

IBM SPSS Modeler Server 15 Guida
della performance e amministrazione



Nota: Prima di utilizzare queste informazioni e il relativo prodotto, leggere le informazioni generali disponibili in Note a pag. .

Questa versione si applica a IBM SPSS Modeler 15 e a tutte le successive versioni e modifiche fino a eventuali disposizioni contrarie indicate in nuove versioni.

Le schermate dei prodotti Adobe sono state ristampate su autorizzazione di Adobe Systems Incorporated.

Le schermate dei prodotti Microsoft sono state ristampate su autorizzazione di Microsoft Corporation.

Materiali concessi in licenza - Proprietà di IBM

© **Copyright IBM Corporation 1994, 2012.**

Tutti i diritti riservati.

Prefazione

IBM® SPSS® Modeler è l'efficace workbench di data mining aziendale di IBM Corp.. SPSS Modeler consente alle organizzazioni di migliorare le relazioni con i clienti e con il pubblico grazie a un'analisi approfondita dei dati. Le organizzazioni potranno utilizzare le informazioni ottenute tramite SPSS Modeler per mantenere i clienti di valore, cogliere opportunità di vendite incrociate, attrarre nuovi clienti, individuare frodi, diminuire i rischi e migliorare l'offerta di servizi a livello statale.

L'interfaccia visiva di SPSS Modeler favorisce l'applicazione di una competenza aziendale specifica da parte degli utenti, grazie alla quale sarà possibile ottenere modelli di previsione più efficaci e una riduzione nei tempi di sviluppo delle soluzioni. SPSS Modeler offre una vasta gamma di tecniche di creazione di modelli, quali previsione, classificazione, segmentazione e algoritmi per l'individuazione delle associazioni. IBM® SPSS® Modeler Solution Publisher consente quindi di distribuire a livello aziendale i modelli creati in modo che vengano utilizzati dai responsabili dei processi decisionali oppure inseriti in un database.

Informazioni su IBM Business Analytics

Il software IBM Business Analytics fornisce informazioni complete, coerenti e accurate a cui i responsabili delle decisioni possono affidarsi per ottimizzare le prestazioni dell'azienda. Un ampio portafoglio di applicazioni di [business intelligence](#), [analisi predittiva](#), [gestione delle prestazioni e delle strategie finanziarie](#) e [analisi](#) offre una panoramica chiara, istantanea e interattiva delle prestazioni attuali e la possibilità di prevedere i risultati futuri. Utilizzato in combinazione con potenti soluzioni di settore, prassi consolidate e servizi professionali, questo software consente alle aziende di tutte le dimensioni di ottimizzare la produttività, automatizzare le decisioni senza problemi e fornire risultati migliori.

Come parte di questo portafoglio, il software IBM SPSS Predictive Analytics consente alle aziende di prevedere gli eventi futuri e di agire tempestivamente in modo da migliorare i risultati delle attività aziendali. Le aziende, gli enti governativi e le università di tutto il mondo si affidano alla tecnologia IBM SPSS perché rappresenta un vantaggio concorrenziale in termini di attrazione, retention e aumento dei clienti, riducendo al tempo stesso le frodi e limitando i rischi. Incorporando il software IBM SPSS nelle attività quotidiane, le aziende diventano imprese in grado di effettuare previsioni e di gestire e automatizzare le decisioni, per raggiungere gli obiettivi aziendali e vantaggi tangibili sulla concorrenza. Per ulteriori informazioni o per contattare un rappresentante, visitare il sito <http://www.ibm.com/spss>.

Supporto tecnico

Il supporto tecnico è a disposizione dei clienti che dispongono di un contratto di manutenzione. I clienti possono contattare il supporto tecnico per richiedere assistenza per l'utilizzo di IBM Corp. o per l'installazione di uno degli ambienti hardware supportati. Per contattare il supporto tecnico, visitare il sito Web IBM Corp. all'indirizzo <http://www.ibm.com/support>. Ricordare che durante la richiesta di assistenza sarà necessario fornire i dati di identificazione personali, i dati relativi alla propria società e il numero del contratto di manutenzione.

Contenuto

1 Informazioni su IBM SPSS Modeler 1

Prodotti IBM SPSS Modeler	1
IBM SPSS Modeler	1
IBM SPSS Modeler Server	2
IBM SPSS Modeler Administration Console	2
IBM SPSS Modeler Batch	2
IBM SPSS Modeler Solution Publisher	2
IBM SPSS Modeler Server Adattatori per IBM SPSS Collaboration and Deployment Services	3
Edizioni di IBM SPSS Modeler	3
Documentazione di IBM SPSS Modeler	4
Documentazione di SPSS Modeler Professional	4
Documentazione di SPSS Modeler Premium	5
Esempi di applicazioni	6
Cartella Demos	7

2 Architettura e suggerimenti relativi all'hardware 8

Architettura di IBM SPSS Modeler	8
Descrizione dell'architettura	8
Suggerimenti relativi all'hardware	10
Spazio temporaneo su disco e RAM	11
Accesso ai dati	13
Riferimento a file di dati	14
Importazione di file di dati di IBM SPSS Statistics	15
Istruzioni di installazione	15

3 Supporto per IBM SPSS Modeler 16

Connessione a IBM SPSS Modeler Server	16
Aggiunta e modifica della connessione di IBM SPSS Modeler Server	18
Ricerca di server in IBM SPSS Collaboration and Deployment Services	19
Dati e file system	20
Autenticazione degli utenti	21
Creazione di file	21
Differenze nei risultati	21

4 Amministrazione di IBM SPSS Modeler Server 23

Avvio e arresto di IBM SPSS Modeler Server	23
Per avviare, arrestare e verificare lo stato di Clementine Server in Windows	23
Per avviare, arrestare e verificare lo stato di Clementine Server in UNIX	23
Gestione dei processi di server bloccati (sistemi UNIX).	24
Amministrazione.	25
Utilizzo di IBM SPSS Modeler Administration Console	26
Avvio di Modeler Administration Console.	26
Configurazione dell'accesso tramite Modeler Administration Console	27
Connessioni a SPSS Modeler Server	28
Configurazione di SPSS Modeler Server	29
Monitoraggio di SPSS Modeler Server.	36
Utilizzo del file options.cfg	36
Chiusura delle connessioni di database inutilizzate.	37
IBM SPSS Statistics Posizione della licenza	38
Utilizzo di SSL per il trasferimento sicuro dei dati	38
Come funziona il protocollo SSL.	38
Protezione delle comunicazioni client-server e server-server con SSL.	39
Configurazione del prefisso dell'URL	43
Protezione di LDAP con SSL	44
Registro del server	44

5 Panoramica delle performance 47

Impostazioni relative alle performance e all'ottimizzazione del server	47
Impostazioni relative alle performance e all'ottimizzazione del client.	47
Ottimizzazione e utilizzo del database	50
Ottimizzazione SQL	50

6 Ottimizzazione SQL 52

Funzionamento della generazione SQL	53
Esempio di generazione SQL	54
Configurazione dell'ottimizzazione SQL	55
Anteprima di SQL generato.	56
Visualizzazione di SQL per gli insiemi di modelli.	58
Suggerimenti per l'ottimizzazione della generazione SQL	58

Nodo con supporto per generazione SQL	59
Espressioni e operatori CLEM con supporto di generazione SQL	63
Utilizzo di funzioni SQL in espressioni CLEM	65
Scrittura di query SQL	66

Appendici

A Configurazione di Oracle per piattaforme UNIX **67**

Configurazione di Oracle per l'ottimizzazione SQL	67
---	----

B Configurazione degli script di avvio di UNIX **69**

Introduzione	69
Script	69
Avvio e arresto automatico di IBM SPSS Modeler Server	70
Avvio e arresto manuale di IBM SPSS Modeler Server	70
Modifica degli script	71
Verifica delle autorizzazioni per la creazione di file	71
IBM SPSS Modeler Server e Data Access Pack	71
Risoluzione dei problemi di una configurazione ODBC	74
Percorsi delle librerie	77
Configurazione di un driver per il nodo Enterprise View	78

C Configurazione ed esecuzione di SPSS Modeler Server come processo non root in UNIX **79**

Introduzione	79
Configurazione non root senza un database privato di password	79
Configurazione non root con un database di password privato	80
Esecuzione di SPSS Modeler Server come utente non root	82
Risoluzione dei problemi di mancata autenticazione degli utenti	82

<i>D</i>	<i>Bilanciamento del carico con cluster di server</i>	<i>84</i>
<i>E</i>	<i>Note</i>	<i>86</i>
	<i>Indice</i>	<i>89</i>

Informazioni su IBM SPSS Modeler

IBM® SPSS® Modeler è un insieme di strumenti di data mining che consente di sviluppare rapidamente modelli predittivi con l'ausilio di competenze aziendali e di eseguirne il deployment nelle operazioni aziendali per migliorare i processi decisionali. Progettato secondo il modello CRISP-DM conforme agli standard di settore, SPSS Modeler supporta l'intero processo di data mining, dai dati a risultati aziendali migliori.

SPSS Modeler offre numerosi metodi di modellazione ricavati dall'apprendimento automatico, dall'intelligenza artificiale e dalla statistica. I metodi disponibili nella palette Modelli consentono di ricavare nuove informazioni dai dati e di sviluppare modelli predittivi. Ogni metodo ha determinati punti di forza e si presta meglio per particolari tipi di problemi.

