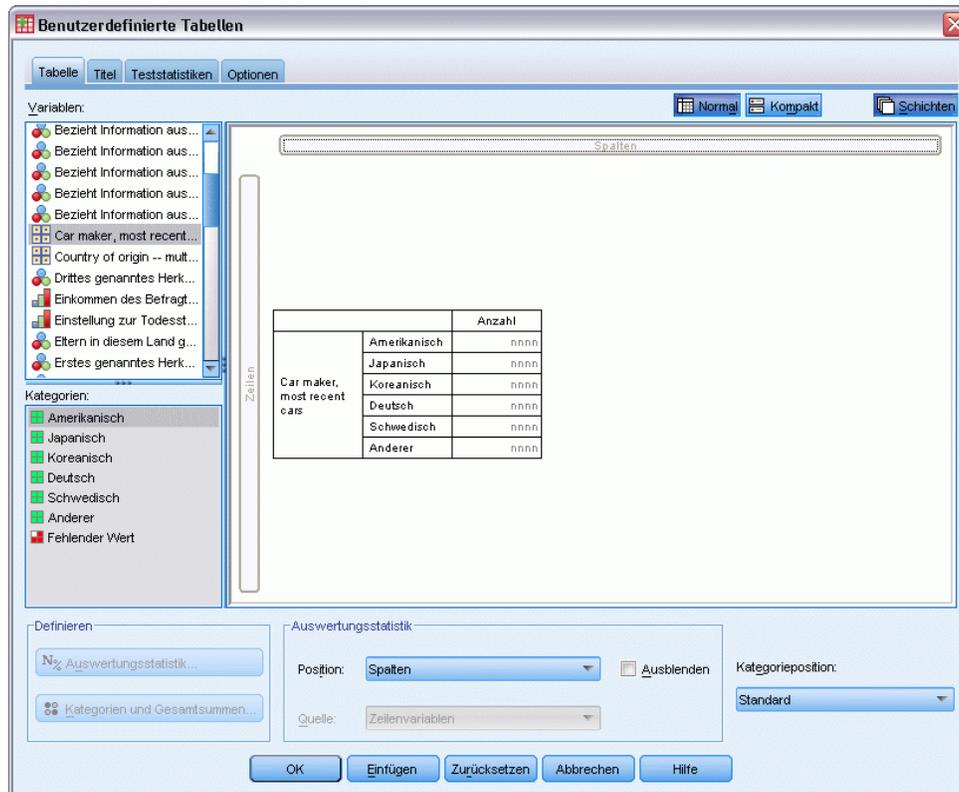
- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie auf Zurücksetzen, um frühere Einstellungen zu löschen.
- ▶ Ziehen Sie die Variable *Autohersteller, aktuelle Kfz-Modelle* aus der Variablenliste auf die Zeichenfläche, und legen Sie sie im Bereich für die Zeilen ab.

Das Symbol neben der Variablen in der Variablenliste kennzeichnet die Variable als Set aus kategorialen Variablen.

Abbildung 11-9  
Symbol “Set aus kategorialen Variablen”



Abbildung 11-10  
Set aus kategorialen Variablen in der Tabellenvorschau



Bei Sets aus kategorialen Variablen repräsentieren die angezeigten Kategorien die Wertelabels, die für alle Variablen im Set definiert sind. (Bei Sets aus dichotomen Variablen stellen die Kategorien hingegen die Variablenlabels für die einzelnen Variablen im Set dar.)

- ▶ Klicken Sie mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf die Variable *Autohersteller, aktuelle Kfz-Modelle* und wählen Sie im Kontextmenü die Option *Kategorien und Gesamtsummen* aus.
- ▶ Wählen Sie im Dialogfeld “Kategorien und Gesamtsummen” die Option *Gesamt* aus, indem Sie darauf klicken, und klicken Sie anschließend auf *Zuweisen*.
- ▶ Klicken Sie mit der rechten Maustaste erneut auf die Variable *Autohersteller, aktuelle Kfz-Modelle* und wählen Sie im Kontextmenü die Option *Auswertungsstatistik* aus.
- ▶ Wählen Sie im Dialogfeld “Auswertungsstatistik” in der Liste “Statistik” die Option *Antworten* aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” zu übertragen.
- ▶ Klicken Sie auf *Der Auswahl zuweisen* und anschließend auf *OK*, um die Tabelle zu erstellen.

Abbildung 11-11

Set aus kategorialen Variablen: Häufigkeiten und Antworten ohne Duplikate

		Anzahl	Antworten
Autohersteller, aktuelle Kfz-Modelle	Amerikanisch	1938	1938
	Japanisch	1327	1327
	Koreanisch	695	695
	Deutsch	693	693
	Schwedisch	360	360
	Andere	343	343
	Gesamt	2832	5356

In der Standardeinstellung werden doppelte Antworten nicht gezählt, daher sind die Werte in dieser Tabelle für die einzelnen Kategorien in den Spalten *Anzahl* und *Antworten* identisch. Lediglich die Gesamtsummen unterscheiden sich.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie auf die Registerkarte Optionen.
- ▶ Wählen Sie die Option Doppeltantworten für Sets aus kategorialen Variablen zählen aus, indem Sie darauf klicken.
- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 11-12

Set aus kategorialen Variablen mit Doppeltantworten

		Anzahl	Antworten
Autohersteller, aktuelle Kfz-Modelle	Amerikanisch	1938	2797
	Japanisch	1327	1717
	Koreanisch	695	760
	Deutsch	693	754
	Schwedisch	360	383
	Andere	343	359
	Gesamt	2832	6770

In dieser Tabelle besteht ein deutlicher Unterschied zwischen den Werten in der Spalte *Anzahl* und den Werten in der Spalte *Antworten*, insbesondere zwischen Werten für amerikanische Autos. Dies ist ein Hinweis darauf, dass viele Befragte Halter mehrerer amerikanischer Autos waren.

## Signifikanztests bei Mehrfachantworten-Sets

Mehrfachantworten-Sets werden bei Signifikanztests im Wesentlichen wie kategoriale Variablen verwendet.

- Bei Tests auf Unabhängigkeit (Chi-Quadrat) oder Vergleichen der Spaltenanteile (Z-Tests) werden die Tests mit den Häufigkeiten durchgeführt. Die Häufigkeiten müssen daher eine der Auswertungsstatistiken sein, die in der Tabelle angezeigt werden.
- Bei Sets aus kategorialen Variablen werden keine Test durchgeführt, in denen Spaltenanteile oder Spaltenmittel verglichen werden (T-Tests), wenn auf der Registerkarte “Optionen” die Option Doppelantworten für multiple Kategoriensets zählen ausgewählt ist. [Für weitere Informationen siehe Thema Benutzerdefinierte Tabellen: Registerkarte “Optionen” in Kapitel 2 auf S. 49.](#)

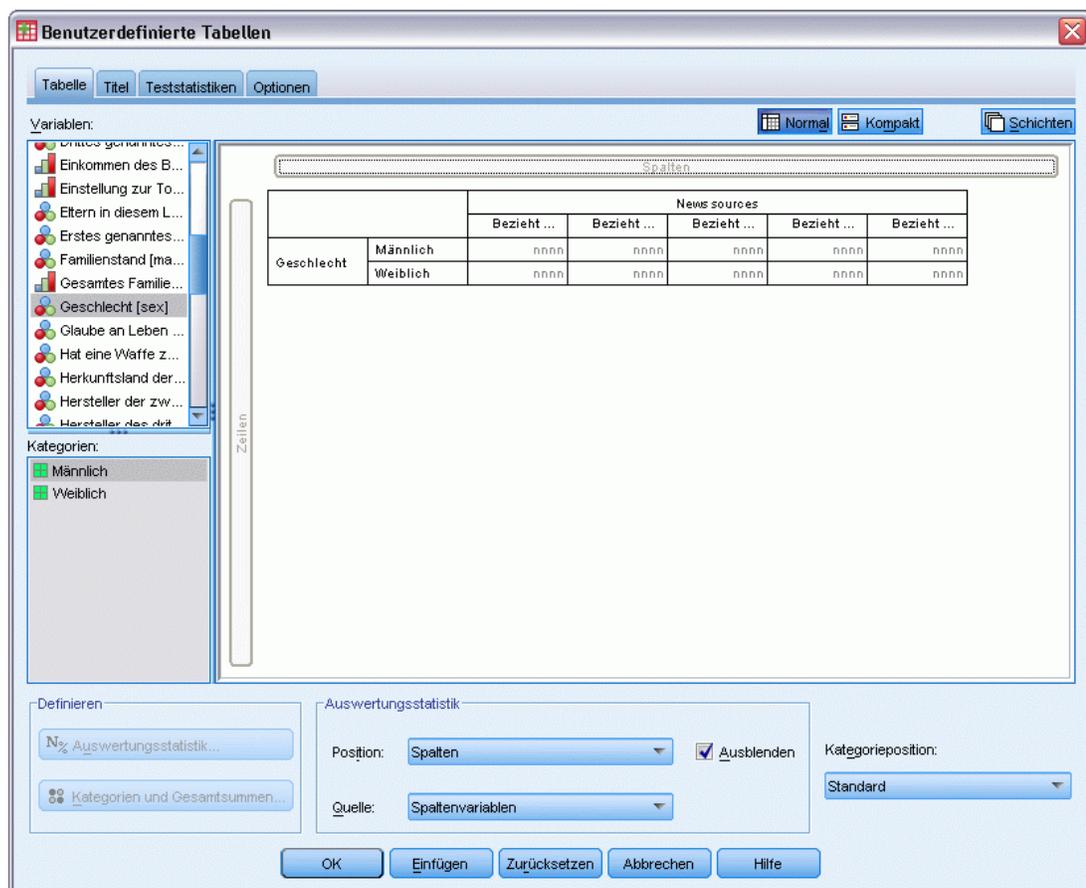
## Test auf Unabhängigkeit bei Mehrfachantworten-Sets

In diesem Beispiel wird eine Kreuztabelle einer kategorialen Variablen und eines Mehrfachantworten-Sets erstellt, und für die Kreuztabelle wird ein Chi-Quadrat-Test auf Unabhängigkeit durchgeführt.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie auf Zurücksetzen, um frühere Einstellungen zu löschen.
- ▶ Ziehen Sie *Nachrichtenquellen*, das aussagekräftige Label für das Set aus dichotomen Variablen *\$mltnews*, aus der Variablenliste auf die Zeichenfläche, und legen Sie es im Bereich für die Spalten ab.
- ▶ Ziehen Sie die Variable *Geschlecht* aus der Variablenliste, und legen Sie sie im Spaltenbereich auf der Zeichenfläche ab.

Abbildung 11-13

*Nachrichtenquellen* (Mehrfachantworten-Set) und *Geschlecht* ausgewählt als Spalten- und Zeilenvariablen



- ▶ Klicken Sie auf die Registerkarte Teststatistiken.

- ▶ Wählen Sie Tests auf Unabhängigkeit (Chi-Quadrat) aus.
- ▶ Falls die Option Mehrfachantwort-Variablen in Tests einschließen nicht bereits ausgewählt ist, wählen Sie sie aus.

Abbildung 11-14  
Registerkarte "Teststatistiken"

**Benutzerdefinierte Tabellen**

Tabellen | Titel | **Teststatistiken** | Optionen

Spaltenmittel vergleichen (T-Tests)  
Alpha: 0.05  
 p-Werte für Mehrfachvergleiche anpassen (Bonferroni-Methode)  
 Varianz nur anhand der verglichenen Kategorien schätzen (immer der Fall bei Mehrfachantwort-Variablen)

Spaltenanteile vergleichen (Z-Tests)  
Alpha: 0.05  
 p-Werte für Mehrfachvergleiche anpassen (Bonferroni-Methode)

Signifikante Unterschiede identifizieren:  
 In einer eigenen Tabelle  
 In der Haupttabelle unter Verwendung von Teilskripts im APA-Stil

Tests auf Unabhängigkeit (Chi-Quadrat) Alpha: 0.05

Zwischenergebnisse anstelle von Kategorien für Zwischenergebnisse verwenden

Mehrfachantwort-Variablen in Tests einschließen

**i** Chi-Quadrat- und Spaltenanteiltests beziehen sich auf Tabellen, die kategoriale Variablen sowohl in den Zeilen als auch in den Spalten enthalten.

Der Spaltenmitteltest bezieht sich auf Tabellen, die metrische Variablen in den Zeilen und kategoriale Variablen in den Spalten enthalten.

Es werden keine Tests für Tabellen durchgeführt, in denen die Kategoriebeschriftungen aus der Standard-Tabellendimension verschoben wurden.

Gesamtsummen werden aus allen Tests ausgeschlossen. Zwischensummen werden nur verwendet, wenn die Kategorien, auf die sie sich beziehen, ausgeblendet sind oder wenn oben angegeben.

Berechnete Kategorien sind von Signifikanztests ausgeschlossen.

OK Einfügen Zurücksetzen Abbrechen Hilfe

- ▶ Klicken Sie auf OK, um die Prozedur auszuführen.

Abbildung 11-15  
Ergebnisse des Chi-Quadrat-Tests

		News sources
Gender	Chi-square	10.266
	df	5
	Sig.	.068

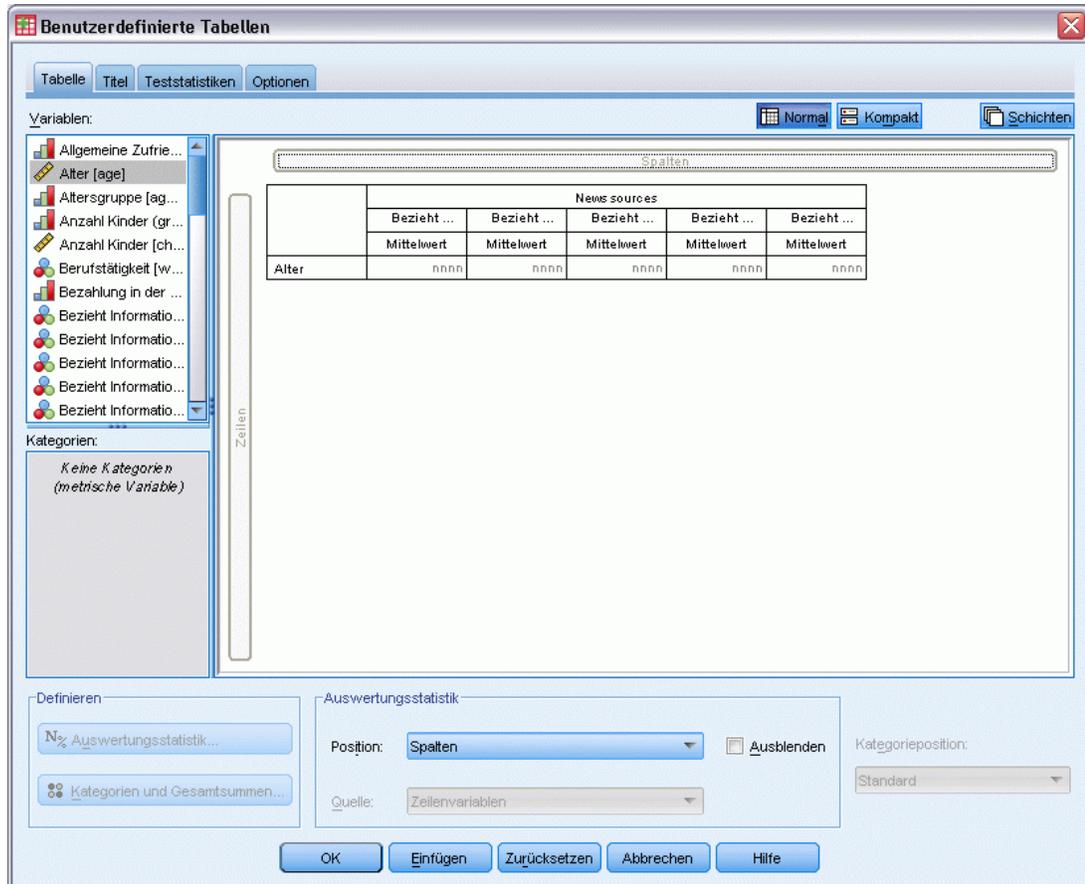
Das Signifikanzniveau von 0,068 für den Chi-Quadrat-Test deutet an, dass sich Männer und Frauen wahrscheinlich nicht signifikant darin unterscheiden, aus welchen Quellen sie ihre Nachrichten beziehen. (Hierbei wird angenommen, dass Sie einen Signifikanzwert von höchstens 0,05 als Kriterium zum Bestimmen der statistischen Signifikanz heranziehen.)