SPSS Modeler può essere acquistato come prodotto autonomo oppure utilizzato come client in combinazione con SPSS Modeler Server. È inoltre disponibile una serie di opzioni, come illustrato nelle sezioni seguenti. Per ulteriori informazioni, vedere <http://www.ibm.com/software/analytics/spss/products/modeler/>.

Prodotti IBM SPSS Modeler

La famiglia di prodotti IBM® SPSS® Modeler e del software associato comprende quanto segue.

- IBM SPSS Modeler
- IBM SPSS Modeler Server
- IBM SPSS Modeler Administration Console
- IBM SPSS Modeler Batch
- IBM SPSS Modeler Solution Publisher
- IBM SPSS Modeler Server adattatori per IBM SPSS Collaboration and Deployment Services

IBM SPSS Modeler

SPSS Modeler è una versione del prodotto con funzionalità complete che viene installata ed eseguita sul proprio PC. È possibile eseguire SPSS Modeler in modalità locale come prodotto autonomo oppure in modalità distribuita assieme a IBM® SPSS® Modeler Server per ottenere una migliore performance su insiemi di dati di grandi dimensioni.

Grazie a SPSS Modeler si possono creare, in modo veloce e intuitivo, modelli predittivi accurati senza ricorrere alla programmazione. La sua avanzata interfaccia visiva permette di visualizzare con facilità il processo di data mining. Grazie alle funzionalità di analisi avanzate incorporate nel prodotto, l'utente potrà rilevare la presenza di pattern e trend, che altrimenti rimarrebbero occulti, all'interno dei dati. La modellazione dei risultati e la comprensione dei fattori che li influenzano consente di beneficiare di maggiori opportunità di business e, al contempo, di ridurre i rischi.

SPSS Modeler è disponibile in due edizioni: SPSS Modeler Professional e SPSS Modeler Premium. [Per ulteriori informazioni, vedere l'argomento Edizioni di IBM SPSS Modeler in *Manuale dell'utente di IBM SPSS Modeler 15*.](#)

IBM SPSS Modeler Server

SPSS Modeler utilizza un'architettura client/server per distribuire le richieste di operazioni che utilizzano molte risorse a potenti componenti software server, con un conseguente miglioramento della performance su insiemi di dati di grandi dimensioni.

SPSS Modeler Server è un prodotto con licenza separata che viene eseguito continuamente in modalità di analisi distribuita su un host server insieme a una o più installazioni IBM® SPSS® Modeler. Una configurazione di questo tipo consente a SPSS Modeler Server di ottenere prestazioni migliori quando si lavora su insiemi di dati di grandi dimensioni, in quanto le operazioni che richiedono un utilizzo consistente della memoria possono essere eseguite sul server senza scaricare i dati sul computer client. IBM® SPSS® Modeler Server offre inoltre il supporto delle funzionalità di ottimizzazione SQL e di modellazione in-database, garantendo ulteriori benefici dal punto di vista delle prestazioni e del livello di automazione.

IBM SPSS Modeler Administration Console

Modeler Administration Console è un'applicazione grafica per la gestione di molte delle opzioni di configurazione di SPSS Modeler Server, la cui configurazione può avvenire, inoltre, mediante un file delle opzioni. L'applicazione fornisce un'interfaccia utente di console per monitorare e configurare le installazioni di SPSS Modeler Server ed è disponibile gratuitamente per i clienti esistenti di SPSS Modeler Server. L'applicazione può essere installata solo sui computer Windows; tuttavia, può gestire un server installato su qualsiasi piattaforma supportata.

IBM SPSS Modeler Batch

Nonostante il data mining sia generalmente un processo di tipo interattivo, è possibile eseguire SPSS Modeler da una riga di comando senza il bisogno di ricorrere all'interfaccia utente grafica. Poniamo, ad esempio, che si debbano svolgere varie operazioni laboriose e ripetitive che non richiedono l'intervento di un utente. SPSS Modeler Batch è una versione speciale del prodotto che supporta l'intera gamma di funzionalità analitiche di SPSS Modeler senza richiedere l'accesso all'interfaccia utente normale. Per utilizzare SPSS Modeler Batch, è necessario disporre di una licenza SPSS Modeler Server.

IBM SPSS Modeler Solution Publisher

SPSS Modeler Solution Publisher è uno strumento che consente di creare una versione a pacchetto di uno stream SPSS Modeler che potrà essere eseguito da un motore di runtime esterno oppure incorporato in una applicazione esterna. Questo permette di pubblicare e sottoporre a deployment stream SPSS Modeler completi in ambienti in cui SPSS Modeler non è installato. SPSS Modeler Solution Publisher è distribuito come parte del servizio IBM SPSS Collaboration and Deployment

Services - Scoring, per cui è necessario procurarsi una licenza separata. Insieme alla licenza, si riceve SPSS Modeler Solution Publisher Runtime, che consente di eseguire gli stream pubblicati.

IBM SPSS Modeler Server Adattatori per IBM SPSS Collaboration and Deployment Services

È disponibile una serie di adattatori per IBM® SPSS® Collaboration and Deployment Services che abilitano l'interazione di SPSS Modeler e SPSS Modeler Server con un repository IBM SPSS Collaboration and Deployment Services. In questo modo, uno stream SPSS Modeler sottoposto a deployment sul repository potrà essere condiviso da più utenti oppure risulterà accessibile dall'applicazione thin client IBM SPSS Modeler Advantage. L'adattatore va installato sul sistema che ospita il repository.

Edizioni di IBM SPSS Modeler

SPSS Modeler è disponibile nelle edizioni seguenti.

SPSS Modeler Professional

SPSS Modeler Professional contiene tutti gli strumenti necessari per utilizzare la maggior parte dei tipi di dati strutturati, quali comportamenti e interazioni registrati in sistemi CRM, dati demografici, dati sulle vendite e sul comportamento d'acquisto.

SPSS Modeler Premium

SPSS Modeler Premium è un prodotto con licenza separata che amplia l'ambito di utilizzo di SPSS Modeler Professional aggiungendo il supporto di dati speciali, quali quelli usati per l'analisi delle entità o dei social network, e di dati di testo non strutturati. SPSS Modeler Premium comprende i seguenti componenti.

IBM® SPSS® Modeler Entity Analytics aggiunge una dimensione completamente nuova alle analisi predittive di IBM® SPSS® Modeler. Se l'analisi predittiva tenta di prevedere il comportamento futuro sulla base di dati precedenti, l'analisi dell'entità si concentra sul miglioramento della coerenza dei dati correnti risolvendo i conflitti tra gli stessi record. Un'identità può essere di un individuo, un'organizzazione, un oggetto o qualsiasi altra entità per cui possa esistere ambiguità. La risoluzione dell'identità può essere essenziale in diversi campi, tra cui la gestione delle relazioni con i clienti, il rilevamento di frodi, il riciclaggio di denaro e la sicurezza nazionale e internazionale.

IBM SPSS Modeler Social Network Analysis trasforma le informazioni sulle relazioni in campi che caratterizzano il comportamento sociale di individui e gruppi. Facendo leva sui dati che descrivono le relazioni esistenti nelle reti sociali, IBM® SPSS® Modeler Social Network Analysis riesce a individuare i leader in grado di influenzare il comportamento degli altri membri della rete. Consente inoltre di stabilire quali individui della rete sono maggiormente influenzati dagli altri membri. La combinazione di questi risultati ad altre misurazioni permette di delineare

profili complessi degli individui su cui basare dei modelli predittivi. I modelli che contengono informazioni sociali generano risultati più accurati rispetto agli altri.

Text Analytics for IBM® SPSS® Modeler utilizza tecnologie linguistiche avanzate e di Natural Language Processing (NLP) per elaborare rapidamente una grande varietà di dati di testo non strutturati, estrarre e organizzare i concetti chiave e raggruppare questi concetti in categorie. È quindi possibile combinare i concetti e le categorie estratti con dati strutturati esistenti, per esempio dati demografici, e applicarli alla modellazione utilizzando la suite completa degli strumenti di data mining di SPSS Modeler per prendere decisioni migliori e più mirate.

Documentazione di IBM SPSS Modeler

La documentazione nel formato guida in linea è disponibile nel menu Aiuto di SPSS Modeler. Sono incluse la documentazione per SPSS Modeler, SPSS Modeler Server e SPSS Modeler Solution Publisher, nonché la Guida alle applicazioni e altro materiale di supporto.

La documentazione completa in formato PDF dei singoli prodotti, istruzioni di installazione comprese, è disponibile nella cartella *Documentation* del DVD di ciascun prodotto. I documenti per l'installazione possono anche essere scaricati dal Web, all'indirizzo <http://www-01.ibm.com/support/docview.wss?uid=swg27023172>.

La documentazione in entrambi i formati è inoltre disponibile presso il Centro informazioni SPSS Modeler all'indirizzo <http://publib.boulder.ibm.com/infocenter/spssmodl/v15r0m0/>.

Documentazione di SPSS Modeler Professional

La documentazione completa di SPSS Modeler Professional, escluse le istruzioni di installazione, è la seguente.