### **Vergleichen von Spaltenmitteln mit Mehrfachantworten-Sets**

In diesem Beispiel werden die Mittelwerte einer metrischen Variablen innerhalb von Kategorien berechnet, die durch ein Mehrfachantworten-Set festgelegt werden. Außerdem wird jedes Kategorienmittel mit jedem anderen Kategorienmittel verglichen, um signifikante Unterschiede zu ermitteln.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie auf Zurücksetzen, um frühere Einstellungen zu löschen.
- ▶ Ziehen Sie *Nachrichtenquellen*, das aussagekräftige Label für das Set aus dichotomen Variablen *\$mltnews*, aus der Variablenliste auf die Zeichenfläche, und legen Sie es im Bereich für die Spalten ab.
- ▶ Ziehen Sie *Alter der befragten Person* auf die Zeichenfläche, und legen Sie es im Bereich für die Zeilen ab.

Abbildung 11-16  
Vergleich des mittleren Alters in den Kategorien eines Mehrfachantwort-Sets



- ▶ Klicken Sie auf die Registerkarte Teststatistiken.
- ▶ Wählen Sie Spaltenmittel vergleichen (T-Tests) aus.
- ▶ Falls die Option Mehrfachantwort-Variablen in Tests einschließen nicht bereits ausgewählt ist, wählen Sie sie aus.
- ▶ Klicken Sie auf OK, um die Prozedur auszuführen.

Abbildung 11-17  
Ergebnisse des Signifikanztests

	Informationsquellen				
	Informationen von Zeitungen	Informationen von Magazinen	Informationen vom Fernsehen	Informationen vom Radio	Informationen vom Internet
	Mittelwert	Mittelwert	Mittelwert	Mittelwert	Mittelwert
Alter des Befragten	52	40	48	40	40

**Vergleich der Spaltenmittelwerte**

	Informationsquellen				
	Informationen von Zeitungen	Informationen von Magazinen	Informationen vom Fernsehen	Informationen vom Radio	Informationen vom Internet
	(A)	(B)	(C)	(D)	(E)
Alter des Befragten	B C D E		B D E		

Ergebnisse basieren auf zweiseitigem Test mit vorausgesetzter Varianz mit Signifikanzlevel 0,05. Für jedes Signifikanzpaar erscheint der Schlüssel für kleiner Kategorie unter der Kategorie des größeren Mittelwertes.

- Jede Kategorie des Mehrfachantworten-Sets wird durch einen Buchstaben (A, B, C, D, E) bezeichnet. Für jede Kategorie, bei der der Mittelwert einer anderen Kategorie sowohl niedriger als der Mittelwert der vorliegenden Kategorie ist als auch signifikant von diesem abweicht, wird der Buchstabe der Kategorie mit dem niedrigeren Mittelwert angezeigt.
- *Nachrichten aus Zeitungen* (A) weist das höchste mittlere Alter auf, und die Mittelwerte aller anderen Kategorien weichen signifikant von diesem Mittelwert ab.
- *Nachrichten aus dem Fernsehen* (C) weist das nächsthöchste mittlere Alter auf, und die Mittelwerte aller weiteren Kategorien (B, D und E) weichen signifikant von diesem Mittelwert ab. (Wie bereits gezeigt, weicht C außerdem signifikant von A ab.)
- Die mittleren Alter von *Nachrichten aus Zeitschriften* (B), *Nachrichten aus dem Radio* (D) und *Nachrichten aus dem Internet* (E) weichen nicht signifikant voneinander ab.

## ***Fehlende Werte (Define Variable)***

Viele Datendateien enthalten eine gewisse Menge an fehlenden Daten. Fehlende Daten können auf die verschiedensten Faktoren zurückgeführt werden. Möglicherweise beantworten Teilnehmer an einer Umfrage nicht jede Frage, bestimmte Variablen sind für einige Fälle nicht geeignet, und Codierfehler können dazu führen, dass einige Werte verworfen werden.

In IBM® SPSS® Statistics gibt es zwei Arten von fehlenden Werten:

- **Benutzerdefiniert fehlende Werte.** Werte, für die definiert wurde, dass sie fehlende Daten enthalten. Diesen Werten können Wertelabels zugewiesen werden, um anzuzeigen, warum die Daten fehlen (z. B. der Code 99 und das Wertelabel *Nicht zutreffend* für Schwangerschaft bei Männern).
- **Systemdefiniert fehlende Werte.** Wenn für eine numerische Variable kein Wert vorhanden ist, wird ihr der systemdefiniert fehlende Wert zugewiesen. Dies wird durch einen Punkt in der Datenansicht des Daten-Editors angegeben.

Es stehen mehrere Funktionen zur Verfügung, mit denen die Auswirkungen von fehlenden Daten kompensiert und sogar Muster in fehlenden Daten analysiert werden können. Dieses Kapitel hat jedoch ein weitaus einfacheres Ziel: Es wird erläutert, wie fehlende Daten in benutzerdefinierten Tabellen verarbeitet werden und welche Auswirkungen fehlende Daten auf die Berechnung von Auswertungsstatistiken haben.

### ***Beispiel-Datendatei***

Für die Beispiele in diesem Kapitel wird die Datendatei *missing\_values.sav* verwendet. [Für weitere Informationen siehe Thema Beispieldateien in Anhang A auf S. 201.](#) Hierbei handelt es sich um eine einfache, frei erfundene Datendatei mit lediglich einer Variablen und zehn Fällen, die nur dazu dienen soll, grundlegende Konzepte zu fehlenden Werten zu veranschaulichen.

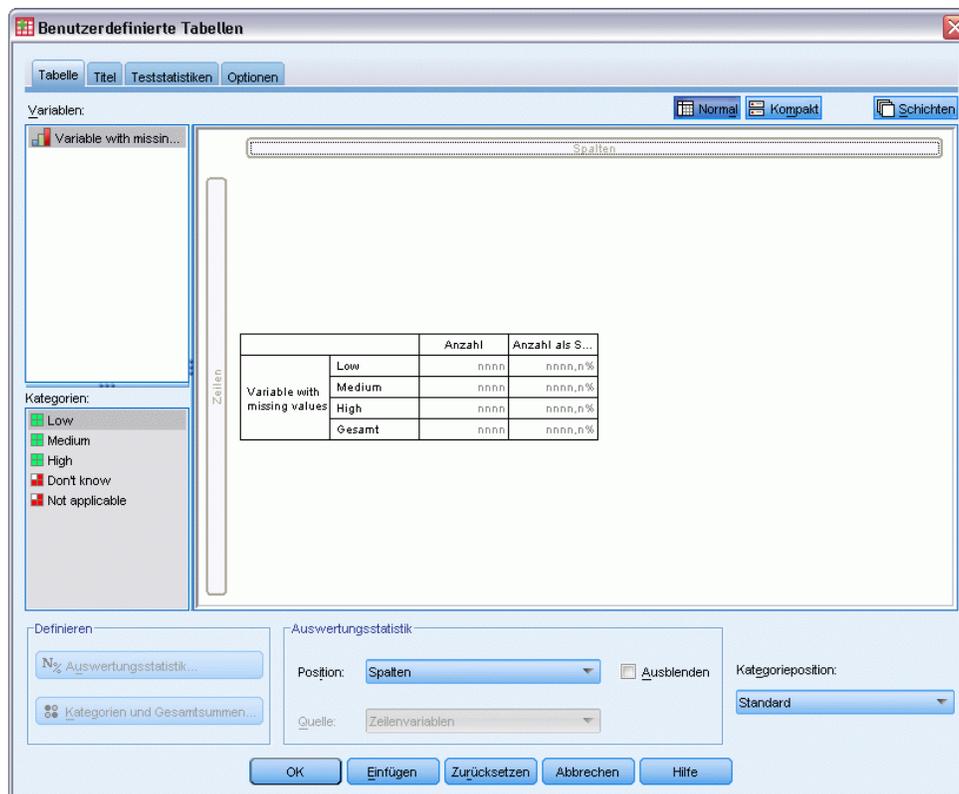
## ***Tabellen ohne fehlende Werte***

In der Standardeinstellung werden benutzerdefiniert fehlende Kategorien in benutzerdefinierten Tabellen nicht angezeigt (systemdefiniert fehlende Werte werden generell nicht angezeigt).

- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
Analysieren > Tabellen > Benutzerdefinierte Tabellen...
- ▶ Ziehen Sie im Dialogfeld zum Konstruieren von Tabellen die *Variable mit fehlenden Werten* (die einzige Variable in der Datei) aus der Variablenliste auf die Zeichenfläche, und legen Sie sie im Bereich für die Zeilen ab.
- ▶ Klicken Sie mit der rechten Maustaste auf der Zeichenfläche auf die Variable und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.

- ▶ Klicken Sie im Dialogfeld “Kategorien und Gesamtsummen” auf die Option Gesamt (aktivieren Sie diese) und klicken Sie anschließend auf Zuweisen.
- ▶ Klicken Sie erneut mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf die *Variable mit fehlenden Werten* und wählen Sie im Kontextmenü die Option Auswertungsstatistik aus.
- ▶ Wählen Sie im Dialogfeld “Auswertungsstatistik” in der Liste “Statistik” die Option Anzahl als Spalten% aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” zu übertragen.
- ▶ Klicken Sie auf Der Auswahl zuweisen.

Abbildung 12-1  
Tabellenvorschau ohne fehlende Werte



Möglicherweise werden Sie eine leichte Diskrepanz zwischen den in der Tabellenvorschau auf der Zeichenfläche angezeigten Kategorien und den in der Liste “Kategorien” (unterhalb der Variablenliste auf der linken Seite des Dialogfelds zum Konstruieren von Tabellen) feststellen. Die Liste “Kategorien” enthält eine Kategorie mit der Beschriftung *Fehlende Werte*, die in der Tabellenvorschau nicht enthalten ist, da Kategorien mit fehlenden Werten in der Standardeinstellung ausgeschlossen sind. Die Beschriftung enthält das Wort “Werte” im Plural. Dadurch wird angegeben, dass die Variable mindestens zwei benutzerdefiniert fehlende Kategorien aufweist.

- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 12-2  
Tabelle ohne fehlende Werte

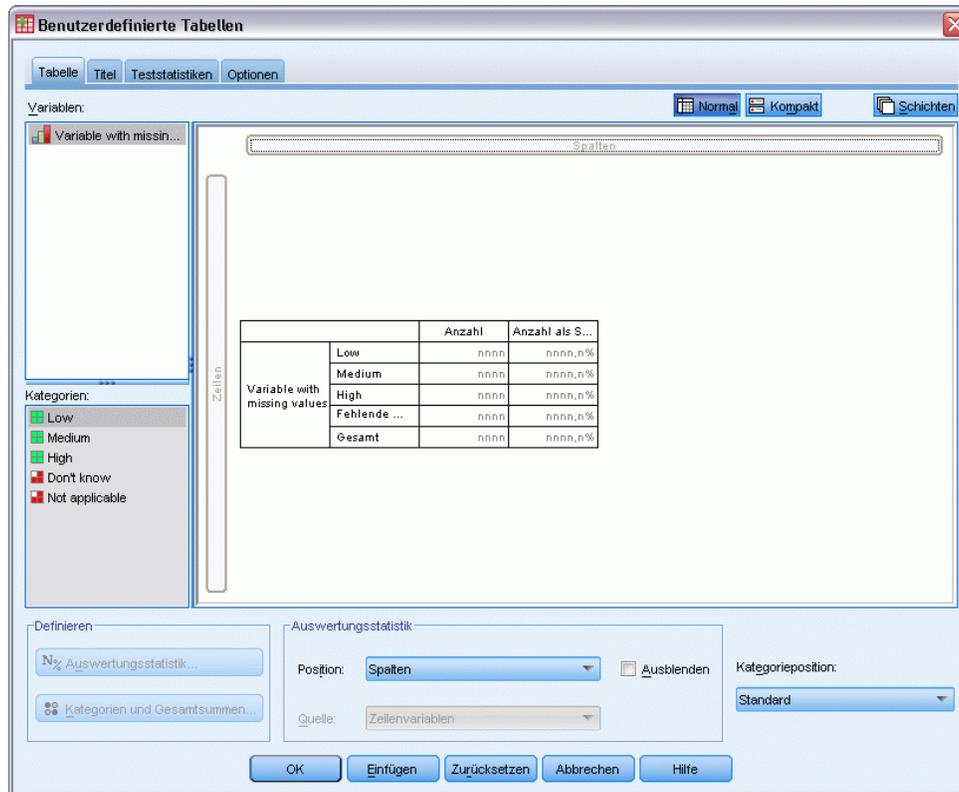
		Anzahl	Spalten%
Variable mit fehlenden Werten	Kleinster Wert	2	28,6%
	Mittel	3	42,9%
	Größter Wert	2	28,6%
	Gesamt	7	100,0%

Diese Tabelle ist anscheinend perfekt. Die Summen der Kategoriewerte ergeben die Gesamtsummen, und die Prozentwerte spiegeln genau die Werte wider, die Sie erhalten würden, wenn Sie die Gesamtanzahl als Grundlage für die Prozentwerte verwenden (z. B.  $3/7 = 0,429$  oder 42,9%). Die Gesamtanzahl entspricht jedoch nicht der Gesamtanzahl von Fällen in der Datendatei, sondern der Gesamtanzahl von Fällen mit **nichtfehlenden** Werten, also Fällen, die für diese Variable keine benutzerdefiniert fehlenden bzw. systemdefiniert fehlenden Werte aufweisen.

## ***Aufnehmen von fehlenden Werten in Tabellen***

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf die *Variable mit fehlenden Werten* und wählen Sie im Kontextmenü die Option Kategorien und Gesamtsummen aus.
- ▶ Klicken Sie im Dialogfeld “Kategorien und Gesamtsummen” auf die Option Fehlende Werte (aktivieren Sie diese) und klicken Sie anschließend auf Zuweisen.

Abbildung 12-3  
Tabellenvorschau mit angezeigter Kategorie für fehlende Werte



Die Tabelle enthält jetzt eine Kategorie für *Fehlende Werte*. Obwohl in der Tabellenvorschau lediglich eine Kategorie für fehlende Werte angezeigt wird, werden in der Tabelle sämtliche benutzerdefiniert fehlenden Kategorien angezeigt.

- ▶ Klicken Sie erneut mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf die *Variable mit fehlenden Werten* und wählen Sie im Kontextmenü die Option *Auswertungsstatistik* aus.
- ▶ Klicken Sie im Dialogfeld "Auswertungsstatistik" auf die Option *Angepaßte Auswertungsstatistik für Gesamtsummen und Zwischenergebnisse* (aktivieren Sie diese).
- ▶ Wählen Sie in der Liste "Statistik" für die angepaßte Auswertungsstatistik die Option *Gültige N* und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste "Anzeigen" zu übertragen.
- ▶ Wiederholen Sie diesen Vorgang für Gesamtanzahl.
- ▶ Klicken Sie auf *Der Auswahl zuweisen* und anschließend in der Tabellenkonstruktion auf *OK*, um die Tabelle zu erstellen.

Abbildung 12-4  
Tabelle mit fehlenden Werten

		Anzahl	Spalten%	Gültige N	Gesamtanzahl
Variable mit fehlenden Werten	Kleinsten Wert	2	22,2%		
	Mittel	3	33,3%		
	Größten Wert	2	22,2%		
	Weiß nicht	1	11,1%		
	Nicht zutreffend	1	11,1%		
	Gesamt	9	100,0%	7	10

Die beiden benutzerdefiniert fehlenden Kategorien *Weiß nicht* und *Nicht zutreffend* werden nun in der Tabelle angezeigt, wobei die Gesamtsumme jetzt 9 statt 7 beträgt. Dies spiegelt die beiden zusätzlichen Fälle mit benutzerdefiniert fehlenden Werten wider (einer pro benutzerdefiniert fehlender Kategorie). Die Prozentwerte in den Spalten wurden ebenfalls geändert, da sie auf der Anzahl der nichtfehlenden und benutzerdefiniert fehlenden Werte beruhen. Nur systemdefiniert fehlende Werte wurden nicht bei der Berechnung der Prozentwerte berücksichtigt.

*Gültige N* zeigt die Gesamtanzahl der nichtfehlenden Fälle (7) an, und *Gesamtanzahl* zeigt die Gesamtanzahl der Fälle einschließlich der benutzerdefiniert fehlenden und systemdefiniert fehlenden Fälle an. Die Gesamtanzahl von Fällen beträgt 10, also ein Fall mehr als die Anzahl von nichtfehlenden und benutzerdefiniert fehlenden Werten, die in der Spalte *Anzahl* als Gesamtanzahl angezeigt werden. Die Ursache dafür ist, dass ein Fall mit einem systemdefiniert fehlenden Wert vorhanden ist.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf die *Variable mit fehlenden Werten* und wählen Sie im Kontextmenü die Option *Auswertungsstatistik* aus.
- ▶ Wählen Sie in der oberen Liste “Statistik” (nicht die Liste im Gruppenfeld “Angepaßte Auswertungsstatistik für Gesamtsummen und Zwischenergebnisse”) die Option *Gültige N als Spalten%* aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” zu übertragen.
- ▶ Wiederholen Sie diesen Vorgang für *Gesamtanzahl als Spalten%*.
- ▶ Sie können diese Auswertungen auch in die entsprechende Liste im Gruppenfeld “Angepaßte Auswertungsstatistik für Gesamtsummen und Zwischenergebnisse” aufnehmen.
- ▶ Klicken Sie auf *Der Auswahl zuweisen* und anschließend auf *OK*, um die Tabelle zu erstellen.

Abbildung 12-5  
Tabelle mit fehlenden Werten, gültigen Prozentwerten und Gesamtprozentwerten

		Anzahl	Spalten%	Gültige N als Spalten%	Gesamtanzahl als Spalten%	Gültige N	Gesamtanzahl
MISSVAR Variable mit fehlenden Werten	Kleinsten Wert	2	22,2%	28,6%	20,0%		
	Mittel	3	33,3%	42,9%	30,0%		
	Größten Wert	2	22,2%	28,6%	20,0%		
	Weiß nicht	1	11,1%	,0%	10,0%		
	Nicht zutreffend	1	11,1%	,0%	10,0%		
	Gesamt	9	100,0%	100,0%	100,0%	7	10

- *Anzahl als Spalten%* gibt den Prozentwert in jeder Kategorie auf der Grundlage der Anzahl nichtfehlender und benutzerdefiniert fehlender Werte an (da benutzerdefiniert fehlende Werte ausdrücklich in die Tabelle aufgenommen wurden).
- *Gültige N als Spalten%* gibt den Prozentwert in jeder Kategorie ausschließlich auf der Grundlage der gültigen, nichtfehlenden Fälle an. Diese Werte entsprechen den Spaltenprozentwerten in der ursprünglichen Tabelle, in der keine benutzerdefiniert fehlenden Werte enthalten waren.
- *Gesamtanzahl als Spalten%* gibt den Prozentwert in jeder Kategorie auf der Grundlage aller Fälle einschließlich der Fälle mit benutzerdefiniert fehlenden und systemdefiniert fehlenden Werten an. Wenn Sie die einzelnen Kategorieprozentwerte in dieser Kategorie addieren, werden Sie feststellen, dass deren Summe lediglich 90% beträgt, da ein Fall der insgesamt 10 Fälle (10 %) den systemdefiniert fehlenden Wert aufweist. Obwohl dieser Fall in der Berechnungsgrundlage für die Prozentwerte berücksichtigt wurde, wird in der Tabelle keine Kategorie mit systemdefiniert fehlenden Werten bereitgestellt.

# Formatieren und Anpassen von Tabellen

Bei benutzerdefinierten Tabellen können Sie verschiedene Eigenschaften für die Tabellenformatierung bereits bei der Tabellenkonstruktion festlegen. Dazu gehören:

- Anzeigeformat und Beschriftungen für Auswertungsstatistiken
- Minimale und maximale Breite für Datenspalten
- Text oder Wert für die Anzeige in leeren Zellen

Diese Einstellungen werden in der Benutzeroberfläche für die Tabellenkonstruktion beibehalten, bis Sie sie ändern, die Einstellungen für die Tabellenkonstruktion zurücksetzen oder eine andere Datendatei öffnen. Auf diese Weise können Sie mehrere Tabellen mit übereinstimmender Formatierung erstellen, ohne sie nach dem Erstellen manuell bearbeiten zu müssen. Sie können diese Formatierungseinstellungen auch zusammen mit allen anderen Tabellenparametern speichern, indem Sie die Befehlssyntax mit der Schaltfläche “Einfügen” in der Benutzeroberfläche für die Tabellenkonstruktion in ein Syntax-Fenster einfügen. Anschließend können Sie die Syntax in einer Datei speichern.

Viele Formatierungseigenschaften von Tabellen können auch nach dem Erstellen geändert werden. Hierzu können Sie alle im Viewer für Pivot-Tabellen verfügbaren Formatierungsfunktionen verwenden. In diesem Kapitel wird jedoch hauptsächlich das Festlegen der Tabellenformatierung vor dem Erstellen der Tabelle behandelt. Weitere Informationen zu Pivot-Tabellen finden Sie im Hilfesystem. Geben Sie als Stichwort auf der Registerkarte “Index” Pivot-Tabellen ein.

## **Beispiel-Datendatei**

Für die Beispiele dieses Kapitels wird die Datendatei *survey\_sample.sav* verwendet. [Für weitere Informationen siehe Thema Beispieldateien in Anhang A auf S. 201.](#)

In allen angegebenen Beispielen werden die Variablenlabels in Dialogfeldern in alphabetischer Reihenfolge angezeigt. Die Anzeigeeigenschaften für Variablenlisten werden im Dialogfeld “Optionen” auf der Registerkarte “Allgemein” festgelegt (Menü “Bearbeiten”, Befehl “Optionen”).

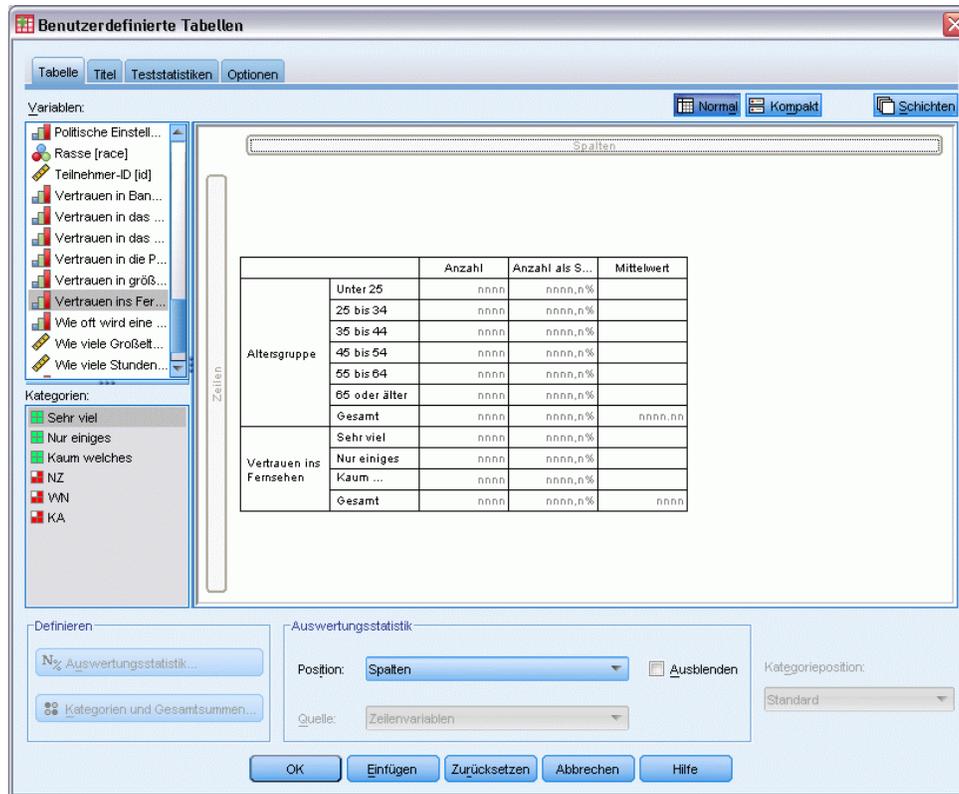
## **Anzeigeformat für Auswertungsstatistiken**

In den benutzerdefinierten Tabellen wird versucht, den Auswertungsstatistiken möglichst geeignete Standardformate zuzuweisen. In manchen Fällen ist es jedoch sinnvoll, diese Formate durch andere zu ersetzen.

- ▶ Wählen Sie die folgenden Befehle aus den Menüs aus:  
Analysieren > Tabellen > Benutzerdefinierte Tabellen...

- ▶ Ziehen Sie die Variable *Alterskategorie* aus der Variablenliste der Tabellenkonstruktion in den Bereich “Zeilen” auf der Zeichenfläche.
- ▶ Ziehen Sie die Variable *Vertrauen in das Fernsehen* unter *Alterskategorie* im Zeilenbereich, und legen Sie sie dort ab. Dadurch werden die beiden Variablen in der Zeilendimension gestapelt.
- ▶ Klicken Sie mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf *Alterskategorie* und wählen Sie im Kontextmenü die Option *Alle Zeilenvariablen auswählen* aus.
- ▶ Klicken Sie mit der rechten Maustaste erneut auf *Alterskategorie* und wählen Sie im Kontextmenü die Option *Kategorien und Gesamtsummen* aus.
- ▶ Wählen Sie im Dialogfeld “Kategorien und Gesamtsummen” die Option *Gesamt* aus (aktivieren Sie diese) und klicken Sie anschließend auf *Zuweisen*.
- ▶ Klicken Sie mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf eine der Variablen und wählen Sie im Kontextmenü die Option *Auswertungsstatistik* aus.
- ▶ Wählen Sie in der Liste “Statistik” die Option *Anzahl als Spalten%* aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” zu übertragen.
- ▶ Wählen Sie die Option *Angepaßte Auswertungsstatistik für Gesamtsummen und Zwischenergebnisse* aus (aktivieren Sie sie).
- ▶ Wählen Sie in der Liste “Statistik” für die angepaßte Auswertungsstatistik die Option *Anzahl als Spalten%* aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” zu übertragen.
- ▶ Wiederholen Sie diesen Vorgang für *Mittelwert*.
- ▶ Klicken Sie dann auf *Allen zuweisen*.

Abbildung 13-1  
Standard-Anzeigeformate in der Tabellenvorschau

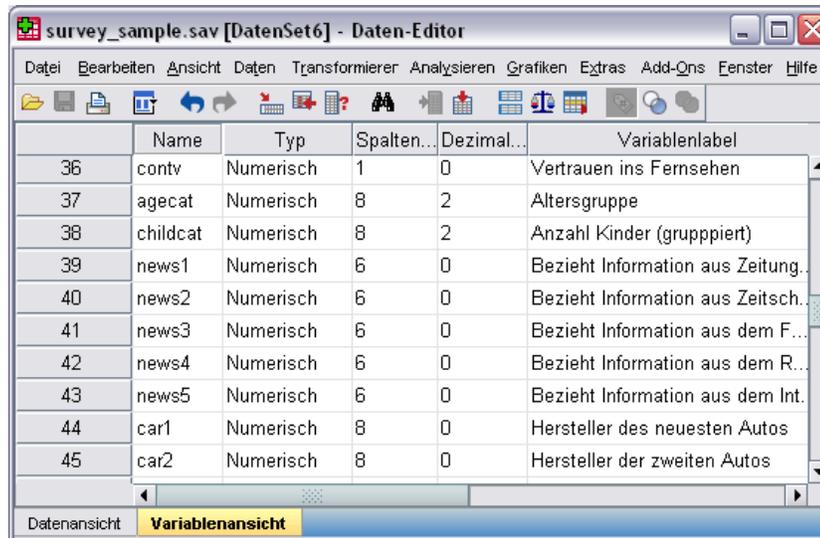


Die Platzhalterwerte in der Tabellenvorschau geben das Standardformat für jede Auswertungsstatistik wieder.

- Für Häufigkeiten lautet das Standard-Anzeigeformat nnnn, also Ganzzahlwerte ohne Dezimalstellen.
- Für Prozentwerte lautet das Standard-Anzeigeformat nnnn.n%, also Zahlen mit einer einzelnen Dezimalstelle und einem Prozentzeichen nach dem Wert.
- Für den Mittelwert ist das Standard-Anzeigeformat für die beiden Variablen *unterschiedlich*.

Bei Auswertungsstatistiken, die keine Form von Häufigkeit (z. B. Gültige N und Gesamtanzahl) oder Prozentwert darstellen, wird als Standard-Anzeigeformat das Anzeigeformat verwendet, das für die Variable im Daten-Editor definiert wurde. Wenn Sie die Variablen in der Variablenansicht des Daten-Editors betrachten, werden Sie feststellen, dass *Alterskategorie* (Variable *alt\_kat*) mit zwei Dezimalstellen definiert ist, während *Vertrauen in das Fernsehen* (Variable *ver\_tv*) ohne (d. h. mit null) Dezimalstellen definiert ist.

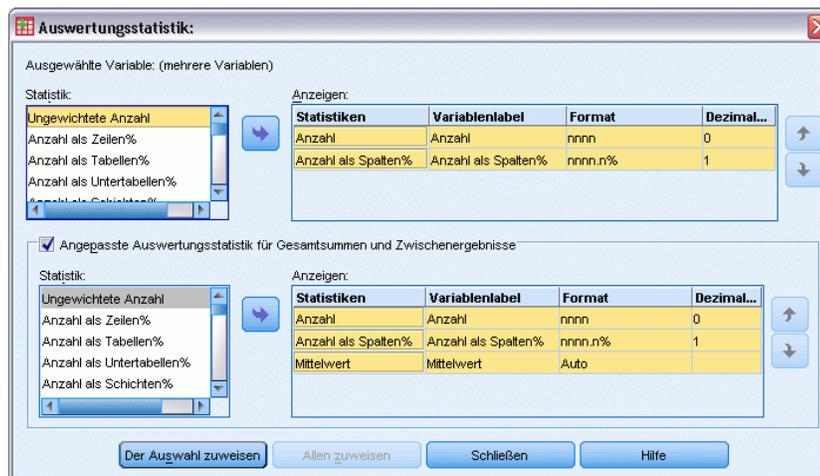
Abbildung 13-2  
Variablenansicht im Daten-Editor



Dies ist wahrscheinlich einer der Fälle, in denen das Standardformat nicht erwünscht ist, da beide Werte möglichst mit derselben Anzahl von Dezimalstellen angezeigt werden sollen.

- Klicken Sie mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf eine der Variablen und wählen Sie im Kontextmenü die Option *Auswertungsstatistik* aus.