- **Manuale dell'utente di IBM SPSS Modeler.** Introduzione generale all'utilizzo di SPSS Modeler che illustra come creare stream di dati, gestire valori mancanti, generare espressioni CLEM, utilizzare progetti e report e assemblare stream per il deployment tramite IBM SPSS Collaboration and Deployment Services, le applicazioni predittive o IBM SPSS Modeler Advantage.
- **IBM SPSS Modeler Source, Process, and Output Nodes.** Descrizioni di tutti i nodi utilizzati per leggere, elaborare e generare dati di output in vari formati, ovvero di nodi ad eccezione dei nodi Modelli.
- **IBM SPSS Modeler Nodi Modelli.** Descrizioni di tutti i nodi utilizzati per creare modelli di data mining. IBM® SPSS® Modeler offre numerosi metodi di modellazione ricavati dall'apprendimento automatico, dall'intelligenza artificiale e dalla statistica. [Per ulteriori informazioni, vedere l'argomento Panoramica sui nodi Modelli in il capitolo 3 in IBM SPSS Modeler 15 Nodi Modelli.](#)
- **IBM SPSS Modeler Algorithms Guide.** Descrizione dei fondamenti di matematica per i metodi di modellazione utilizzati in SPSS Modeler. Questa guida è disponibile solo in formato PDF.

- **IBM SPSS Modeler Guida alle applicazioni.** Gli esempi inclusi in questa guida forniscono indicazioni mirate e sintetiche su specifici metodi e tecniche di modellazione. Una versione in linea di questa guida è inoltre disponibile dal menu Aiuto. [Per ulteriori informazioni, vedere l'argomento Esempi di applicazioni in Manuale dell'utente di IBM SPSS Modeler 15.](#)
- **IBM SPSS Modeler Script e automazione.** Informazioni sulle modalità di automazione del sistema tramite script, incluse le proprietà che è possibile utilizzare per manipolare nodi e stream.
- **IBM SPSS Modeler Deployment Guide.** Informazioni sull'esecuzione di stream e scenari SPSS Modeler come fasi dell'elaborazione di lavori in IBM® SPSS® Collaboration and Deployment Services Deployment Manager.
- **IBM SPSS Modeler Guida per lo sviluppatore CLEF.** CLEF consente di integrare programmi di terze parti (quali routine di elaborazione di dati o algoritmi di modellazione) come nodi in SPSS Modeler.
- **IBM SPSS Modeler Guida alla modellazione in-database.** Informazioni sulle modalità per utilizzare al meglio la potenza del database in uso al fine di ottenere prestazioni migliori ed estendere la gamma di funzionalità analitiche tramite algoritmi di terze parti.
- **IBM SPSS Modeler Server Guida della performance e amministrazione.** Informazioni su come configurare e amministrare IBM® SPSS® Modeler Server.
- **Manuale dell'utente di IBM SPSS Modeler Administration Console.** Informazioni sull'installazione e l'utilizzo dell'interfaccia utente della console per il monitoraggio e la configurazione di SPSS Modeler Server. La console viene implementata come plug-in dell'applicazione Deployment Manager.
- **IBM SPSS Modeler Solution Publisher Guide.** SPSS Modeler Solution Publisher è un componente aggiuntivo che consente di pubblicare gli stream al di fuori dell'ambiente SPSS Modeler standard.
- **Guida CRISP-DM di IBM SPSS Modeler.** Guida passo a passo al data mining tramite la metodologia CRISP-DM con SPSS Modeler.
- **Manuale dell'utente di IBM SPSS Modeler Batch.** Guida completa all'utilizzo di IBM SPSS Modeler in modalità batch, contenente dettagli per l'esecuzione della modalità batch e gli argomenti della riga di comando. Questa guida è disponibile solo in formato PDF.

Documentazione di SPSS Modeler Premium

La documentazione completa di SPSS Modeler Premium, escluse le istruzioni di installazione, è la seguente.

- **Manuale dell'utente di IBM SPSS Modeler Entity Analytics.** Contiene informazioni per l'utilizzo dell'analisi delle entità con SPSS Modeler; descrive l'installazione e la configurazione di repository, i nodi Entity Analytics e le attività amministrative.
- **Manuale dell'utente di IBM SPSS Modeler Social Network Analysis.** Guida che spiega come eseguire l'analisi dei social network con SPSS Modeler; comprende l'analisi di gruppo e l'analisi di diffusione.

- **Manuale dell'utente di Text Analytics for SPSS Modeler.** Contiene informazioni per l'utilizzo di analisi di testo con SPSS Modeler; descrive i nodi di text mining, il workbench interattivo, i modelli e altre risorse.
- **Manuale dell'utente di Text Analytics for IBM SPSS Modeler Administration Console.** Informazioni sull'installazione e l'utilizzo dell'interfaccia utente della console per il monitoraggio e la configurazione di IBM® SPSS® Modeler Server per l'utilizzo con Text Analytics for SPSS Modeler. La console viene implementata come plug-in dell'applicazione Deployment Manager.

Esempi di applicazioni

Mentre gli strumenti per il data mining di SPSS Modeler consentono di risolvere un'ampia gamma di problemi a livello aziendale e organizzativo, gli esempi di applicazioni forniscono indicazioni mirate e sintetiche su specifici metodi e tecniche di modellazione. Gli insiemi di dati utilizzati negli esempi hanno dimensioni molto più limitate rispetto agli enormi archivi di dati gestiti da alcuni data miner, ma i concetti e i metodi coinvolti sono rapportabili alle applicazioni del mondo reale.

È possibile accedere agli esempi facendo clic su Esempi di applicazioni nel menu Aiuto di SPSS Modeler. I file di dati e gli stream di esempio sono installati nella cartella *Demos* nella directory di installazione del prodotto. [Per ulteriori informazioni, vedere l'argomento Cartella Demos in *Manuale dell'utente di IBM SPSS Modeler 15*.](#)

Esempi di modellazione in-database. Vedere gli esempi nella *IBM SPSS Modeler Guida alla modellazione in-database*.

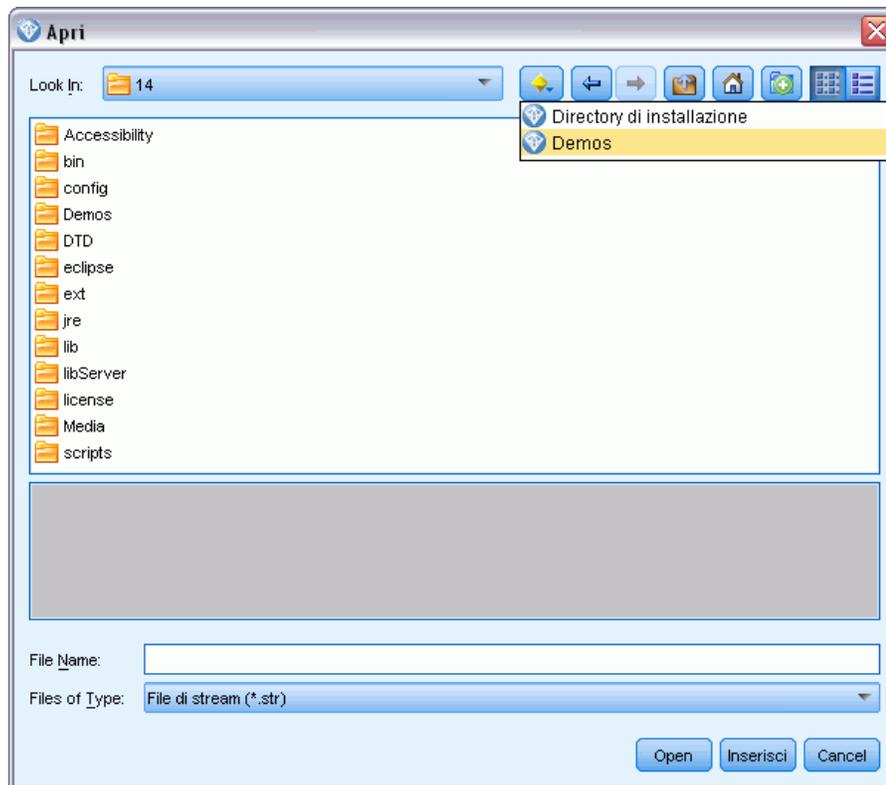
Esempi di script. Vedere gli esempi nella *IBM SPSS Modeler Guida per script e automazione*.

Cartella Demos

I file di dati e gli stream di esempio utilizzati negli esempi di applicazioni sono installati nella cartella *Demos* nella directory di installazione del prodotto. A questa cartella è possibile accedere anche dal gruppo di programmi IBM SPSS Modeler 15 nel menu Start di Windows oppure facendo clic su *Demos* nell'elenco delle directory recenti nella finestra di dialogo Apri file.

Figura 1-1

Selezione della cartella *Demos* dall'elenco delle directory utilizzate di recente



Architettura e suggerimenti relativi all'hardware

Architettura di IBM SPSS Modeler

Questa sezione descrive l'architettura di IBM® SPSS® Modeler Server, che comprende il software server, il software client e il database. Sono presentate informazioni sulla progettazione di SPSS Modeler Server, volta a ottenere prestazioni ottimali, nonché suggerimenti per ottimizzare tali prestazioni mediante la scelta di componenti hardware correttamente dimensionati. Il capitolo si conclude con una sezione dedicata all'accesso ai dati, che descrive dove impostare i driver ODBC necessari.