Abbildung 13-3  
Dialogfeld "Auswertungsstatistik"



Für den Mittelwert wird in der Zelle "Format" der Liste "Anzeigen" das Format *Auto* gezeigt. Dies bedeutet, dass das für die Variable definierte Anzeigeformat verwendet wird und die Zelle "Dezimalstellen" deaktiviert ist. Um die Anzahl der Dezimalstellen festzulegen, müssen Sie zunächst ein anderes Format auswählen.

- Klicken Sie in der Liste "Anzeigen" für die angepasste Auswertungsstatistik auf die Zelle "Format" für den Mittelwert und wählen Sie in der Dropdown-Liste der Formate die Option *nnnn* aus.

- ▶ Geben Sie in der Zelle “Dezimalstellen” den Wert 1 ein.
- ▶ Klicken Sie dann auf Allen zuweisen, um diese Einstellung beiden Variablen zu zuweisen.

Abbildung 13-4

Tabellenvorschau mit benutzerdefinierten Anzeigeformaten für Auswertungsstatistiken

The screenshot shows the 'Benutzerdefinierte Tabellen' dialog box. The 'Auswertungsstatistik' section is set to 'Spalten' and 'Zellenvariablen'. The preview table shows the following data:

Altersgruppe	Anzahl	Anzahl als S...	Mittelwert
Unter 25	nnnn	nnnn,n%	nnnn,n
25 bis 34	nnnn	nnnn,n%	nnnn,n
35 bis 44	nnnn	nnnn,n%	nnnn,n
45 bis 54	nnnn	nnnn,n%	nnnn,n
55 bis 64	nnnn	nnnn,n%	nnnn,n
65 oder älter	nnnn	nnnn,n%	nnnn,n
Gesamt	nnnn	nnnn,n%	nnnn,n
Vertrauen ins Fernsehen	Sehr viel	nnnn	nnnn,n%
	Nur einiges	nnnn	nnnn,n%
	Kaum ...	nnnn	nnnn,n%
	Gesamt	nnnn	nnnn,n

Aus der Tabellenvorschau ist nun ersichtlich, dass beide Mittelwerte mit einer Dezimalstelle angezeigt werden. (Sie könnten diese Tabelle nun unmittelbar erstellen. Der “Mittelwert” für *Alterskategorie* ist jedoch möglicherweise etwas schwierig zu interpretieren, da die eigentlichen numerischen Codes für diese Variable nur im Bereich von 1 bis 6 liegen.)

## Anzeigen von Beschriftungen für Auswertungsstatistiken

Neben den Anzeigeformaten für Auswertungsstatistiken können Sie auch die Beschriftungen für jede Auswertungsstatistik festlegen.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie auf Zurücksetzen, um frühere Einstellungen in der Tabellenkonstruktion zu löschen.
- ▶ Ziehen Sie die Variable *Alterskategorie* aus der Variablenliste der Tabellenkonstruktion in den Bereich “Zeilen” auf der Zeichenfläche.

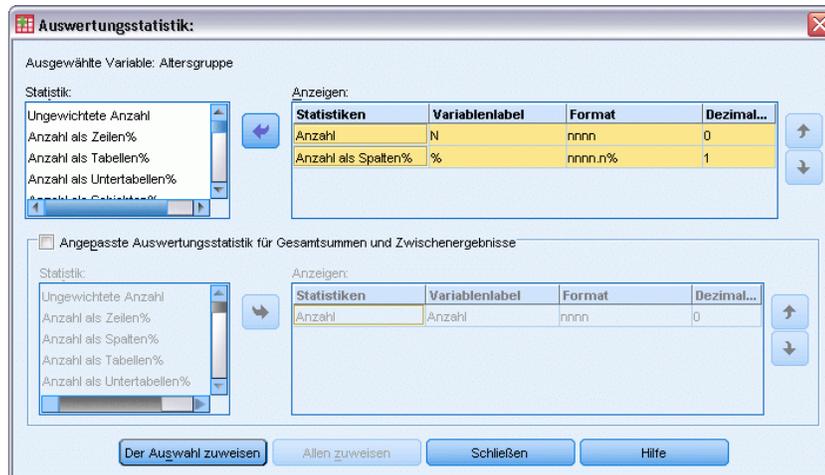
- ▶ Ziehen Sie die Variable *Zahlungsart letzte Woche* aus der Variablenliste auf die Zeichenfläche, und legen Sie sie im Spaltenbereich ab.
- ▶ Klicken Sie mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf die Variable *Alterskategorie* und wählen Sie im Kontextmenü die Option *Auswertungsstatistik* aus.
- ▶ Wählen Sie in der Liste "Statistik" die Option *Anzahl als Spalten%* aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste "Anzeigen" zu übertragen.
- ▶ Doppelklicken Sie in der Liste "Anzeigen" in der Zelle "Label" auf das Wort *Spalte*, um den Inhalt der Zelle zu bearbeiten. Löschen Sie das Wort *Spalte* aus der Beschriftung, und ersetzen Sie es durch *%*.
- ▶ Bearbeiten Sie die Zelle "Label" für *Anzahl* auf entsprechende Weise, und ersetzen Sie den Eintrag durch *N*.

Ändern Sie bei dieser Gelegenheit das Format der Statistik "Anzahl als Spalten%", um das nun überflüssige Prozentzeichen zu entfernen. (Die Spaltenbeschriftung gibt jetzt an, dass die Spalte Prozentwerte enthält.)

- ▶ Klicken Sie auf die Zelle "Format" für *Anzahl als Spalten%* und wählen Sie in der Dropdown-Liste der Formate die Option *nnnn.n*.

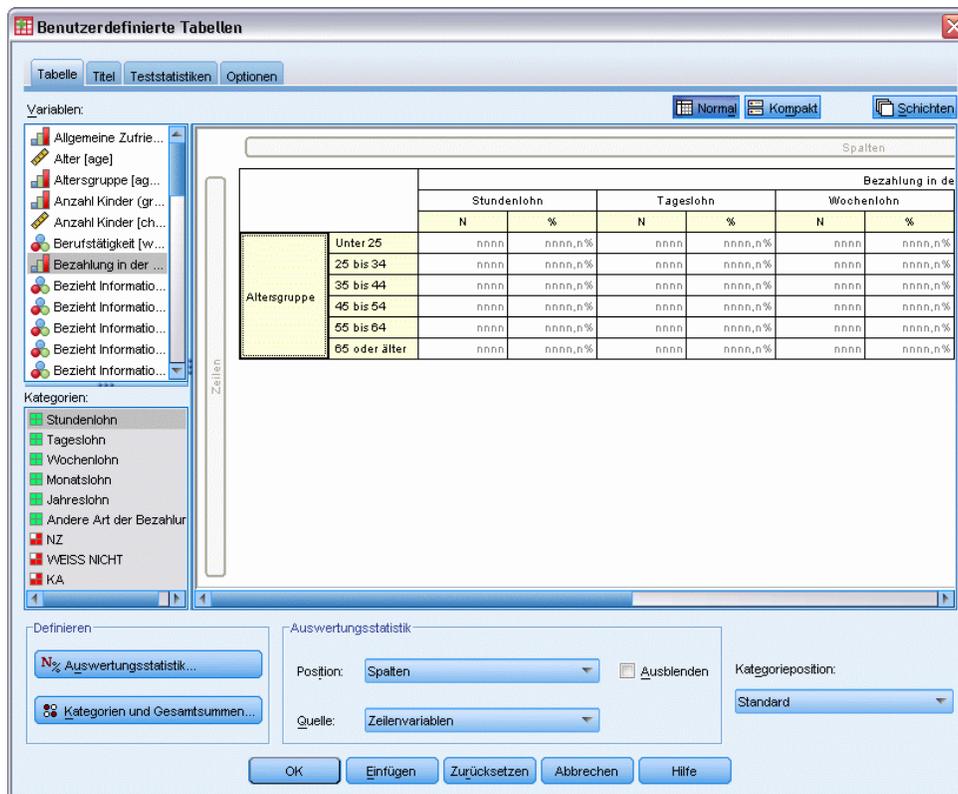
Abbildung 13-5

Dialogfeld "Auswertungsstatistik" mit geänderten Beschriftungen und Formaten



- ▶ Klicken Sie dann auf *Der Auswahl zuweisen*.

Abbildung 13-6  
Tabellenvorschau mit geänderten Beschriftungen für Auswertungsstatistiken



In der Tabellenvorschau werden das geänderte Anzeigeformat und die geänderten Beschriftungen angezeigt.

- Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 13-7  
Tabelle mit geänderten Beschriftungen für Auswertungsstatistiken

		Zahlungsart letzte Woche											
		Stundenlohn		Tagespauschale		Wochenlohn		Monatsgehalt		Jahresgehalt		Andere Zahlungsart	
		N	%	N	%	N	%	N	%	N	%	N	%
Alterskategorie	Weniger als 25	91	14,0	0	,0	12	9,7	3	2,0	7	3,1	14	7,7
	25 bis 34	175	26,9	5	29,4	33	26,6	37	24,8	63	28,0	31	17,1
	35 bis 44	185	28,5	5	29,4	42	33,9	45	30,2	66	29,3	61	33,7
	45 bis 54	124	19,1	5	29,4	25	20,2	38	25,5	58	25,8	41	22,7
	55 bis 64	52	8,0	0	,0	10	8,1	23	15,4	29	12,9	19	10,5
	65 oder älter	23	3,5	2	11,8	2	1,6	3	2,0	2	,9	15	8,3

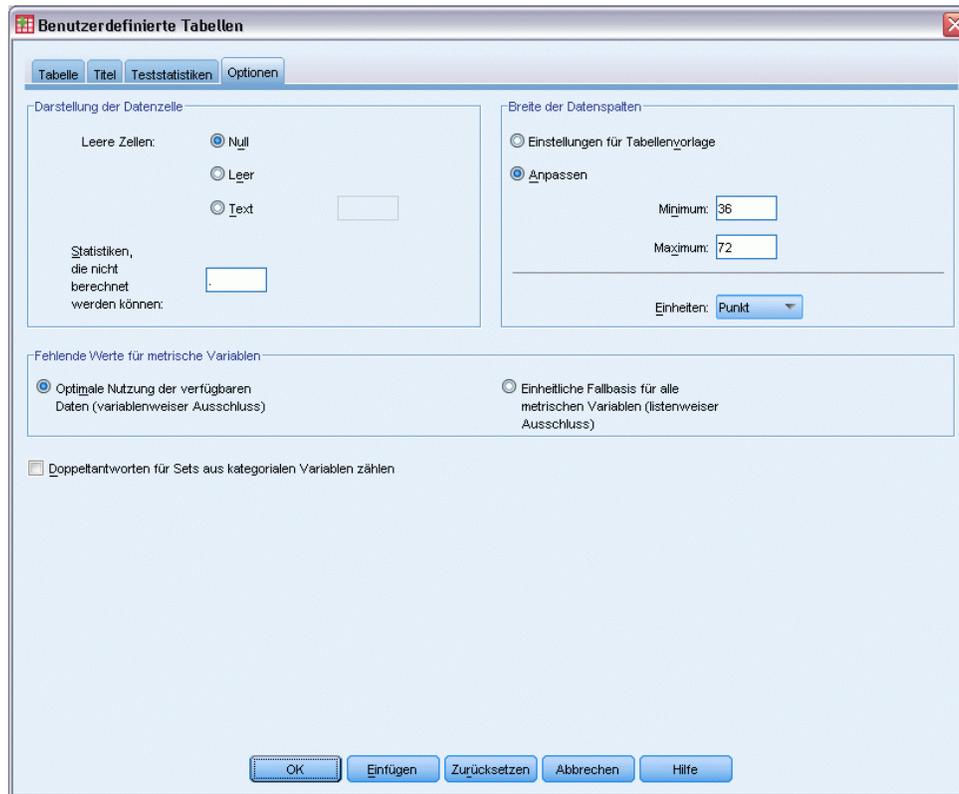
## Spaltenbreite

Sie werden bemerkt haben, dass die Tabelle im vorangegangenen Beispiel recht breit ist. Eine Lösung für dieses Problem besteht darin, einfach die Zeilen- und Spaltenvariablen zu tauschen. Alternativ dazu können die Spaltenbreiten reduziert werden, da sie wesentlich breiter als notwendig erscheinen. (Die Beschriftungen für die Auswertungsstatistik wurden nämlich gekürzt, damit die Spalten schmaler gemacht werden können.)

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie auf die Registerkarte Optionen.
- ▶ Wählen Sie im Gruppenfeld “Breite der Datenspalten” die Option Anpassen.
- ▶ Geben Sie im Feld “Maximum” den Wert 36 ein. (Stellen Sie sicher, dass im Feld “Einheiten” die Option Punkt ausgewählt ist.)

Abbildung 13-8

Dialogfeld “Benutzerdefinierte Tabellen”, Registerkarte “Optionen”



- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 13-9

Tabelle mit reduzierten Spaltenbreiten

		Zahlungsart letzte Woche											
		Stundenlohn		Tagespauschale		Wochenlohn		Monatsgehalt		Jahresgehalt		Andere Zahlungsart	
		N	%	N	%	N	%	N	%	N	%	N	%
Alterskategorie	Weniger als 25	91	14,0	0	,0	12	9,7	3	2,0	7	3,1	14	7,7
	25 bis 34	175	26,9	5	29,4	33	26,6	37	24,8	63	28,0	31	17,1
	35 bis 44	185	28,5	5	29,4	42	33,9	45	30,2	66	29,3	61	33,7
	45 bis 54	124	19,1	5	29,4	25	20,2	38	25,5	58	25,8	41	22,7
	55 bis 64	52	8,0	0	,0	10	8,1	23	15,4	29	12,9	19	10,5
	65 oder älter	23	3,5	2	11,8	2	1,6	3	2,0	2	,9	15	8,3

Die Tabelle ist nun wesentlich übersichtlicher.

## Anzeigewert für leere Zellen

In der Standardeinstellung wird in leeren Zellen (Zellen, die keine Fälle enthalten) eine 0 angezeigt. In diesen Zellen kann jedoch auch nichts (d. h., die Zellen werden leer gelassen) oder ein bestimmter Text angezeigt werden.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Klicken Sie auf die Registerkarte Optionen.
- ▶ Wählen Sie im Gruppenfeld “Darstellung der Datenzelle” neben “Leere Zellen” die Option Text und geben Sie Kein ein.
- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 13-10

Tabelle, bei der in leeren Zellen “Kein” angezeigt wird

		Zahlungsart letzte Woche											
		Stundenlohn		Tagespauschale		Wochenlohn		Monatsgehalt		Jahresgehalt		Andere Zahlungsart	
		N	%	N	%	N	%	N	%	N	%	N	%
Alterskategorie	Weniger als 25	91	14,0	Kein	Kein	12	9,7	3	2,0	7	3,1	14	7,7
	25 bis 34	175	26,9	5	29,4	33	26,6	37	24,8	63	28,0	31	17,1
	35 bis 44	185	28,5	5	29,4	42	33,9	45	30,2	66	29,3	61	33,7
	45 bis 54	124	19,1	5	29,4	25	20,2	38	25,5	58	25,8	41	22,7
	55 bis 64	52	8,0	Kein	Kein	10	8,1	23	15,4	29	12,9	19	10,5
	65 oder älter	23	3,5	2	11,8	2	1,6	3	2,0		,9	15	8,3

In den vier leeren Zellen der Tabelle wird nun statt des Werts 0 der Text *Kein* angezeigt.

## Anzeigewert für fehlende Statistiken

Wenn eine Statistik nicht berechnet werden kann, wird in der Standardeinstellung ein Punkt (.) angezeigt. Dies ist das Symbol, mit dem der systemdefinierte fehlende Werte angegeben wird. Dieser Wert unterscheidet sich von einer “leeren” Zelle, und daher wird der Anzeigewert für fehlende Statistiken getrennt vom Anzeigewert für Zellen ohne Fälle festgelegt.

- ▶ Öffnen Sie das Dialogfeld zum Konstruieren von Tabellen (Menü “Analysieren”, Untermenü “Tabellen”, Befehl “Benutzerdefinierte Tabellen”).
- ▶ Ziehen Sie die Variable *Täglicher Fernsehkonsum in Stunden* aus der Variablenliste auf die Zeichenfläche, und legen Sie am Anfang des Spaltenbereichs über *Zahlungsart letzte Woche* ab.

Da *Täglicher Fernsehkonsum in Stunden* eine metrische Variable ist, wird sie automatisch als Quellvariable der Statistik verwendet, und die Auswertungsstatistik wird auf den Mittelwert gesetzt.

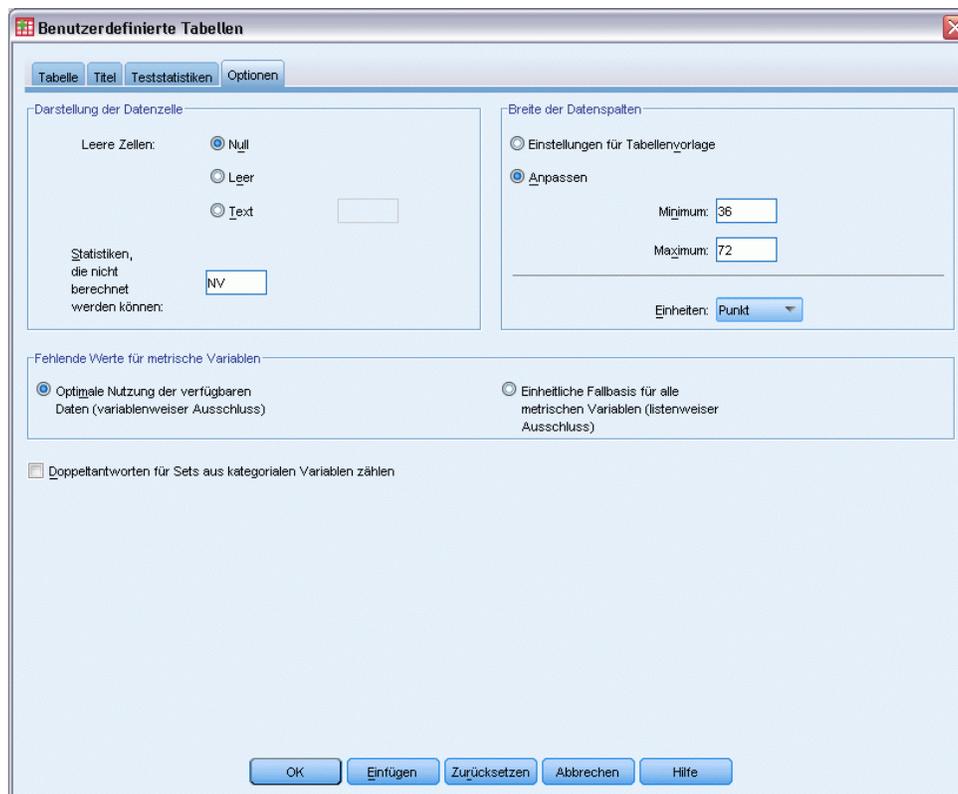
- ▶ Klicken Sie mit der rechten Maustaste in der Tabellenvorschau auf der Zeichenfläche auf die Variable *Täglicher Fernsehkonsum in Stunden* und wählen Sie im Kontextmenü die Option Auswertungsstatistik aus.
- ▶ Wählen Sie in der Liste “Statistik” die Option Gültige N aus und klicken Sie auf die Pfeilschaltfläche, um sie in die Liste “Anzeigen” zu übertragen.

Abbildung 13-11  
Dialogfeld "Auswertungsstatistik" für metrische Variablen



- ▶ Klicken Sie auf Der Auswahl zuweisen.
- ▶ Klicken Sie auf die Registerkarte Optionen.
- ▶ Geben Sie im Textfeld "Statistiken, die nicht berechnet werden können" den Text NV ein.

Abbildung 13-12  
Ändern des Anzeigewerts für Statistiken, die nicht berechnet werden können



- ▶ Klicken Sie auf OK, um die Tabelle zu erstellen.

Abbildung 13-13  
Tabelle, bei der für fehlende Statistiken "NV" angezeigt wird

		Wie viele Stunden Fernsehen pro Tag											
		Bezahlung in der vorigen Woche											
		Stundenlohn		Tageslohn		Wochenlohn		Monatslohn		Jahreslohn		Andere Art der Bezahlung	
		Mittelwert	Gültige N	Mittelwert	Gültige N	Mittelwert	Gültige N	Mittelwert	Gültige N	Mittelwert	Gültige N	Mittelwert	Gültige N
Altersgruppe	Unter 25	3	71	NV	Kein	2	10	2	3	2	6	2	8
	25 bis 34	3	134	4	2	2	30	2	29	2	52	2	22
	35 bis 44	3	136	2	5	3	30	2	34	2	47	3	46
	45 bis 54	2	90	2	4	2	22	2	36	2	45	2	34
	55 bis 64	3	40	NV	Kein	3	7	2	15	2	23	3	15
	65 oder älter	3	18	2	2	1	1	NV	0	1	2	3	11

Der Text *NV* wird in drei Zellen der Tabelle als Mittelwert angezeigt. Der entsprechende Wert für *Gültige N* liefert jeweils die Erklärung hierfür: Es sind keine Fälle vorhanden, anhand deren der Mittelwert berechnet werden kann.

Sie werden jedoch eine leichte Diskrepanz bemerken: Einer der drei Werte für "Gültige N" wird nicht, wie für Zellen ohne Fälle anzunehmen, mit der Beschriftung *Kein*, sondern als 0 angezeigt. Die Ursache hierfür ist, dass zwar keine gültigen Fälle zum Berechnen des Mittelwerts vorhanden sind, die Kategorie jedoch nicht völlig leer ist. Wenn Sie nochmals die ursprüngliche Tabelle mit nur zwei kategorialen Variablen betrachten, werden Sie erkennen, dass in dieser Kategorie eigentlich drei Fälle vorhanden sind. Es sind jedoch keine gültigen Fälle vorhanden, da alle drei Fälle fehlende Werte für die metrische Variable *Täglicher Fernsehkonsum in Stunden* aufweisen.

# Beispieldateien

Die zusammen mit dem Produkt installierten Beispieldateien finden Sie im Unterverzeichnis *Samples* des Installationsverzeichnisses. Für jeder der folgenden Sprachen gibt es einen eigenen Ordner innerhalb des Unterverzeichnisses "Samples": Englisch, Französisch, Deutsch, Italienisch, Japanisch, Koreanisch, Polnisch, Russisch, Vereinfachtes Chinesisch, Spanisch und Traditionelles Chinesisch.

Nicht alle Beispieldateien stehen in allen Sprachen zur Verfügung. Wenn eine Beispieldatei nicht in einer Sprache zur Verfügung steht, enthält der jeweilige Sprachordner eine englische Version der Beispieldatei.

## **Beschreibungen**

Im Folgenden finden Sie Kurzbeschreibungen der in den verschiedenen Beispielen in der Dokumentation verwendeten Beispieldateien.

- **accidents.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um eine Versicherungsgesellschaft geht, die alters- und geschlechtsabhängige Risikofaktoren für Autounfälle in einer bestimmten Region untersucht. Jeder Fall entspricht einer Kreuzklassifikation von Alterskategorie und Geschlecht.
- **adl.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um Bemühungen geht, die Vorteile einer vorgeschlagenen Therapieform für Schlaganfallpatienten zu ermitteln. Ärzte teilten weibliche Schlaganfallpatienten nach dem Zufallsprinzip jeweils einer von zwei Gruppen zu. Die erste Gruppe erhielt die physische Standardtherapie, die zweite erhielt eine zusätzliche Emotionaltherapie. Drei Monate nach den Behandlungen wurden die Fähigkeiten der einzelnen Patienten, übliche Alltagsaktivitäten auszuführen, als ordinale Variablen bewertet.
- **advert.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen eines Einzelhändlers geht, die Beziehungen zwischen den in Werbung investierten Beträgen und den daraus resultierenden Umsätzen zu untersuchen. Zu diesem Zweck hat er die Umsätze vergangener Jahre und die zugehörigen Werbeausgaben zusammengestellt.
- **aflatoxin.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um Tests von Maisernten auf Aflatoxin geht, ein Gift, dessen Konzentration stark zwischen und innerhalb von Ernteerträgen schwankt. Ein Kornverarbeitungsbetrieb hat aus 8 Ernteerträgen je 16 Proben erhalten und das Aflatoxinniveau in Teilen pro Milliarde (parts per billion, PPB) gemessen.
- **anorectic.sav.** Bei der Ausarbeitung einer standardisierten Symptomatologie anorektischen/bulimischen Verhaltens führten Forscher ) eine Studie mit 55 Jugendlichen mit bekannten Ess-Störungen durch. Jeder Patient wurde vier Mal über einen Zeitraum von vier Jahren untersucht, es fanden also insgesamt 220 Beobachtungen statt. Bei jeder Beobachtung erhielten die Patienten Scores für jedes von 16 Symptomen. Die Symptomwerte fehlen für

Patient 71 zum Zeitpunkt 2, Patient 76 zum Zeitpunkt 2 und Patient 47 zum Zeitpunkt 3, wodurch 217 gültige Beobachtungen verbleiben.

- **bankloan.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen einer Bank geht, den Anteil der nicht zurückgezahlten Kredite zu reduzieren. Die Datei enthält Informationen zum Finanzstatus und demografischen Hintergrund von 850 früheren und potenziellen Kunden. Bei den ersten 700 Fällen handelt es sich um Kunden, denen bereits ein Kredit gewährt wurde. Bei den letzten 150 Fällen handelt es sich um potenzielle Kunden, deren Kreditrisiko die Bank als gering oder hoch einstufen möchte.
- **bankloan\_binning.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die Informationen zum Finanzstatus und demografischen Hintergrund von 5.000 früheren Kunden enthält.
- **behavior.sav.** In einem klassischen Beispiel () wurden 52 Schüler/Studenten gebeten, die Kombinationen aus 15 Situationen und 15 Verhaltensweisen auf einer 10-Punkte-Skala von 0 = “ausgesprochen angemessen” bis 9 = “ausgesprochen unangemessen” zu bewerten. Die Werte werden über die einzelnen Personen gemittelt und als Unähnlichkeiten verwendet.
- **behavior\_ini.sav.** Diese Datendatei enthält eine Ausgangskonfiguration für eine zweidimensionale Lösung für *behavior.sav*.
- **brakes.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Qualitätskontrolle in einer Fabrik geht, die Scheibenbremsen für Hochleistungsautomobile herstellt. Die Datendatei enthält Messungen des Durchmessers von 16 Scheiben aus 8 Produktionsmaschinen. Der Zieldurchmesser für die Scheiben ist 322 Millimeter.
- **breakfast.sav.** In einer klassischen Studie () wurden 21 MBA-Studenten der Wharton School mit ihren Lebensgefährten darum gebeten, 15 Frühstücksartikel in der Vorzugsreihenfolge von 1 = “am meisten bevorzugt” bis 15 = “am wenigsten bevorzugt” zu ordnen. Die Bevorzugungen wurden in sechs unterschiedlichen Szenarien erfasst, von “Overall preference” (Allgemein bevorzugt) bis “Snack, with beverage only” (Imbiss, nur mit Getränk).
- **breakfast-overall.sav.** Diese Datei enthält die Daten zu den bevorzugten Frühstücksartikeln, allerdings nur für das erste Szenario, “Overall preference” (Allgemein bevorzugt).
- **broadband\_1.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die die Anzahl der Abonnenten eines Breitband-Service, nach Region geordnet, enthält. Die Datendatei enthält die monatlichen Abonentenzahlen für 85 Regionen über einen Zeitraum von vier Jahren.
- **broadband\_2.sav** Diese Datendatei stimmt mit *broadband\_1.sav* überein, enthält jedoch Daten für weitere drei Monate.
- **car\_insurance\_claims.sav.** Ein an anderer Stelle () vorgestelltes und analysiertes Daten-Set bezieht sich auf Schadensansprüche für Autos. Die durchschnittliche Höhe der Schadensansprüche lässt sich mit Gamma-Verteilung modellieren. Dazu wird eine inverse Verknüpfungsfunktion verwendet, um den Mittelwert der abhängigen Variablen mit einer linearen Kombination aus Alter des Versicherungsnehmers, Fahrzeugtyp und Fahrzeugalter in Bezug zu setzen. Die Anzahl der eingereichten Schadensansprüche kann als Skalierungsgewicht verwendet werden.
- **car\_sales.sav.** Diese Datendatei enthält hypothetische Verkaufsschätzer, Listenpreise und physische Spezifikationen für verschiedene Fahrzeugfabrikate und -modelle. Die Listenpreise und physischen Spezifikationen wurden von *edmunds.com* und Hersteller-Websites entnommen.

- **car\_sales\_uprepared.sav.** Hierbei handelt es sich um eine modifizierte Version der Datei *car\_sales.sav*, die keinerlei transformierte Versionen der Felder enthält.
- **carpet.sav** In einem beliebigen Beispiel möchte einen neuen Teppichreiniger vermarkten und dazu den Einfluss von fünf Faktoren auf die Bevorzugung durch den Verbraucher untersuchen: Verpackungsgestaltung, Markenname, Preis, Gütesiegel, *Good Housekeeping* und Geld-zurück-Garantie. Die Verpackungsgestaltung setzt sich aus drei Faktorenebenen zusammen, die sich durch die Position der Auftragebürste unterscheiden. Außerdem gibt es drei Markennamen (*K2R*, *Glory* und *Bissell*), drei Preisstufen sowie je zwei Ebenen (Nein oder Ja) für die letzten beiden Faktoren. 10 Kunden stufen 22 Profile ein, die durch diese Faktoren definiert sind. Die Variable *Preference* enthält den Rang der durchschnittlichen Einstufung für die verschiedenen Profile. Ein niedriger Rang bedeutet eine starke Bevorzugung. Diese Variable gibt ein Gesamtmaß der Bevorzugung für die Profile an.
- **carpet\_prefs.sav.** Diese Datendatei beruht auf denselben Beispielen, wie für *carpet.sav* beschrieben, enthält jedoch die tatsächlichen Einstufungen durch jeden der 10 Kunden. Die Kunden wurden gebeten, die 22 Produktprofile in der Reihenfolge ihrer Präferenzen einzustufen. Die Variablen *PREF1* bis *PREF22* enthalten die IDs der zugeordneten Profile, wie in *carpet\_plan.sav* definiert.
- **catalog.sav.** Diese Datendatei enthält hypothetische monatliche Verkaufszahlen für drei Produkte, die von einem Versandhaus verkauft werden. Daten für fünf mögliche Einflussvariablen wurden ebenfalls aufgenommen.
- **catalog\_seasfac.sav.** Diese Datendatei ist mit *catalog.sav* identisch, außer, dass ein Set von saisonalen Faktoren, die mithilfe der Prozedur "Saisonale Zerlegung" berechnet wurden, sowie die zugehörigen Datumsvariablen hinzugefügt wurden.
- **cellular.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen eines Mobiltelefonunternehmens geht, die Kundenabwanderung zu verringern. Scores für die Abwanderungsneigung (von 0 bis 100) werden auf die Kunden angewendet. Kunden mit einem Score von 50 oder höher streben vermutlich einen Anbieterwechsel an.
- **ceramics.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen eines Herstellers geht, der ermitteln möchte, ob ein neue, hochwertige Keramiklegierung eine größere Hitzebeständigkeit aufweist als eine Standardlegierung. Jeder Fall entspricht einem Test einer der Legierungen; die Temperatur, bei der das Keramikwälzlager versagte, wurde erfasst.
- **cereal.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um eine Umfrage geht, bei der 880 Personen nach ihren Frühstücksgewohnheiten befragt wurden. Außerdem wurden Alter, Geschlecht, Familienstand und Vorliegen bzw. Nichtvorliegen eines aktiven Lebensstils (auf der Grundlage von mindestens zwei Trainingseinheiten pro Woche) erfasst. Jeder Fall entspricht einem Teilnehmer.
- **clothing\_defects.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Qualitätskontrolle in einer Bekleidungsfabrik geht. Aus jeder in der Fabrik produzierten Charge entnehmen die Kontrolleure eine Stichprobe an Bekleidungsartikeln und zählen die Anzahl der Bekleidungsartikel die inakzeptabel sind.
- **coffee.sav.** Diese Datendatei enthält Daten zum wahrgenommenen Image von sechs Eiskaffeearten. Bei den 23 Attributen des Eiskaffee-Image sollten die Teilnehmer jeweils alle Marken auswählen, die durch dieses Attribut beschrieben werden. Die sechs Marken

werden als “AA”, “BB”, “CC”, “DD”, “EE” und “FF” bezeichnet, um Vertraulichkeit zu gewährleisten.