Descrizione dell'architettura

IBM® SPSS® Modeler Server utilizza un'architettura distribuita a tre livelli. Le operazioni del software sono condivise tra il computer client e il computer server. I vantaggi offerti dall'installazione e dall'uso di SPSS Modeler Server (rispetto a IBM® SPSS® Modeler come applicazione autonoma) sono numerosi, soprattutto quando occorre gestire insiemi di dati voluminosi:

- SPSS Modeler Server può essere eseguito su sistemi UNIX, oltre che Windows, garantendo quindi maggiori possibilità di scelta per quanto riguarda l'installazione. Su qualsiasi piattaforma è possibile destinare un computer server più veloce e potente ai processi di data mining.
- SPSS Modeler Server è ottimizzato per garantire prestazioni veloci. Quando non è possibile inviare le operazioni al database, SPSS Modeler Server archivia i risultati intermedi su disco come file temporanei anziché nella RAM. Poiché in genere i server dispongono di notevole spazio libero su disco, SPSS Modeler Server è in grado di eseguire operazioni di ordinamento, unione e aggregazione su insiemi di dati molto grandi.
- Grazie all'architettura client-server è possibile centralizzare i processi di data mining all'interno della propria organizzazione. Questo può contribuire a ufficializzare il ruolo del data mining nell'ambito dei processi aziendali.
- Con l'utilizzo di strumenti di amministrazione quali IBM® SPSS® Modeler Administration Console (fornito con SPSS Modeler Server) e IBM® SPSS® Collaboration and Deployment Services (venduto separatamente), è possibile monitorare i processi di data mining, garantendo la disponibilità di capacità di calcolo adeguate. Con IBM SPSS Collaboration and Deployment Services è possibile automatizzare alcune attività di data mining, gestire l'accesso ai modelli di dati e condividere i risultati in tutta l'organizzazione.

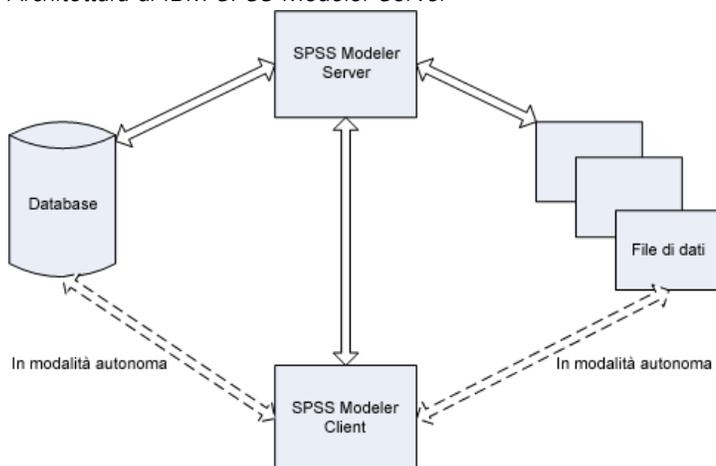
I componenti dell'architettura distribuita di IBM® SPSS® Modeler sono raffigurati nello "Architettura di IBM SPSS Modeler Server" schema a pag. 9.

- **SPSS Modeler.** Il software client, installato sul computer dell'utente finale, costituisce l'interfaccia utente e visualizza i risultati del data mining. L'applicazione client è un'installazione completa del software di SPSS Modeler il cui motore di esecuzione è però inattivo quando questo è connesso a SPSS Modeler Server per l'analisi distribuita. SPSS Modeler funziona solo su sistemi operativi Windows.
- **SPSS Modeler Server** Il software server è installato su un computer server, con connettività di rete sia alle installazioni di SPSS Modeler sia al database. SPSS Modeler Server viene eseguito come servizio (in Windows) o come processo daemon (in UNIX) che attende la connessione dei client. L'applicazione gestisce l'esecuzione degli stream e degli script creati mediante SPSS Modeler.
- **Server del database.** Il server del database può essere un data warehouse in tempo reale (ad esempio Oracle su un server UNIX) o, per ridurre l'impatto su altri sistemi operativi, un data mart su un server locale/di reparto (ad esempio, SQL Server in Windows).

Architettura di IBM SPSS Modeler Server

Figura 2-1

Architettura di IBM SPSS Modeler Server



Con l'architettura distribuita, gran parte dell'elaborazione avviene sul server. Quando l'utente finale esegue uno stream, SPSS Modeler invia al server una descrizione dello stream il server determina quali operazioni possono essere eseguite in SQL e crea le query appropriate. Le query vengono eseguite nel database e i dati risultanti vengono trasferiti al server per tutte le elaborazioni che non è possibile esprimere con SQL. Al termine dell'elaborazione, al client vengono restituiti solo i risultati pertinenti.

Se necessario, SPSS Modeler Server può eseguire tutte le operazioni di SPSS Modeler all'esterno del database. Per la memorizzazione dei dati da manipolare, l'applicazione ripartisce automaticamente in modo uniforme l'utilizzo della RAM e della memoria su disco. Grazie a questo procedimento, SPSS Modeler Server risulta totalmente compatibile con i file piatti.

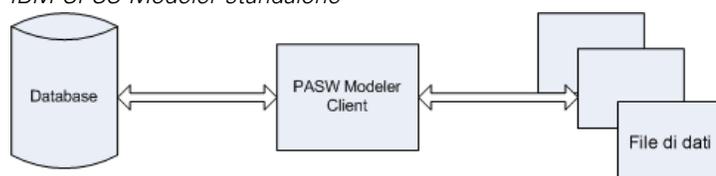
È inoltre disponibile il bilanciamento del carico utilizzando un cluster di server per l'elaborazione. L'utilizzo di cluster di server è possibile a partire da IBM SPSS Collaboration and Deployment Services 3.5 tramite il plug-in Coordinator of Processes. [Per ulteriori informazioni, vedere l'argomento Bilanciamento del carico con cluster di server in l'appendice D a pag. 84.](#) È possibile connettersi direttamente a un server o a un cluster di server gestiti in Coordinator of

Processes tramite la finestra di dialogo Login server di SPSS Modeler. [Per ulteriori informazioni, vedere l'argomento Connessione a IBM SPSS Modeler Server in il capitolo 3 in Manuale dell'utente di IBM SPSS Modeler 15.](#)

Client standalone

SPSS Modeler può essere configurato anche per essere eseguito come applicazione desktop autonoma, come illustrato nello "IBM SPSS Modeler standalone" schema sotto. [Per ulteriori informazioni, vedere l'argomento Supporto per IBM SPSS Modeler in il capitolo 3 a pag. 16.](#)

Figura 2-2
IBM SPSS Modeler standalone



Suggerimenti relativi all'hardware

Nel pianificare l'installazione di IBM® SPSS® Modeler Server è opportuno prendere in considerazione l'hardware che sarà utilizzato. Benché SPSS Modeler Server sia progettato per operare velocemente, è possibile ottimizzarne l'efficienza utilizzando componenti hardware di capacità e potenza adeguate alle attività di data mining che saranno eseguite. Un aggiornamento dell'hardware è spesso il sistema più semplice ed economico per migliorare complessivamente le prestazioni.

Server dedicato. Installare SPSS Modeler Server su un server dedicato per evitare che il programma sia costretto a condividere le risorse con altre applicazioni, compresi i database con cui SPSS Modeler Server effettua la connessione. In particolare, le operazioni di creazione dei modelli assorbono molte risorse e la loro esecuzione è molto più veloce quando non devono competere con altre applicazioni.

Nota: sebbene l'installazione di SPSS Modeler Server sullo stesso computer del database possa ridurre i tempi di trasferimento dei dati tra il database e il server evitando l'overhead di rete, nella maggior parte dei casi la migliore configurazione consiste nell'installare il server e il database su computer diversi per evitare qualsiasi competizione nella ripartizione delle risorse. Per ridurre al minimo il costo di trasferimento dei dati è opportuno stabilire una connessione veloce tra i due computer.

Processori. Il numero dei processori presenti nel computer non deve essere inferiore al numero di attività concomitanti (stream che vengono eseguiti contemporaneamente) che si prevede di eseguire con regolarità. In generale, un numero maggiore di processori garantisce migliori prestazioni.

- Una singola istanza di SPSS Modeler Server può accettare connessioni da più client (utenti), e ogni connessione client può avviare l'esecuzione di diversi stream. Pertanto, in un server possono essere in esecuzione diverse operazioni contemporaneamente.

- Come regola empirica, prevedere un processore per uno o due utenti, due processori per un massimo di quattro utenti e quattro processori per un massimo di otto utenti. Aggiungere un processore ogni due-quattro utenti aggiuntivi, a seconda delle operazioni di cui si compone il lavoro.
- Nei casi in cui è possibile rinviare al database parte dell'elaborazione mediante l'ottimizzazione SQL, può essere possibile condividere una CPU fra due o più utenti con una riduzione minima in termini di performance.
- Grazie alle funzionalità di multithreading, l'esecuzione di una singola attività può essere ripartita fra più processori: l'aggiunta di altri processori può quindi migliorare le prestazioni persino nei casi in cui venga eseguita una sola attività alla volta. In genere, il multithreading viene utilizzato per la creazione di modelli C5.0 e per alcune operazioni di preparazione dei dati (ordinamento, aggregazione e unione).