- **contacts.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Kontaktlisten einer Gruppe von Vertretern geht, die Computer an Unternehmen verkaufen. Die einzelnen Kontaktpersonen werden anhand der Abteilung, in der sie in ihrem Unternehmen arbeiten und anhand ihrer Stellung in der Unternehmenshierarchie in Kategorien eingeteilt. Außerdem werden der Betrag des letzten Verkaufs, die Zeit seit dem letzten Verkauf und die Größe des Unternehmens, in dem die Kontaktperson arbeitet, aufgezeichnet.
- **creditpromo.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen eines Kaufhauses geht, die Wirksamkeit einer kürzlich durchgeführten Kreditkarten-Werbeaktion einzuschätzen. Dazu wurden 500 Karteninhaber nach dem Zufallsprinzip ausgewählt. Die Hälfte erhielt eine Werbebeilage, die einen reduzierten Zinssatz für Einkäufe in den nächsten drei Monaten ankündigte. Die andere Hälfte erhielt eine Standard-Werbebeilage.
- **customer\_dbase.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen eines Unternehmens geht, das die Informationen in seinem Data Warehouse nutzen möchte, um spezielle Angebote für Kunden zu erstellen, die mit der größten Wahrscheinlichkeit darauf ansprechen. Nach dem Zufallsprinzip wurde eine Untergruppe des Kundenstamms ausgewählt. Diese Gruppe erhielt die speziellen Angebote und die Reaktionen wurden aufgezeichnet.
- **customer\_information.sav.** Eine hypothetische Datendatei mit Kundenmailingdaten wie Name und Adresse.
- **customer\_subset.sav.** Eine Teilmenge von 80 Fällen aus der Datei *customer\_dbase.sav*.
- **debate.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die gepaarte Antworten auf eine Umfrage unter den Zuhörern einer politischen Debatte enthält (Antworten vor und nach der Debatte). Jeder Fall entspricht einem Befragten.
- **debate\_aggregate.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, in der die Antworten aus *debate.sav* aggregiert wurden. Jeder Fall entspricht einer Kreuzklassifikation der bevorzugten Politiker vor und nach der Debatte.
- **demo.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um eine Kundendatenbank geht, die zum Zwecke der Zusendung monatlicher Angebote erworben wurde. Neben verschiedenen demografischen Informationen ist erfasst, ob der Kunde auf das Angebot geantwortet hat.
- **demo\_cs\_1.sav.** Hierbei handelt es sich um eine hypothetische Datendatei für den ersten Schritt eines Unternehmens, das eine Datenbank mit Umfrageinformationen zusammenstellen möchte. Jeder Fall entspricht einer anderen Stadt. Außerdem sind IDs für Region, Provinz, Landkreis und Stadt erfasst.
- **demo\_cs\_2.sav.** Hierbei handelt es sich um eine hypothetische Datendatei für den zweiten Schritt eines Unternehmens, das eine Datenbank mit Umfrageinformationen zusammenstellen möchte. Jeder Fall entspricht einem anderen Stadtteil aus den im ersten Schritt ausgewählten Städten. Außerdem sind IDs für Region, Provinz, Landkreis, Stadt, Stadtteil und Wohneinheit erfasst. Die Informationen zur Stichprobenziehung aus den ersten beiden Stufen des Stichprobenplans sind ebenfalls enthalten.

- **demo\_cs.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die Umfrageinformationen enthält die mit einem komplexen Stichprobenplan erfasst wurden. Jeder Fall entspricht einer anderen Wohneinheit. Es sind verschiedene Informationen zum demografischen Hintergrund und zur Stichprobenziehung erfasst.
- **dmdata.sav.** Dies ist eine hypothetische Datendatei, die demografische und kaufbezogene Daten für ein Direktmarketingunternehmen enthält. *dmdata2.sav* enthält Informationen für eine Teilmenge von Kontakten, die ein Testmailing erhalten. *dmdata3.sav* enthält Informationen zu den verbleibenden Kontakten, die kein Testmailing erhalten.
- **dietstudy.sav.** Diese hypothetische Datendatei enthält die Ergebnisse einer Studie der “Stillman-Diät”. Jeder Fall entspricht einem Teilnehmer und enthält dessen Gewicht vor und nach der Diät in amerikanischen Pfund sowie mehrere Messungen des Triglyceridspiegels (in mg/100 ml).
- **dvdplayer.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Entwicklung eines neuen DVD-Spielers geht. Mithilfe eines Prototyps hat das Marketing-Team Zielgruppendaten erfasst. Jeder Fall entspricht einem befragten Benutzer und enthält demografische Daten zu dem Benutzer sowie dessen Antworten auf Fragen zum Prototyp.
- **german\_credit.sav.** Diese Daten sind aus dem Daten-Set “German credit” im Repository of Machine Learning Databases () an der Universität von Kalifornien in Irvine entnommen.
- **grocery\_1month.sav.** Bei dieser hypothetischen Datendatei handelt es sich um die Datendatei *grocery\_coupons.sav*, wobei die wöchentlichen Einkäufe zusammengefasst sind, sodass jeder Fall einem anderen Kunden entspricht. Dadurch entfallen einige der Variablen, die wöchentlichen Änderungen unterworfen waren, und der verzeichnete ausgegebene Betrag ist nun die Summe der Beträge, die in den vier Wochen der Studie ausgegeben wurden.
- **grocery\_coupons.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die Umfragedaten enthält, die von einer Lebensmittelkette erfasst wurden, die sich für die Kaufgewohnheiten ihrer Kunden interessiert. Jeder Kunde wird über vier Wochen beobachtet, und jeder Fall entspricht einer Kundenwoche und enthält Informationen zu den Geschäften, in denen der Kunde einkauft sowie zu anderen Merkmalen, beispielsweise welcher Betrag in der betreffenden Woche für Lebensmittel ausgegeben wurde.
- **guttman.sav.** Bell () legte eine Tabelle zur Darstellung möglicher sozialer Gruppen vor. Guttman () verwendete einen Teil dieser Tabelle, bei der fünf Variablen, die Aspekte beschreiben, wie soziale Interaktion, das Gefühl der Gruppenzugehörigkeit, die physische Nähe der Mitglieder und die Formalität der Beziehung, mit sieben theoretischen sozialen Gruppen gekreuzt wurden: “crowds” (Menschenmassen, beispielsweise die Zuschauer eines Fußballspiels), “audience” (Zuhörerschaften, beispielsweise die Personen im Theater oder bei einer Vorlesung), “public” (Öffentlichkeit, beispielsweise Zeitungsleser oder Fernsehzuschauer), “mobs” (Mobs, wie Menschenmassen, jedoch mit wesentlich stärkerer Interaktion), “primary groups” (Primärgruppen, vertraulich), “secondary groups” (Sekundärgruppen, freiwillig) und “modern community” (die moderne Gesellschaft, ein lockerer Zusammenschluss, der aus einer engen physischen Nähe und dem Bedarf an spezialisierten Dienstleistungen entsteht).
- **health\_funding.sav.** Hierbei handelt es sich um eine hypothetische Datei, die Daten zur Finanzierung des Gesundheitswesens (Betrag pro 100 Personen), Krankheitsraten (Rate pro 10.000 Personen der Bevölkerung) und Besuche bei medizinischen Einrichtungen/Ärzten (Rate pro 10.000 Personen der Bevölkerung) enthält. Jeder Fall entspricht einer anderen Stadt.

- **hivassay.sav.** Hierbei handelt es sich um eine hypothetische Datendatei zu den Bemühungen eines pharmazeutischen Labors, einen Schnelltest zur Erkennung von HIV-Infektionen zu entwickeln. Die Ergebnisse des Tests sind acht kräftiger werdende Rotschattierungen, wobei kräftigeren Schattierungen auf eine höhere Infektionswahrscheinlichkeit hindeuten. Bei 2.000 Blutproben, von denen die Hälfte mit HIV infiziert war, wurde ein Labortest durchgeführt.
- **hourlywagedata.sav.** Hierbei handelt es sich um eine hypothetische Datendatei zum Stundenlohn von Pflegepersonal in Praxen und Krankenhäusern mit unterschiedlich langer Berufserfahrung.
- **insurance\_claims.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um eine Versicherungsgesellschaft geht, die ein Modell zur Kennzeichnung verdächtiger, potenziell betrügerischer Ansprüche erstellen möchte. Jeder Fall entspricht einem Anspruch.
- **insure.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um eine Versicherungsgesellschaft geht, die die Risikofaktoren untersucht, die darauf hinweisen, ob ein Kunde die Leistungen einer mit einer Laufzeit von 10 Jahren abgeschlossenen Lebensversicherung in Anspruch nehmen wird. Jeder Fall in der Datendatei entspricht einem Paar von Verträgen, je einer mit Leistungsforderung und der andere ohne, wobei die beiden Versicherungsnehmer in Alter und Geschlecht übereinstimmen.
- **judges.sav.** Hierbei handelt es sich um eine hypothetische Datendatei mit den Wertungen von ausgebildeten Kampfrichtern (sowie eines Sportliebhabers) zu 300 Kunstturnleistungen. Jede Zeile stellt eine Leistung dar; die Kampfrichter bewerteten jeweils dieselben Leistungen.
- **kinship\_dat.sav.** Rosenberg und Kim () haben 15 Bezeichnungen für den Verwandtschaftsgrad untersucht (Tante, Bruder, Cousin, Tochter, Vater, Enkelin, Großvater, Großmutter, Enkel, Mutter, Nefte, Nichte, Schwester, Sohn, Onkel). Die beiden Analytiker baten vier Gruppen von College-Studenten (zwei weibliche und zwei männliche Gruppen), diese Bezeichnungen auf der Grundlage der Ähnlichkeiten zu sortieren. Zwei Gruppen (eine weibliche und eine männliche Gruppe) wurden gebeten, die Bezeichnungen zweimal zu sortieren; die zweite Sortierung sollte dabei nach einem anderen Kriterium erfolgen als die erste. So wurden insgesamt sechs "Quellen" erzielt. Jede Quelle entspricht einer Ähnlichkeitsmatrix mit  $15 \times 15$  Elementen. Die Anzahl der Zellen ist dabei gleich der Anzahl der Personen in einer Quelle minus der Anzahl der gemeinsamen Platzierungen der Objekte in dieser Quelle.
- **kinship\_ini.sav.** Diese Datendatei enthält eine Ausgangskonfiguration für eine dreidimensionale Lösung für *kinship\_dat.sav*.
- **kinship\_var.sav.** Diese Datendatei enthält die unabhängigen Variablen *gender* (Geschlecht), *gener*(Generation) und *degree* (Verwandtschaftsgrad), die zur Interpretation der Dimensionen einer Lösung für *kinship\_dat.sav* verwendet werden können. Insbesondere können sie verwendet werden, um den Lösungsraum auf eine lineare Kombination dieser Variablen zu beschränken.
- **marketvalues.sav.** Diese Datendatei betrifft Hausverkäufe in einem Neubaugebiet in Algonquin, Illinois, in den Jahren 1999–2000. Diese Verkäufe sind in Grundbucheinträgen dokumentiert.
- **nhis2000\_subset.sav.** Die "National Health Interview Survey (NHIS)" ist eine große, bevölkerungsbezogene Umfrage in unter der US-amerikanischen Zivilbevölkerung. Es werden persönliche Interviews in einer landesweit repräsentativen Stichprobe von Haushalten durchgeführt. Für die Mitglieder jedes Haushalts werden demografische Informationen und Beobachtungen zum Gesundheitsverhalten und Gesundheitsstatus eingeholt. Diese Datendatei

enthält eine Teilmenge der Informationen aus der Umfrage des Jahres 2000. National Center for Health Statistics. National Health Interview Survey, 2000. Datendatei und Dokumentation öffentlich zugänglich. [ftp://ftp.cdc.gov/pub/Health\\_Statistics/NCHS/Datasets/NHIS/2000/](ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHIS/2000/). Zugriff erfolgte 2003.

- **ozone.sav.** Die Daten enthalten 330 Beobachtungen zu sechs meteorologischen Variablen zur Vorhersage der Ozonkonzentration aus den übrigen Variablen. Bei früheren Untersuchungen (, ) fanden Wissenschaftler einige Nichtlinearitäten unter diesen Variablen, die die Standardverfahren bei der Regression behindern.
- **pain\_medication.sav.** Diese hypothetische Datendatei enthält die Ergebnisse eines klinischen Tests für ein entzündungshemmendes Medikament zur Schmerzbehandlung bei chronischer Arthritis. Von besonderem Interesse ist die Zeitdauer, bis die Wirkung des Medikaments einsetzt und wie es im Vergleich mit bestehenden Medikamenten abschneidet.
- **patient\_los.sav.** Diese hypothetische Datendatei enthält die Behandlungsaufzeichnungen zu Patienten, die wegen des Verdachts auf Herzinfarkt in das Krankenhaus eingeliefert wurden. Jeder Fall entspricht einem Patienten und enthält diverse Variablen in Bezug auf den Krankenhausaufenthalt.
- **patlos\_sample.sav.** Diese hypothetische Datendatei enthält die Behandlungsaufzeichnungen für eine Stichprobe von Patienten, denen während der Behandlung eines Herzinfarkts Thrombolytika verabreicht wurden. Jeder Fall entspricht einem Patienten und enthält diverse Variablen in Bezug auf den Krankenhausaufenthalt.
- **poll\_cs.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um Bemühungen geht, die öffentliche Unterstützung für einen Gesetzentwurf zu ermitteln, bevor er im Parlament eingebracht wird. Die Fälle entsprechen registrierten Wählern. Für jeden Fall sind County, Gemeinde und Wohnviertel des Wählers erfasst.
- **poll\_cs\_sample.sav.** Diese hypothetische Datendatei enthält eine Stichprobe der in *poll\_cs.sav* aufgeführten Wähler. Die Stichprobe wurde gemäß dem in der Plandatei *poll.csplan* angegebenen Stichprobenplan gezogen und in dieser Datendatei sind die Einschlusswahrscheinlichkeiten und Stichprobengewichtungen erfasst. Beachten Sie jedoch Folgendes: Da im Stichprobenplan die PPS-Methode (PPS: probability proportional to size; Wahrscheinlichkeit proportional zur Größe) verwendet wird, gibt es außerdem eine Datei mit den gemeinsamen Auswahlwahrscheinlichkeiten (*poll\_jointprob.sav*). Die zusätzlichen Variablen zum demografischen Hintergrund der Wähler und ihrer Meinung zum vorgeschlagenen Gesetzentwurf wurden nach der Ziehung der Stichprobe erfasst und zur Datendatei hinzugefügt.
- **property\_assess.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, in der es um die Bemühungen eines für einen Bezirk (County) zuständigen Immobilienbewerbers geht, trotz eingeschränkter Ressourcen die Einschätzungen des Werts von Immobilien auf dem aktuellsten Stand zu halten. Die Fälle entsprechen den Immobilien, die im vergangenen Jahr in dem betreffenden County verkauft wurden. Jeder Fall in der Datendatei enthält die Gemeinde, in der sich die Immobilie befindet, den Bewerter, der die Immobilie besichtigt hat, die seit dieser Bewertung verstrichene Zeit, den zu diesem Zeitpunkt ermittelten Wert sowie den Verkaufswert der Immobilie.
- **property\_assess\_cs.sav** Hierbei handelt es sich um eine hypothetische Datendatei, in der es um die Bemühungen eines für einen US-Bundesstaat zuständigen Immobilienbewerbers geht, trotz eingeschränkter Ressourcen die Einschätzungen des Werts von Immobilien auf dem aktuellsten Stand zu halten. Die Fälle entsprechen den Immobilien in dem betreffenden

Bundesstaat. Jeder Fall in der Datendatei enthält das County, die Gemeinde und das Wohnviertel, in dem sich die Immobilie befindet, die seit der letzten Bewertung verstrichene Zeit sowie zu diesem Zeitpunkt ermittelten Wert.