Piattaforme a 64-bit. Se si prevede di elaborare o creare modelli in base a volumi di dati molto grandi, utilizzare per SPSS Modeler Server una piattaforma Solaris, Windows o Linux a 64-bit e ottimizzare la quantità di RAM del computer. Quando gli insiemi di dati sono particolarmente consistenti, il server può esaurire rapidamente la quantità di memoria allocata a ogni processo imposta dalle piattaforme a 32-bit, costringendo il sistema a riversare i dati su disco e aumentando considerevolmente i tempi di esecuzione. Le implementazioni su server a 64-bit possono usufruire di una quantità di memoria RAM aggiuntiva; si consiglia un minimo di 8 gigabyte (GB). Il supporto a 64-bit è disponibile per le piattaforme Solaris, Windows e Linux.

Esigenze future. Nei limiti del possibile, fare in modo che la memoria e le CPU del server siano espandibili, per adeguarsi sia a un uso più intensivo (per esempio per gestire un maggior numero di utenti contemporaneamente o per far fronte all'aumento delle esigenze di elaborazione degli utenti esistenti), sia al futuro potenziamento delle funzionalità di multithreading di SPSS Modeler Server.

Spazio temporaneo su disco e RAM

IBM® SPSS® Modeler Server utilizza lo spazio temporaneo su disco per elaborare grandi volumi di dati. La quantità di spazio temporaneo necessaria dipende dalla quantità e dal tipo di dati da elaborare e dal tipo di operazioni che si eseguono. La quantità dei dati è proporzionale sia al numero di righe, *sia* al numero di colonne. Maggiore è il numero di righe e di colonne elaborate, maggiore è la quantità di spazio su disco necessaria.

Questa sezione descrive le condizioni in cui è necessario disporre di spazio temporaneo su disco e di una maggiore quantità di RAM, nonché le modalità per valutare la quantità necessaria. Si noti che questa sezione non riguarda il fabbisogno di spazio temporaneo su disco per i processi che si verificano in un database, poiché tale fabbisogno è specifico per ogni singolo database.

Condizioni che richiedono spazio temporaneo su disco

La potente funzione di ottimizzazione SQL di IBM® SPSS® Modeler Server consente di eseguire l'elaborazione all'interno del database (anziché sul server) ogniqualvolta ciò sia possibile. Tuttavia, quando si verifica una qualsiasi delle condizioni elencate di seguito, l'ottimizzazione SQL non può essere utilizzata:

- I dati da elaborare sono contenuti in un file piatto anziché in un database.

- L'ottimizzazione SQL è disattivata.
- L'operazione di elaborazione non può essere ottimizzata mediante SQL.

Nei casi in cui è impossibile utilizzare l'ottimizzazione SQL, i seguenti nodi di manipolazione dei dati e le seguenti funzioni CLEM creano copie temporanee su disco di tutti i dati o parte di essi. Se gli stream utilizzati dall'utente contengono questi comandi di elaborazione o queste funzioni, può essere necessario riservare una maggiore quantità di spazio su disco sul server.

- nodo Aggregazione
- nodo Elimina duplicati
- nodo Discretizza
- Nodo Unione quando si utilizza l'opzione di unione per chiave
- Qualsiasi nodo Modelli
- nodo Ordina
- Nodo di output Tabella
- Funzioni @OFFSET in cui la condizione di ricerca utilizza @THIS.
- Tutte le funzioni @ , per esempio @MIN, @MAX e @AVE, in cui viene calcolato il parametro della distanza.

Calcolo della quantità di spazio temporaneo su disco

In generale, IBM® SPSS® Modeler Server deve essere in grado di scrivere un file temporaneo almeno *tre volte più grande* dell'insieme di dati originario. Per esempio, se il file di dati è di 2 GB e non si utilizza la generazione di SQL, SPSS Modeler Server necessita di 6 GB di spazio su disco per elaborare i dati. Poiché ciascuno degli account utente simultanei crea i propri file temporanei, sarà necessario aumentare di conseguenza lo spazio su disco per ogni utente simultaneo.

Se la propria organizzazione utilizza spesso file temporanei di grandi dimensioni, valutare se sia opportuno utilizzare un file system distinto per i file temporanei di IBM® SPSS® Modeler, creato su un disco separato. Per ottenere i migliori risultati è possibile utilizzare un insieme di dati con striping RAID 0 ripartito su più dischi fisici in modo da accelerare le operazioni su disco, preferibilmente affidando ogni disco del file system con striping a un controller separato.

Requisiti della RAM

Per quasi tutte le operazioni di elaborazione che non è possibile eseguire nel database, IBM® SPSS® Modeler Server archivia i risultati intermedi come file temporanei su disco anziché nella memoria RAM. Tuttavia, se possibile, per i nodi Modello viene utilizzata la RAM. I nodi Rete neurale, Kohonen e K-Means impegnano grandi quantità di RAM per l'elaborazione. Se questi nodi sono di uso frequente nella propria organizzazione, può essere opportuno espandere la RAM del server.

In genere, il numero di byte di RAM necessari può essere calcolato con la formula

$(\text{numero_di_record} * \text{numero_di_celle_per_record}) * \text{numero_di_byte_per_cella}$

dove `number_of_cells_per_record` può aumentare notevolmente se sono presenti campi nominali.

Per i requisiti di RAM correnti, fare riferimento alla sezione sui requisiti di sistema nel manuale di installazione del server. Per quattro o più utenti simultanei, si consiglia di aumentare la RAM disponibile. Poiché la memoria deve essere condivisa tra attività concomitanti, è necessario aumentarla di conseguenza. In molti casi è probabile che l'espansione della memoria sia uno dei sistemi più efficaci e convenienti per migliorare le prestazioni complessive del sistema.

Accesso ai dati

Per leggere o scrivere su un database, occorre che un'origine dati ODBC sia installata e configurata per il database in questione, con le relative autorizzazioni di lettura e scrittura. IBM® SPSS® Data Access Pack contiene una serie di driver ODBC utili a questo scopo e che sono disponibili sul DVD IBM SPSS Data Access Pack o scaricabili dal sito. Per domande sulla creazione o l'impostazione di autorizzazioni per sorgenti dati ODBC, rivolgersi all'amministratore del database.

Il supporto dei database in IBM® SPSS® Modeler viene classificato in base a tre livelli, ciascuno dei quali rappresenta un diverso livello di supporto per il push back e l'ottimizzazione SQL, a seconda del produttore del database. I diversi livelli di supporto sono implementati tramite un numero di impostazioni di sistema, che possono essere personalizzate nell'ambito di un contratto di servizio.

I tre livelli di supporto database sono i seguenti:

Tabella 2-1

Livelli di supporto database

Livello di supporto	Descrizione
Livello 1	Push back SQL interamente disponibile con ottimizzazione SQL specifica per database.
Livello 2	Push back SQL per gran parte disponibile con ottimizzazione SQL non specifica per database.
Livello 3	Nessuna ottimizzazione o push back disponibile (solo supporto di lettura e scrittura database).

Driver ODBC supportati

Per le ultime informazioni sui database e i driver ODBC che sono supportati e testati per l'uso con SPSS Modeler 15, vedere le matrici di compatibilità del prodotto sul sito aziendale dedicato all'assistenza clienti (<http://www.ibm.com/support>).

Dove installare i driver

I driver ODBC devono essere installati e configurati su ogni computer utilizzato.

- Se IBM® SPSS® Modeler viene eseguito in modalità locale (standalone), i driver devono essere installati sul computer locale.
- Se SPSS Modeler viene eseguito in modalità distribuita su un IBM® SPSS® Modeler Server remoto, i driver ODBC devono essere installati nel computer in cui è installato SPSS Modeler Server. Per SPSS Modeler Server in sistemi UNIX, vedere anche "Configurazione dei driver ODBC in sistemi UNIX" più avanti in questa sezione.

- Se occorre accedere alle stesse origini dati sia da SPSS Modeler che da SPSS Modeler Server, i driver ODBC devono essere installati su entrambi i computer.
- Se SPSS Modeler è in esecuzione su Terminal Services, i driver ODBC devono essere installati sul server Terminal Services in cui è stato installato SPSS Modeler.
- Se si utilizza IBM® SPSS® Modeler Solution Publisher Runtime per eseguire flussi pubblicati su un computer a parte, occorre installare e configurare i driver ODBC anche su questo computer.

Nota: se si utilizza SPSS Modeler Server su UNIX per accedere a un database Teradata, è necessario utilizzare il gestore dei driver ODBC installato con il driver di ODBC Teradata. Per poter apportare questa modifica a SPSS Modeler Server, specificare un valore per `ODBC_DRIVER_MANAGER_PATH` all'inizio dello script `modelersrv.sh` dove indicato dai commenti. Questa variabile di ambiente deve essere impostata sul percorso del gestore dei driver ODBC, che viene fornito con il driver ODBC Teradata (`/usr/odbc/lib` in un'installazione predefinita del driver ODBC Teradata). Perché la modifica abbia effetto occorre riavviare SPSS Modeler Server. Per i dettagli relativi alle piattaforme SPSS Modeler Server che offrono supporto all'accesso Teradata, e la versione di driver ODBC Teradata supportata, vedere il sito aziendale dedicato al supporto tecnico all'indirizzo <http://www.ibm.com/support>.