- **property\_assess\_cs\_sample.sav.** Diese hypothetische Datendatei enthält eine Stichprobe der in *property\_assess\_cs.sav* aufgeführten Immobilien. Die Stichprobe wurde gemäß dem in der Plandatei *property\_assess.csplan* angegebenen Stichprobenplan gezogen und in dieser Datendatei sind die Einschlusswahrscheinlichkeiten und Stichprobengewichtungen erfasst. Die zusätzliche Variable *Current value* (Aktueller Wert) wurde nach der Ziehung der Stichprobe erfasst und zur Datendatei hinzugefügt.
- **recidivism.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen einer Strafverfolgungsbehörde geht, einen Einblick in die Rückfallraten in ihrem Zuständigkeitsbereich zu gewinnen. Jeder Fall entspricht einem früheren Straftäter und erfasst Daten zu dessen demografischen Hintergrund, einige Details zu seinem ersten Verbrechen sowie die Zeit bis zu seiner zweiten Festnahme, sofern diese innerhalb von zwei Jahren nach der ersten Festnahme erfolgte.
- **recidivism\_cs\_sample.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen einer Strafverfolgungsbehörde geht, einen Einblick in die Rückfallraten in ihrem Zuständigkeitsbereich zu gewinnen. Jeder Fall entspricht einem früheren Straftäter, der im Juni 2003 erstmals aus der Haft entlassen wurde, und erfasst Daten zu dessen demografischen Hintergrund, einige Details zu seinem ersten Verbrechen sowie die Daten zu seiner zweiten Festnahme, sofern diese bis Ende Juni 2006 erfolgte. Die Straftäter wurden aus per Stichprobenziehung ermittelten Polizeidirektionen ausgewählt (gemäß dem in *recidivism\_cs.csplan* angegebenen Stichprobenplan). Da hierbei eine PPS-Methode (PPS: probability proportional to size; Wahrscheinlichkeit proportional zur Größe) verwendet wird, gibt es außerdem eine Datei mit den gemeinsamen Auswahlwahrscheinlichkeiten (*recidivism\_cs\_jointprob.sav*).
- **rfm\_transactions.sav.** Eine hypothetische Datendatei mit Kauftransaktionsdaten wie Kaufdatum, gekauften Artikeln und Geldbetrag für jede Transaktion.
- **salesperformance.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um Bewertung von zwei neuen Verkaufsschulungen geht. 60 Mitarbeiter, die in drei Gruppen unterteilt sind, erhalten jeweils eine Standardschulung. Zusätzlich erhält Gruppe 2 eine technische Schulung und Gruppe 3 eine Praxisschulung. Die einzelnen Mitarbeiter wurden am Ende der Schulung einem Test unterzogen und die erzielten Punkte wurden erfasst. Jeder Fall in der Datendatei stellt einen Lehrgangsteilnehmer dar und enthält die Gruppe, der der Lehrgangsteilnehmer zugeteilt wurde sowie die von ihm in der Prüfung erreichte Punktzahl.
- **satisf.sav.** Hierbei handelt es sich um eine hypothetische Datendatei zu einer Zufriedenheitsumfrage, die von einem Einzelhandelsunternehmen in 4 Filialen durchgeführt wurde. Insgesamt wurden 582 Kunden befragt. Jeder Fall gibt die Antworten eines einzelnen Kunden wieder.
- **screws.sav.** Diese Datendatei enthält Informationen zu den Eigenschaften von Schrauben, Bolzen, Muttern und Reißnägeln ().
- **shampoo\_ph.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Qualitätskontrolle in einer Fabrik für Haarpflegeprodukte geht. In regelmäßigen Zeitabständen werden Messwerte von sechs separaten Ausgangschargen erhoben und ihr pH-Wert erfasst. Der Zielbereich ist 4,5–5,5.

- **ships.sav.** Ein an anderer Stelle () vorgestelltes und analysiertes Daten-Set bezieht sich auf die durch Wellen verursachten Schäden an Frachtschiffen. Die Vorfalshäufigkeiten können unter Angabe von Schiffstyp, Konstruktionszeitraum und Betriebszeitraum gemäß einer Poisson-Rate modelliert werden. Das Aggregat der Betriebsmonate für jede Zelle der durch die Kreuzklassifizierung der Faktoren gebildeten Tabelle gibt die Werte für die Risikoanfälligkeit an.
- **site.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen eines Unternehmens geht, neue Standorte für die betriebliche Expansion auszuwählen. Das Unternehmen beauftragte zwei Berater unabhängig voneinander mit der Bewertung der Standorte. Neben einem umfassenden Bericht gaben die Berater auch eine zusammenfassende Wertung für jeden Standort als “good” (gut) “fair” (mittelmäßig) oder “poor” (schlecht) ab.
- **smokers.sav.** Diese Datendatei wurde aus der Umfrage “National Household Survey of Drug Abuse” aus dem Jahr 1998 abstrahiert und stellt eine Wahrscheinlichkeitsstichprobe US-amerikanischer Haushalte dar. (<http://dx.doi.org/10.3886/ICPSR02934>) Daher sollte der erste Schritt bei der Analyse dieser Datendatei darin bestehen, die Daten entsprechend den Bevölkerungstrends zu gewichten.
- **stocks.sav** Diese hypothetische Datendatei umfasst Börsenkurse und -volumina für ein Jahr.
- **stroke\_clean.sav.** Diese hypothetische Datendatei enthält den Zustand einer medizinischen Datenbank, nachdem diese mithilfe der Prozeduren in der Option “Data Preparation” bereinigt wurde.
- **stroke\_invalid.sav.** Diese hypothetische Datendatei enthält den ursprünglichen Zustand einer medizinischen Datenbank, der mehrere Dateneingabefehler aufweist.
- **stroke\_survival.** In dieser hypothetischen Datendatei geht es um die Überlebenszeiten von Patienten, die nach einem Rehabilitationsprogramm wegen eines ischämischen Schlaganfalls mit einer Reihe von Problemen zu kämpfen haben. Nach dem Schlaganfall werden das Auftreten von Herzinfarkt, ischämischem Schlaganfall und hämorrhagischem Schlaganfall sowie der Zeitpunkt des Ereignisses aufgezeichnet. Die Stichprobe ist auf der linken Seite abgeschnitten, da sie nur Patienten enthält, die bis zum Ende des Rehabilitationprogramms, das nach dem Schlaganfall durchgeführt wurde, überlebten.
- **stroke\_valid.sav.** Diese hypothetische Datendatei enthält den Zustand einer medizinischen Datenbank, nachdem diese mithilfe der Prozedur “Daten validieren” überprüft wurde. Sie enthält immer noch potenziell anomale Fälle.
- **survey\_sample.sav.** Diese Datendatei enthält Umfragedaten einschließlich demografischer Daten und verschiedener Meinungskennzahlen. Sie beruht auf einer Teilmenge der Variablen aus der NORC General Social Survey aus dem Jahr 1998. Allerdings wurden zu Demonstrationszwecken einige Daten abgeändert und weitere fiktive Variablen hinzugefügt.
- **telco.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen eines Telekommunikationsunternehmens geht, die Kundenabwanderung zu verringern. Jeder Fall entspricht einem Kunden und enthält verschiedene Informationen zum demografischen Hintergrund und zur Servicenutzung.
- **telco\_extra.sav.** Diese Datendatei ähnelt der Datei *telco.sav*, allerdings wurden die Variablen “tenure” und die Log-transformierten Variablen zu den Kundenausgaben entfernt und durch standardisierte Log-transformierte Variablen ersetzt.

- **telco\_missing.sav.** Diese Datendatei ist eine Untermenge der Datendatei *telco.sav*, allerdings wurde ein Teil der demografischen Datenwerte durch fehlende Werte ersetzt.
- **testmarket.sav.** Diese hypothetische Datendatei bezieht sich auf die Pläne einer Fast-Food-Kette, einen neuen Artikel in ihr Menü aufzunehmen. Es gibt drei mögliche Kampagnen zur Verkaufsförderung für das neue Produkt. Daher wird der neue Artikel in Filialen in mehreren zufällig ausgewählten Märkten eingeführt. An jedem Standort wird eine andere Form der Verkaufsförderung verwendet und die wöchentlichen Verkaufszahlen für das neue Produkt werden für die ersten vier Wochen aufgezeichnet. Jeder Fall entspricht einer Standort-Woche.
- **testmarket\_1month.sav.** Bei dieser hypothetischen Datendatei handelt es sich um die Datendatei *testmarket.sav*, wobei die wöchentlichen Verkaufszahlen zusammengefasst sind, sodass jeder Fall einem Standort entspricht. Dadurch entfallen einige der Variablen, die wöchentlichen Änderungen unterworfen waren, und die verzeichneten Verkaufszahlen sind nun die Summe der Verkaufszahlen während der vier Wochen der Studie.
- **tree\_car.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die demografische Daten sowie Daten zum Kaufpreis von Fahrzeugen enthält.
- **tree\_credit.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die demografische Daten sowie Daten zu früheren Bankkrediten enthält.
- **tree\_missing\_data.sav** Hierbei handelt es sich um eine hypothetische Datendatei, die demografische Daten sowie Daten zu früheren Bankkrediten enthält und eine große Anzahl fehlender Werte aufweist.
- **tree\_score\_car.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die demografische Daten sowie Daten zum Kaufpreis von Fahrzeugen enthält.
- **tree\_textdata.sav.** Eine einfache Datendatei mit nur zwei Variablen, die vor allem den Standardzustand von Variablen vor der Zuweisung von Messniveau und Wertelabels zeigen soll.
- **tv-survey.sav.** Hierbei handelt es sich um eine hypothetische Datendatei zu einer Studie, die von einem Fernsehstudio durchgeführt wurde, das überlegt, ob die Laufzeit eines erfolgreichen Programms verlängert werden soll. 906 Personen wurden gefragt, ob sie das Programm unter verschiedenen Bedingungen ansehen würden. Jede Zeile entspricht einem Befragten; jede Spalte entspricht einer Bedingung.
- **ulcer\_recurrence.sav.** Diese Datei enthält Teilm Informationen aus einer Studie zum Vergleich der Wirksamkeit zweier Therapien zur Vermeidung des Wiederauftretens von Geschwüren. Es stellt ein gutes Beispiel für intervallzensierte Daten dar und wurde an anderer Stelle () vorgestellt und analysiert.
- **ulcer\_recurrence\_recoded.sav.** In dieser Datei sind die Daten aus *ulcer\_recurrence.sav* so umstrukturiert, dass das Modell der Ereigniswahrscheinlichkeit für jedes Intervall der Studie berechnet werden kann und nicht nur die Ereigniswahrscheinlichkeit am Ende der Studie. Sie wurde an anderer Stelle () vorgestellt und analysiert.
- **verd1985.sav.** Diese Datendatei enthält eine Umfrage (). Die Antworten von 15 Subjekten auf 8 Variablen wurden aufgezeichnet. Die relevanten Variablen sind in drei Sets unterteilt. Set 1 umfasst *alter* und *heirat*, Set 2 besteht aus *pet* und *news* und in Set 3 finden sich *music* und *live*. Die Variable *pet* wird mehrfach nominal skaliert und die Variable *Alter* ordinal. Alle anderen Variablen werden einzeln nominal skaliert.

- **virus.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen eines Internet-Diensteanbieters geht, der die Auswirkungen eines Virus auf seine Netzwerke ermitteln möchte. Dabei wurde vom Moment der Virusentdeckung bis zu dem Zeitpunkt, zu dem die Virusinfektion unter Kontrolle war, der (ungefähre) prozentuale Anteil infizierter E-Mail in den Netzwerken erfasst.
- **wheeze\_steubenville.sav.** Hierbei handelt es sich um eine Teilmenge der Daten aus einer Langzeitstudie zu den gesundheitlichen Auswirkungen der Luftverschmutzung auf Kinder (). Die Daten enthalten wiederholte binäre Messungen des Keuchens von Kindern aus Steubenville, Ohio, im Alter von 7, 8, 9 und 10 Jahren sowie eine unveränderliche Angabe, ob die Mutter im ersten Jahr der Studie rauchte oder nicht.
- **workprog.sav.** Hierbei handelt es sich um eine hypothetische Datendatei zu einem Arbeitsprogramm der Regierung, das versucht, benachteiligten Personen bessere Arbeitsplätze zu verschaffen. Eine Stichprobe potenzieller Programmteilnehmer wurde beobachtet. Von diesen Personen wurden nach dem Zufallsprinzip einige für die Teilnahme an dem Programm ausgewählt. Jeder Fall entspricht einem Programmteilnehmer.
- **worldsales.sav** Diese hypothetische Datendatei enthält Verkaufserlöse nach Kontinent und Produkt.

## **Hinweise**

Diese Informationen wurden für weltweit angebotene Produkte und Dienstleistungen erarbeitet.

IBM bietet die in diesem Dokument behandelten Produkte, Dienstleistungen oder Merkmale möglicherweise nicht in anderen Ländern an. Informationen zu den derzeit in Ihrem Land erhältlichen Produkten und Dienstleistungen erhalten Sie bei Ihrem zuständigen IBM-Mitarbeiter vor Ort. Mit etwaigen Verweisen auf Produkte, Programme oder Dienste von IBM soll nicht behauptet oder impliziert werden, dass nur das betreffende Produkt oder Programm bzw. der betreffende Dienst von IBM verwendet werden kann. Stattdessen können alle funktional gleichwertigen Produkte, Programme oder Dienste verwendet werden, die keine geistigen Eigentumsrechte von IBM verletzen. Es obliegt jedoch der Verantwortung des Benutzers, die Funktionsweise von Produkten, Programmen oder Diensten von Drittanbietern zu bewerten und zu überprüfen.

IBM verfügt möglicherweise über Patente oder hat Patentanträge gestellt, die sich auf in diesem Dokument beschriebene Inhalte beziehen. Durch die Bereitstellung dieses Dokuments werden Ihnen keinerlei Lizenzen an diesen Patenten gewährt. Lizenzanfragen können schriftlich an folgende Adresse gesendet werden:

*IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785, U.S.A.*

Bei Lizenzanfragen in Bezug auf DBCS-Daten (Double-Byte Character Set) wenden Sie sich an die für geistiges Eigentum zuständige Abteilung von IBM in Ihrem Land. Schriftliche Anfragen können Sie auch an folgende Adresse senden:

*Intellectual Property Licensing, Legal and Intellectual Property Law, IBM Japan Ltd., 1623-14, Shimotsuruma, Yamato-shi, Kanagawa 242-8502 Japan.*

**Der folgende Abschnitt findet in Großbritannien und anderen Ländern keine Anwendung, in denen solche Bestimmungen nicht mit der örtlichen Gesetzgebung vereinbar sind:** INTERNATIONAL BUSINESS MACHINES STELLT DIESE VERÖFFENTLICHUNG IN DER VERFÜGBAREN FORM OHNE GARANTIEN BEREIT, SEIEN ES AUSDRÜCKLICHE ODER STILLSCHWEIGENDE, EINSCHLIESSLICH JEDOCH NICHT NUR DER GARANTIEN BEZÜGLICH DER NICHT-RECHTSVERLETZUNG, DER GÜTE UND DER EIGNUNG FÜR EINEN BESTIMMTEN ZWECK. Manche Rechtsprechungen lassen den Ausschluss ausdrücklicher oder implizierter Garantien bei bestimmten Transaktionen nicht zu, sodass die oben genannte Ausschlussklausel möglicherweise nicht für Sie relevant ist.

Diese Informationen können technische Ungenauigkeiten oder typografische Fehler aufweisen. An den hierin enthaltenen Informationen werden regelmäßig Änderungen vorgenommen. Diese Änderungen werden in neuen Ausgaben der Veröffentlichung aufgenommen. IBM kann jederzeit und ohne vorherige Ankündigung Optimierungen und/oder Änderungen an den Produkten und/oder Programmen vornehmen, die in dieser Veröffentlichung beschrieben werden.

Jegliche Verweise auf Drittanbieter-Websites in dieser Information werden nur der Vollständigkeit halber bereitgestellt und dienen nicht als Befürwortung dieser. Das Material auf diesen Websites ist kein Bestandteil des Materials zu diesem IBM-Produkt und die Verwendung erfolgt auf eigene Gefahr.

IBM kann die von Ihnen angegebenen Informationen verwenden oder weitergeben, wie dies angemessen erscheint, ohne Ihnen gegenüber eine Verpflichtung einzugehen.

Lizenznehmer dieses Programms, die Informationen dazu benötigen, wie (i) der Austausch von Informationen zwischen unabhängig erstellten Programmen und anderen Programmen und (ii) die gegenseitige Verwendung dieser ausgetauschten Informationen ermöglicht wird, wenden sich an:

*IBM Software Group, Attention: Licensing, 233 S. Wacker Dr., Chicago, IL 60606, USA.*

Derartige Informationen stehen ggf. in Abhängigkeit von den jeweiligen Geschäftsbedingungen sowie in einigen Fällen der Zahlung einer Gebühr zur Verfügung.

Das in diesem Dokument beschriebene lizenzierte Programm und sämtliche dafür verfügbaren lizenzierten Materialien werden von IBM gemäß dem IBM-Kundenvertrag, den Internationalen Nutzungsbedingungen für Programmpakete der IBM oder einer anderen zwischen uns getroffenen Vereinbarung bereitgestellt.

Informationen zu Produkten von Drittanbietern wurden von den Anbietern des jeweiligen Produkts, aus deren veröffentlichten Ankündigungen oder anderen, öffentlich verfügbaren Quellen bezogen. IBM hat diese Produkte nicht getestet und kann die Genauigkeit bezüglich Leistung, Kompatibilität oder anderen Behauptungen nicht bestätigen, die sich auf Drittanbieter-Produkte beziehen. Fragen bezüglich der Funktionen von Drittanbieter-Produkten sollten an die Anbieter der jeweiligen Produkte gerichtet werden.

Diese Informationen enthalten Beispiele zu Daten und Berichten, die im täglichen Geschäftsbetrieb Verwendung finden. Um diese so vollständig wie möglich zu illustrieren, umfassen die Beispiele Namen von Personen, Unternehmen, Marken und Produkten. Alle diese Namen sind fiktiv und jegliche Ähnlichkeit mit Namen und Adressen realer Unternehmen ist rein zufällig.

Unter Umständen werden Fotografien und farbige Abbildungen nicht angezeigt, wenn Sie diese Informationen nicht in gedruckter Form verwenden.

### **Marken**

IBM, das IBM-Logo, ibm.com und SPSS sind Marken der IBM Corporation und in vielen Ländern weltweit registriert. Eine aktuelle Liste der IBM-Marken finden Sie im Internet unter <http://www.ibm.com/legal/copytrade.shtml>.

Adobe, das Adobe-Logo, PostScript und das PostScript-Logo sind eingetragene Marken oder Marken von 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 der Tochtergesellschaften des Unternehmens in den USA und anderen Ländern.

Java und alle Java-basierten Marken sowie Logos sind Marken von Sun Microsystems, Inc. in den USA, anderen Ländern oder beidem.

Linux ist eine eingetragene Marke von Linus Torvalds in den USA, anderen Ländern oder beidem.

Microsoft, Windows, Windows NT und das Windows-Logo sind Marken der Microsoft Corporation in den USA, anderen Ländern oder beidem.

UNIX ist eine eingetragene Marke der The Open Group in den USA und anderen Ländern.

In diesem Produkt wird WinWrap Basic verwendet, Copyright 1993-2007, Polar Engineering and Consulting, <http://www.winwrap.com>.

Andere Produkt- und Servicenamen können Marken von IBM oder anderen Unternehmen sein.

Screenshots von Adobe-Produkten werden mit Genehmigung von Adobe Systems Incorporated abgedruckt.

Screenshots von Microsoft-Produkten werden mit Genehmigung der Microsoft Corporation abgedruckt.



- Angepaßte Auswertungsstatistik für Gesamtsummen, 130
- Anzahl
  - und gültige N, 139
- Anzeigeformate, 63
  - Auswertungsstatistiken in benutzerdefinierten Tabellen, 41, 190
- Anzeigen von Kategoriewerten, 134
- Ausblenden der Beschriftungen für die Statistiken in benutzerdefinierten Tabellen, 57
- Ausschließen von Kategorien
  - Custom Tables (Benutzerdefinierte Tabellen), 42
  - Tabellen, 65
- Auswertungsstatistik, 121
  - Ändern des Beschriftungstexts, 194
  - Angepaßte Auswertungsstatistik für Gesamtsummen, 130
  - Anzeigeformat, 190
  - gestapelte Tabellen, 127
  - Quelldimension, 122
  - Quellevariable, 122
  - unterschiedliche Auswertungen für unterschiedliche Variablen in gestapelten Tabellen, 141
- Beispieldateien
  - Speicherort, 201
- Benutzerdefinierte fehlende Werte, 184
- benutzerdefinierte Tabellen
  - Verarbeitung aufgeteilter Dateien, 32
- Benutzerdefinierte Tabellen
  - Ändern der Dimension von Auswertungsstatistiken, 40
  - Ändern des Messniveaus, 27
  - angepaßte Gesamtsummen, 40
  - Anzeigeformate, 35
  - Anzeigeformate für Auswertungsstatistiken, 41
  - Ausblenden der Kategorien für das Zwischenergebnis, 96
  - Ausschließen von fehlenden Werten bei der Auswertung metrischer Variablen, 49
  - Ausschließen von Kategorien, 42
  - Auswertungsstatistik, 36, 38–39
  - Berechnete Kategorien, 42, 100
  - Comperimeter-Tabellen, 48, 111
  - Eckenbeschriftungen, 51
  - Einblenden und Ausblenden von Variablenamen und -labels, 34
  - Erklärungen, 51
  - Gesamtsummen, 42, 87
  - Häufigkeitstabellen, 48, 111
  - kategoriale Variablen, 27
  - Konstruieren einer Tabelle, 30
  - Leere Zellen, 49
  - Mehrfachantworten-Sets, 27
  - Metrische Variablen, 27
  - Postcompute-Kategorien, 100
  - Prozentwerte, 36, 38
  - Prozentwerte für Mehrfachantworten-Sets, 38
  - Reduzieren von Kategorien, 96
  - Sets aus kategorialen Variablen, 49
  - Spaltenbreite, 49
  - Steuern der angezeigten Dezimalstellen, 35
  - Tabellen für Mittelwert und Häufigkeit, 40
  - Tabellen für Variablen mit gemeinsamen Kategorien, 48, 111
  - Teststatistiken, 52
  - Titel, 51
  - Umordnen von Kategorien, 42
  - Wertelabels für kategoriale Variablen, 27
  - Zwischenergebnisse, 42, 87
- Berechnete Kategorien
  - Anzeigeformate, 48
  - aus Zwischenergebnissen, 104
  - Ausblenden von Kategorien im Ausdruck, 103
  - Benutzerdefinierte Tabellen, 100
  - Custom Tables (Benutzerdefinierte Tabellen), 46
- Bereich
  - Custom Tables (Benutzerdefinierte Tabellen), 39
- Beschriftungen
  - Ändern des Beschriftungstexts für Auswertungsstatistiken, 194
- Chi-Quadrat
  - Custom Tables (Benutzerdefinierte Tabellen), 148
- Comperimeter-Tabellen, 48, 111
- Custom Tables (Benutzerdefinierte Tabellen)
  - Berechnete Kategorien, 46
  - Drucken von Tabellen mit Schichten, 85
  - Kompaktansicht, 78
  - Postcompute-Kategorien, 46
  - Schichtvariablen, 81, 83, 85
  - Signifikanztests und Mehrfachantworten, 168
  - Stapeln von Variablen, 70–71
  - Teststatistiken, 148
  - Verschachteln von Schichtvariablen, 85
  - Verschachtelte Variablen, 73, 78
  - Vertauschen von Zeilen- und Spaltenvariablen, 80
- Datum
  - Einfügen des aktuellen Datums in benutzerdefinierte Tabellen, 51
- Dezimalstellen
  - Steuern der in benutzerdefinierten Tabellen angezeigten Dezimalstellen, 35, 63, 190
- Eckenbeschriftungen
  - Custom Tables (Benutzerdefinierte Tabellen), 51
- Erklärungen
  - Custom Tables (Benutzerdefinierte Tabellen), 51
- Fehlende Werte, 139, 184
  - Auswirkungen auf die Berechnung von Prozentwerten, 186

- in benutzerdefinierte Tabellen aufnehmen, 186
- Gesamtanzahl, 186
- Gesamtsummen
  - Anzeigeposition, 90
  - ausgeschlossene Kategorien, 88
  - Benutzerdefinierte Tabellen, 87
  - Custom Tables (Benutzerdefinierte Tabellen), 42
  - Gesamtsummen für Gruppen, 90
  - Randsummen in benutzerdefinierten Tabellen, 64
  - Schichten, 92
  - Tabellen, 60
  - verschachtelte Tabellen, 90
- Gesamtsummen für Gruppen, 90
- Gesamtsummen für Untergruppen, 90
- Gruppierte Auswertungen
  - Metrische Variablen, 143
- Gültige Werte, 139, 186
  - Custom Tables (Benutzerdefinierte Tabellen), 39
- Häufigkeitstabellen
  - Benutzerdefinierte Tabellen, 111
  - Custom Tables (Benutzerdefinierte Tabellen), 48
- Kreuztabelle
  - Tabellen, 61
- Leere Zellen
  - angezeigter Wert in benutzerdefinierten Tabellen, 49, 198
- Löschen von Kategorien
  - Custom Tables (Benutzerdefinierte Tabellen), 42
  - Tabellen, 65
- Marken, 213
- Maximum
  - Custom Tables (Benutzerdefinierte Tabellen), 39
- Median, 138
  - Custom Tables (Benutzerdefinierte Tabellen), 39
- Mehrfachantworten-Sets, 169
  - Doppeltantworten in Sets aus kategorialen Variablen, 49
  - Prozentwerte, 38
  - Signifikanztests, 168, 179, 181
- Messniveau
  - in benutzerdefinierten Tabellen ändern, 27
- Metrische Variablen
  - Auswertungen gruppiert nach kategorialen Zeilen- und Spaltenvariablen, 144
  - Auswertungsstatistik, 137
  - Gruppierte Auswertungen, 143
  - mehrere Auswertungsstatistiken, 138
  - Stapeln, 137
  - Verschachtelung, 147
- Minimum
  - Custom Tables (Benutzerdefinierte Tabellen), 39
- Mittelwert, 137
  - Custom Tables (Benutzerdefinierte Tabellen), 39
- Modalwert
  - Custom Tables (Benutzerdefinierte Tabellen), 39
- Postcompute-Kategorien
  - Benutzerdefinierte Tabellen, 100
  - Custom Tables (Benutzerdefinierte Tabellen), 46
- Prozente
  - Fehlende Werte, 186
  - in benutzerdefinierten Tabellen, 36, 38, 58, 62
  - Mehrfachantworten-Sets, 38
- Quellvariable für Auswertungsstatistik
  - Metrische Variablen, 147
- Rechtliche Hinweise, 212
- Reduzieren von Kategorien
  - Benutzerdefinierte Tabellen, 96
- Schichtvariablen
  - Custom Tables (Benutzerdefinierte Tabellen), 81, 83, 85
  - Drucken von Tabellen mit Schichten, 85
  - Stapeln von Schichtvariablen, 83
  - Verschachteln von Schichtvariablen, 85
- Signifikanztests
  - Custom Tables (Benutzerdefinierte Tabellen), 52
  - Mehrfachantworten-Sets, 179, 181
- Sortieren von Kategorien
  - Tabellen, 65
- Spaltenanteilestatistik
  - benutzerdefinierte Tabellen, 160
- Spaltenbreite
  - Steuern in benutzerdefinierten Tabellen, 49, 196
- Spaltenmittelstatistik
  - benutzerdefinierte Tabellen, 153
- Standardabweichung
  - Custom Tables (Benutzerdefinierte Tabellen), 39
- Stapeln von Variablen
  - Custom Tables (Benutzerdefinierte Tabellen), 70–71
  - mehrere Quellvariablen für Auswertungsstatistiken, 127
  - Metrische Variablen, 137
  - Stapeln von Schichtvariablen, 83
  - Unterschiedliche Auswertungsstatistiken für unterschiedliche Variablen, 141
- Statistik
  - Angepaßte Auswertungsstatistik für Gesamtsummen, 130
  - Auswertungsstatistik, 121
  - gestapelte Tabellen, 127
- Steuern der angezeigten Dezimalstellen, 63
- Summe
  - Custom Tables (Benutzerdefinierte Tabellen), 39

systemdefinierte fehlende Werte, 184

## Tabellen

Ändern der Beschriftung für die Auswertungsstatistik, 58

Ausblenden der Beschriftung für die Statistik, 57

Ausschließen von Kategorien, 65

Custom Tables (Benutzerdefinierte Tabellen), 27

einfache Tabellen für kategoriale Variablen, 57

Gesamtsummen, 60

Gesamtsummen in Tabellen mit ausgeschlossenen Kategorien, 65

Kreuztabelle, 61

Mehrfachantworten-Sets, 169

Prozente, 58, 62

Quelldimension der Statistik, 62

Randsummen, 64

Sortieren von Kategorien, 65

Zeilen- und Spaltenprozentwerte, 58

Tabellen für Mittelwert und Häufigkeit, 40, 130

Tabellen mit Schichten, 85

Teststatistiken

Custom Tables (Benutzerdefinierte Tabellen), 52, 148

Titel

Custom Tables (Benutzerdefinierte Tabellen), 51

Umordnen von Kategorien

Custom Tables (Benutzerdefinierte Tabellen), 42

Unterschiedliche Auswertungsstatistiken für unterschiedliche Variablen

gestapelte Tabellen, 141

Variablenlabels

Unterdrücken der Anzeige in benutzerdefinierten Tabellen, 34

Varianz

Custom Tables (Benutzerdefinierte Tabellen), 39

Verarbeitung aufgeteilter Dateien

benutzerdefinierte Tabellen, 32

Verschachtelte Variablen

Custom Tables (Benutzerdefinierte Tabellen), 73, 78

Metrische Variablen, 147

Weglassen von Kategorien

Tabellen, 65

Werte

Anzeigen von Kategoriebeschriftungen und -werten, 134

Werte und Wertelabels, 134

Zeit

Einfügen der aktuellen Uhrzeit in benutzerdefinierte Tabellen, 51

Zwischenergebnisse, 94

Ausblenden der Kategorien für das Zwischenergebnis, 96

Benutzerdefinierte Tabellen, 87

Custom Tables (Benutzerdefinierte Tabellen), 42