Configurazione dei driver ODBC in sistemi UNIX

Di default, il gestore dei driver DataDirect non è configurato per SPSS Modeler Server nei sistemi UNIX. Per configurare UNIX per il caricamento del gestore dei driver DataDirect, immettere i comandi seguenti:

```
cd directory_installazione_server_modeler/bin
rm -f libspssodbc.so
ln -s libspssodbc_datadirect.so libspssodbc.so
```

In questo modo si rimuove il collegamento di default e si crea un collegamento al gestore dei driver DataDirect.

Nota: le regole di cui sopra sono valide esclusivamente per l'accesso ai dati di un database. Le operazioni sui file di altro tipo, quali l'apertura e il salvataggio di stream, progetti, modelli, nodi, file PMML, di output e di script vengono sempre eseguite sul lato client e sono sempre indicate come relative al file system del computer client. Inoltre, il comando Imposta directory di SPSS Modeler imposta la directory di lavoro per gli oggetti *locali* dei client (per esempio gli stream), ma non modifica la directory di lavoro del server.

UNIX. Per informazioni su come configurare SPSS Modeler Server su sistemi UNIX per il funzionamento con la tecnologia IBM® SPSS® Statistics Data Access, vedere [Configurazione degli script di avvio di UNIX a pag. 69](#).

Riferimento a file di dati

Windows. Se i dati vengono archiviati sul computer di IBM® SPSS® Modeler Server, si consiglia di assegnare il percorso ai dati in base al punto di vista del server (per esempio, `C:\DatiServer\Vendite 1998.csv`). Le prestazioni migliorano quando non viene utilizzata la rete per individuare il file.

Se i dati vengono archiviati su un host diverso, si consiglia di utilizzare riferimenti UNC per i file (per esempio, \\mioserverdati\DatiServer\Vendite 1998.csv). Si noti che i nomi UNC funzionano solo quando il percorso contiene il nome di una risorsa di rete condivisa. Il computer di riferimento deve essere autorizzato per la lettura del file specificato. Se si passa spesso dalla modalità di analisi distribuita a quella locale, è opportuno utilizzare i riferimenti UNC per i file poiché questi funzionano indipendentemente dalla modalità.

UNIX. Per fare riferimento a file di dati residenti su un server UNIX, specificare il percorso completo del file utilizzando le barre (/) (per esempio, /pubblico/dati/DatiServer/Vendite 1998.csv). Evitare l'uso della barra rovesciata nella directory di UNIX e nei nomi dei file di dati utilizzati con SPSS Modeler Server. L'utilizzo del formato UNIX o DOS in un file di testo non è rilevante, poiché entrambi i formati vengono gestiti automaticamente.

Importazione di file di dati di IBM SPSS Statistics

Se la propria organizzazione utilizza anche IBM® SPSS® Statistics Server, gli utenti possono avere necessità di importare o esportare dati SPSS Statistics quando lavorano in modalità distribuita. Tenere presente che, quando viene eseguito in modalità distribuita, IBM® SPSS® Modeler presenta il file system del server. Il client SPSS Statistics funziona in modo analogo. Per consentire l'importazione e l'esportazione tra le due applicazioni è necessario che entrambi i client funzionino nella stessa modalità. In caso contrario, i rispettivi file system saranno visualizzati in modo diverso e non sarà possibile condividere i file. I nodi SPSS Statistics di IBM® SPSS® Modeler possono avviare automaticamente il client SPSS Statistics, ma gli utenti devono prima verificare che quest'ultimo sia in esecuzione nella stessa modalità di SPSS Modeler.

Istruzioni di installazione

Per informazioni sull'installazione di IBM® SPSS® Modeler Server, vedere le istruzioni contenute nella cartella \documentation\installation\ del CD del Server. Per i sistemi Windows e UNIX sono disponibili documenti a parte.

Per informazioni dettagliate sull'installazione e sull'uso di IBM® SPSS® Modeler, consultare il CD fornito insieme all'applicazione.

Supporto per IBM SPSS Modeler

Questa sezione si rivolge agli amministratori e al personale degli help desk che forniscono supporto agli utenti di IBM® SPSS® Modeler. Vengono trattati i seguenti argomenti:

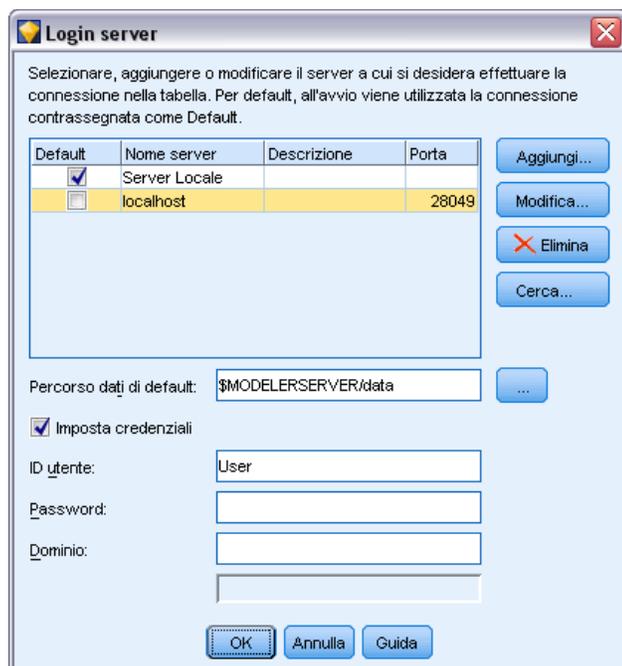
- Modalità di accesso a IBM® SPSS® Modeler Server (oppure esecuzione della versione standalone interrompendo la connessione a un server)
- Informazioni su dati e file system di cui possono necessitare gli utenti
- Account utente e autorizzazioni per i file relativi a SPSS Modeler Server
- Differenze nei risultati riscontrati dagli utenti nel passaggio da SPSS Modeler Server a IBM® SPSS® Modeler

Connessione a IBM SPSS Modeler Server

È possibile eseguire IBM® SPSS® Modeler come applicazione autonoma oppure come client connesso direttamente a IBM® SPSS® Modeler Server o a SPSS Modeler Server o a un cluster di server tramite il plug-in Coordinator of Processes disponibile in IBM® SPSS® Collaboration and Deployment Services. Lo stato di connessione corrente è visualizzato nella parte inferiore sinistra della finestra di SPSS Modeler.

Per connettersi a un server, è possibile immettere manualmente il nome del server al quale connettersi oppure selezionare un nome definito in precedenza. Tuttavia, se si dispone di IBM SPSS Collaboration and Deployment Services, è possibile cercare nell'elenco di server o cluster di server disponibili nella finestra di dialogo Login server. Il plug-in Coordinator of Processes consente di spostarsi tra i servizi Statistics in esecuzione in una rete. [Per ulteriori informazioni, vedere l'argomento Bilanciamento del carico con cluster di server in l'appendice D a pag. 84.](#)

Figura 3-1
Finestra di dialogo Login server



Per connettersi a un server

- ▶ Nel menu Strumenti, fare clic su Login server. Viene visualizzata la finestra di dialogo Login server. In alternativa, fare doppio clic sull'area dello stato di connessione nella finestra di SPSS Modeler.
- ▶ Nella finestra di dialogo specificare le opzioni per la connessione al server locale oppure selezionare una connessione dalla tabella.
 - Fare clic su Aggiungi o Modifica per aggiungere o modificare una connessione. [Per ulteriori informazioni, vedere l'argomento Aggiunta e modifica della connessione di IBM SPSS Modeler Server in Manuale dell'utente di IBM SPSS Modeler 15.](#)
 - Fare clic su Cerca per accedere al server o al cluster di server in Coordinator of Processes. [Per ulteriori informazioni, vedere l'argomento Ricerca di server in IBM SPSS Collaboration and Deployment Services in Manuale dell'utente di IBM SPSS Modeler 15.](#)

Tabella Server. Questa tabella contiene l'insieme di connessioni server definite. Nella tabella sono visualizzati la connessione di default, il nome del server, la descrizione e il numero di porta. È possibile aggiungere manualmente una nuova connessione, nonché selezionare o cercare una connessione esistente. Per impostare un determinato server come connessione di default, selezionare la casella di controllo nella colonna Default della tabella per la connessione desiderata.

Percorso dati di default. Specificare il percorso utilizzato per i dati nel server. Fare clic sul pulsante con i puntini di sospensione (...) per passare alla posizione richiesta.

Imposta credenziali. Lasciare questa casella deselezionata per abilitare la funzione di **Single Sign-On**, che cercherà di far accedere l'utente al server utilizzando il nome utente e la password del computer locale. Se la funzione di Single Sign-On non è disponibile oppure se si seleziona la casella per disabilitare Single Sign-On (per esempio, per accedere a un account amministratore), vengono attivati i seguenti campi per l'inserimento delle credenziali.

ID utente. Immettere il nome utente con il quale accedere al server.

Password. Immettere la password associata al nome utente specificato.

Dominio. Specificare il dominio utilizzato per accedere al server. È necessario specificare un nome di dominio solo quando il computer server risiede in un dominio Windows diverso da quello del computer client.

- Fare clic su OK per completare la connessione.

Per disconnettersi da un server

- Nel menu Strumenti, fare clic su Login server. Viene visualizzata la finestra di dialogo Login server. In alternativa, fare doppio clic sull'area dello stato di connessione nella finestra di SPSS Modeler.
- Nella finestra di dialogo, selezionare il Server locale e fare clic su OK.

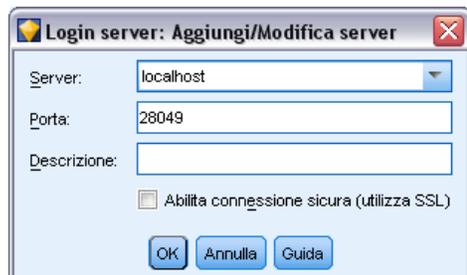
Aggiunta e modifica della connessione di IBM SPSS Modeler Server

È possibile modificare o aggiungere manualmente una connessione server nella finestra di dialogo Login server. Facendo clic su Aggiungi, è possibile accedere a una finestra di dialogo Aggiungi/Modifica server vuota in cui è possibile immettere i dettagli di connessione al server. Dopo aver selezionato una connessione esistente e aver fatto clic su Modifica nella finestra di dialogo Login server, viene visualizzata la finestra di dialogo Aggiungi/Modifica server con i dettagli della connessione da modificare.

Nota: non è possibile modificare una connessione server che è stata aggiunta da IBM® SPSS® Collaboration and Deployment Services, poiché il nome, la porta e altri dettagli sono definiti in IBM SPSS Collaboration and Deployment Services.

Figura 3-2

Finestra di dialogo Login server: Aggiungi/Modifica server



Per aggiungere connessioni server

- Nel menu Strumenti, fare clic su Login server. Viene visualizzata la finestra di dialogo Login server.

- ▶ Nella finestra di dialogo, fare clic su **Aggiungi**. Viene visualizzata la finestra di dialogo **Login server: Aggiungi/Modifica server**.
- ▶ Immettere i dettagli della connessione server e fare clic su **OK** per salvare la connessione e ritornare alla finestra di dialogo **Login server**.
 - **Server**. Specificare uno dei server disponibili o selezionarne uno dall'elenco. È possibile identificare il computer server tramite un nome alfanumerico, per esempio *serverpersonale*, oppure tramite l'indirizzo IP assegnato al computer server, per esempio 202.123.456.78.
 - **Porta**. Indica il numero di porta su cui il server rimane in ascolto. Se l'impostazione predefinita non funziona, chiedere il numero di porta corretto all'amministratore di sistema.
 - **Descrizione**. Immettere una descrizione facoltativa della connessione al server.
 - **Abilita connessione sicura (utilizza SSL)**. Specifica se è necessario utilizzare una connessione SSL (**Secure Sockets Layer**). SSL è un protocollo molto diffuso per la gestione della sicurezza dei dati trasmessi in rete. Per utilizzare questa funzione è necessario che il protocollo SSL sia stato attivato sul server di hosting di IBM® SPSS® Modeler Server. Se sono necessarie ulteriori informazioni, rivolgersi all'amministratore di sistema.

Per modificare connessioni server

- ▶ Nel menu **Strumenti**, fare clic su **Login server**. Viene visualizzata la finestra di dialogo **Login server**.
- ▶ Nella finestra di dialogo, selezionare la connessione da modificare, quindi fare clic su **Modifica**. Viene visualizzata la finestra di dialogo **Login server: Aggiungi/Modifica server**.
- ▶ Modificare i dettagli della connessione server e fare clic su **OK** per salvare le modifiche e ritornare alla finestra di dialogo **Login server**.

Ricerca di server in IBM SPSS Collaboration and Deployment Services

Anziché immettere una connessione server manualmente, è possibile selezionare un server o un cluster di server della rete tramite **Coordinator of Processes**, disponibile in **IBM® SPSS® Collaboration and Deployment Services**. Un cluster di server è un gruppo di server tra i quali **Coordinator of Processes** stabilisce quello più adatto a rispondere a una richiesta di elaborazione. [Per ulteriori informazioni, vedere l'argomento **Bilanciamento del carico con cluster di server** in l'appendice D a pag. 84.](#)

Sebbene nella finestra di dialogo **Login server** sia possibile aggiungere server manualmente, cercando i server disponibili è possibile connettersi agli stessi senza doverne conoscere il nome e il numero di porta. Queste informazioni sono infatti fornite automaticamente. È in ogni caso necessario disporre delle informazioni di accesso corrette, quali nome utente, dominio e password.

Nota: se non si dispone della funzionalità **Coordinator of Processes**, è comunque possibile immettere manualmente il nome del server a cui connettersi oppure selezionare un nome definito in precedenza. [Per ulteriori informazioni, vedere l'argomento **Aggiunta e modifica della connessione di IBM SPSS Modeler Server** in *Manuale dell'utente di IBM SPSS Modeler 15*.](#)

Figura 3-3
Finestra di dialogo Cerca server



Per cercare server e cluster

- ▶ Nel menu Strumenti, fare clic su Login server. Viene visualizzata la finestra di dialogo Login server.
- ▶ Nella finestra di dialogo, fare clic su Cerca per aprire la finestra di dialogo Cerca server. Se non si è ancora connessi a IBM SPSS Collaboration and Deployment Services quando si tenta di sfogliare Coordinator of Processes, viene richiesto di connettersi. [Per ulteriori informazioni, vedere l'argomento Connessione al repository in il capitolo 9 in Manuale dell'utente di IBM SPSS Modeler 15.](#)
- ▶ Selezionare un server o un cluster di server dall'elenco.
- ▶ Fare clic su OK per chiudere la finestra di dialogo e aggiungere la connessione alla tabella nella finestra di dialogo Login server.

Dati e file system

Gli utenti che utilizzano IBM® SPSS® Modeler Server dovranno probabilmente accedere ai file di dati e ad altre sorgenti dati attraverso la rete, nonché salvare i file in rete. A seconda dei casi, tali utenti possono necessitare delle seguenti informazioni:

- **Sorgente dati ODBC.** Se gli utenti necessitano di accedere alle sorgenti dati ODBC definite sul server, dovranno conoscere i nomi, le descrizioni e i dati per il login (compresi gli ID e le password per l'accesso ai database) relative alle sorgenti dati.
- **Accesso ai file di dati.** Se gli utenti necessitano di accedere ai file di dati sul server o su altri dispositivi in rete, dovranno conoscere i nomi e i percorsi di tali file di dati.
- **Posizione dei file salvati.** Quando salvano dei dati durante la connessione a SPSS Modeler Server, gli utenti possono tentare di salvare i file sul server. Spesso, tuttavia, la posizione di salvataggio sul server è protetta da scrittura. In questo caso, comunicare agli utenti dove devono salvare i file di dati (in genere nelle rispettive home directory).

Autenticazione degli utenti

Per autenticare gli utenti che si connettono al server, IBM® SPSS® Modeler Server utilizza il sistema operativo installato sul server stesso. Quando un utente si connette a SPSS Modeler Server, tutte le operazioni effettuate per conto dell'utente vengono eseguite nel contesto di sicurezza di quell'utente. L'accesso alle tabelle dei database deve sottostare ai privilegi di utente e/o password del database stesso.

Windows. In Windows, qualsiasi utente provvisto di un account valido sulla rete dell'host può effettuare l'accesso. Con l'autenticazione di default, gli utenti devono disporre dei diritti di accesso in modifica alla directory `<modeler_server_install>\Tmp`. Senza questi diritti, gli utenti non possono accedere a SPSS Modeler Server dal client con l'autenticazione di default di Windows.

UNIX. Si presuppone per default che in UNIX SPSS Modeler Server venga eseguito come processo root. In questo modo, qualunque utente provvisto di un account valido sulla rete dell'host può effettuare l'accesso, e l'accesso degli utenti è limitato solo ai file e alle directory di loro competenza. È possibile tuttavia configurare SPSS Modeler Server perché venga eseguito senza privilegi root. In questo caso è necessario creare un database privato di password da utilizzare per l'autenticazione e tutti gli utenti di IBM® SPSS® Modeler devono condividere un unico account utente di UNIX (e di conseguenza l'accesso ai file di dati). [Per ulteriori informazioni, vedere l'argomento Configurazione non root con un database di password privato in l'appendice C a pag. 80.](#)

Sulle piattaforme Solaris, HP-UX, Linux e AIX, SPSS Modeler Server utilizza PAM per l'autenticazione. Se necessario, è possibile utilizzare il nome del servizio `modelerserver` per configurare i moduli PAM per SPSS Modeler Server.

Creazione di file

Quando accede ai dati e li elabora, IBM® SPSS® Modeler Server deve spesso conservare una copia temporanea di quei dati su disco. La quantità di spazio su disco utilizzata per i file temporanei dipende dalle dimensioni del file di dati analizzato dall'utente finale e dal tipo di analisi eseguita. [Per ulteriori informazioni, vedere l'argomento Spazio temporaneo su disco e RAM in il capitolo 2 a pag. 11.](#)

UNIX. Le versioni UNIX di SPSS Modeler Server utilizzano il comando UNIX `umask` per impostare le autorizzazioni relative ai file temporanei. Le autorizzazioni di default del server possono essere sovrascritte. [Per ulteriori informazioni, vedere l'argomento Verifica delle autorizzazioni per la creazione di file in l'appendice B a pag. 71.](#)

Differenze nei risultati

Gli utenti che eseguono analisi in entrambe le modalità possono notare lievi differenze nei risultati ottenuti con IBM® SPSS® Modeler e IBM® SPSS® Modeler Server. La discrepanza si verifica in genere a causa dell'ordine dei record o delle differenze negli arrotondamenti.

Ordine dei record. A meno che uno stream non ordini esplicitamente i record, l'ordine in cui vengono presentati i record negli stream eseguiti localmente può variare rispetto a quelli eseguiti sul server. È inoltre possibile che esistano delle differenze di ordine tra le operazioni eseguite

in un database e quelle eseguite in SPSS Modeler Server. Tali differenze sono dovute ai diversi algoritmi utilizzati da ogni sistema per implementare le funzioni che possono riordinare i record, come per esempio l'aggregazione. Si noti, inoltre, che SQL non specifica l'ordine in cui vengono restituiti i record da un database nei casi in cui non vi sia un'operazione di ordinamento esplicita.

Arrotondamento delle differenze. Quando viene eseguito in modalità locale, IBM® SPSS® Modeler utilizza un formato interno diverso da quello di SPSS Modeler Server per archiviare i valori a virgola mobile. A causa delle differenze di arrotondamento, i risultati ottenuti con le due versioni possono variare leggermente.

Amministrazione di IBM SPSS Modeler Server

Questo capitolo contiene informazioni su come avviare e arrestare IBM® SPSS® Modeler Server, come configurarne le varie opzioni e come interpretare il file registro. Vengono descritte inoltre le modalità di utilizzo di IBM® SPSS® Modeler Administration Console, un'applicazione che semplifica la configurazione e il monitoraggio del server. Per informazioni sull'installazione di questo componente, consultare le istruzioni di installazione di SPSS Modeler Server fornite insieme al prodotto.

Avvio e arresto di IBM SPSS Modeler Server

IBM® SPSS® Modeler Server viene eseguito come servizio in Windows e come processo daemon in UNIX.

Nota per la pianificazione: l'arresto di SPSS Modeler Server provoca la disconnessione degli utenti finali e ne termina le sessioni. È opportuno quindi cercare di pianificare i riavvii del server nei periodi di minor utilizzo. Se questo non è possibile, ricordare di informare gli utenti prima di arrestare il server.

Per avviare, arrestare e verificare lo stato di Clementine Server in Windows

In Windows, IBM® SPSS® Modeler Server viene controllato tramite la finestra di dialogo Servizi del Pannello di controllo (Strumenti di amministrazione).

- ▶ **Windows XP.** Aprire il menu Start. Scegliere Impostazioni e quindi Pannello di controllo. Fare doppio clic su Strumenti di amministrazione e quindi su Servizi.

Windows 2003 o 2008. Aprire il menu Start. Scegliere Pannello di controllo, Strumenti di amministrazione, Servizi.

- ▶ Selezionare il servizio IBM SPSS Modeler Server <nn.n>. A questo punto è possibile verificarne lo stato, avviarlo o arrestarlo e modificarne i parametri di avvio, in funzione delle esigenze.

Per default, il servizio è configurato per l'esecuzione automatica: pertanto, quando viene arrestato dall'utente, sarà riavviato automaticamente al riavvio del computer. Se avviato in questo modo, il servizio viene eseguito in modo invisibile all'utente e la disconnessione dal server non influisce in alcun modo su di esso.

Per avviare, arrestare e verificare lo stato di Clementine Server in UNIX

In UNIX, IBM® SPSS® Modeler Server viene avviato e arrestato eseguendo lo script *modelersrv.sh* nella directory di installazione di SPSS Modeler Server.

- ▶ Passare alla directory di installazione di SPSS Modeler Server. Per esempio, a un prompt dei comandi di UNIX, digitare

```
cd /usr/modelersrv
```

in cui *modelersrv* è la directory di installazione di SPSS Modeler Server.

- ▶ Per avviare il server, al prompt dei comandi digitare

```
./modelersrv.sh start
```

- ▶ Per arrestare il server, al prompt dei comandi digitare

```
./modelersrv.sh stop
```

- ▶ Per verificare lo stato di SPSS Modeler Server, al prompt dei comandi di UNIX digitare

```
./modelersrv.sh list
```

ed esaminare l'output, che è simile a quello generato dal comando UNIX `ps`. Il primo processo dell'elenco è il processo daemon SPSS Modeler Server, mentre gli altri processi sono le sessioni di IBM® SPSS® Modeler.

Il programma di installazione di SPSS Modeler Server contiene uno script (*auto.sh*) che configura il sistema per avviare automaticamente il processo daemon del server all'avvio del computer. Se il server viene arrestato dopo che è stato eseguito questo script, il processo daemon del server sarà riavviato automaticamente al riavvio del computer. [Per ulteriori informazioni, vedere l'argomento Avvio e arresto automatico di IBM SPSS Modeler Server in l'appendice B a pag. 70.](#)

limiti del kernel UNIX

È necessario verificare che i limiti del kernel sul sistema siano sufficienti a consentire il funzionamento di SPSS Modeler Server. I parametri `ulimit` relativi a dati, memoria e file sono particolarmente importanti e dovrebbero essere impostati su un valore illimitato nell'ambiente SPSS Modeler Server. A questo scopo, eseguire le operazioni seguenti:

- ▶ Aggiungere i seguenti comandi a *modelersrv.sh*:

```
ulimit -d unlimited
```

```
ulimit -m unlimited
```

```
ulimit -f unlimited
```

- ▶ Riavviare SPSS Modeler Server.

Gestione dei processi di server bloccati (sistemi UNIX)

I processi di IBM® SPSS® Modeler Server possono bloccarsi per svariati motivi, per esempio, quando eseguono una chiamata di sistema o una chiamata al driver ODBC che rimane bloccata (la chiamata non ritorna o impiega molto tempo a tornare). I processi UNIX che entrano in questo stato possono essere risolti utilizzando il comando `kill` di UNIX (le interruzioni iniziate dal client IBM® SPSS® Modeler o la chiusura del client SPSS Modeler non avranno alcun

effetto). Il comando kill viene fornito come alternativa al normale comando stop e consente a un amministratore di utilizzare *modelersrv.sh* per emettere il comando kill appropriato.

Nei sistemi soggetti ad accumulo di processi di server inutilizzabili (processi “zombie”), si consiglia di arrestare e riavviare SPSS Modeler Server a intervalli regolari, utilizzando la sequenza di comandi seguente:

```
cd directory_installazione_server_modeler
./modelersrv.sh stop
./modelersrv.sh kill
```

I processi di SPSS Modeler che vengono terminati utilizzando il comando *modelersrv.sh kill* si lasciano alle spalle dei file temporanei (nella directory temporanea) che dovranno essere rimossi manualmente. I file temporanei possono essere “lasciati” nel sistema anche in altre occasioni, per esempio nei casi di arresti anomali dell’applicazione dovuti a esaurimento delle risorse, nei casi di interruzione da parte dell’utente, di arresti anomali del sistema e in altre circostanze. Pertanto, nell’ambito del processo di riavvio di SPSS Modeler Server a intervalli regolari, si consiglia anche di rimuovere tutti i file rimanenti nella directory temporanea di SPSS Modeler.

Dopo che sono stati chiusi tutti i processi server e che sono stati rimossi tutti i file temporanei, è possibile riavviare SPSS Modeler Server in tutta sicurezza.

Amministrazione

IBM® SPSS® Modeler Server dispone di numerose opzioni configurabili che controllano il modo in cui opera il programma. Tali opzioni possono essere impostate in due modi:

- Utilizzare l’applicazione IBM® SPSS® Modeler Administration Console, che è disponibile gratuitamente per i clienti IBM® SPSS® Modeler. [Per ulteriori informazioni, vedere l'argomento Utilizzo di IBM SPSS Modeler Administration Console a pag. 26.](#)
- Utilizzare il file di testo *options.cfg*, contenuto nella directory *[percorso di installazione server]/config*. [Per ulteriori informazioni, vedere l'argomento Utilizzo del file options.cfg a pag. 36.](#)

Si consiglia di installare Modeler Administration Console e di utilizzarlo come strumento di amministrazione anziché modificare il file *options.cfg*. Per la modifica di questo file è necessario accedere al file system di SPSS Modeler Server, mentre Modeler Administration Console consente di autorizzare chiunque sia provvisto di un account utente alla modifica delle opzioni configurabili. Inoltre, Modeler Administration Console fornisce informazioni aggiuntive sui processi del server, consentendo il monitoraggio dell’utilizzo e delle prestazioni. A differenza di quanto accade con il file di configurazione, quasi tutte le opzioni di configurazione possono essere modificate senza riavviare SPSS Modeler Server.

Ulteriori informazioni sull’utilizzo di Modeler Administration Console e del file *options.cfg* sono contenute nelle sezioni che seguono.

Utilizzo di IBM SPSS Modeler Administration Console

L'applicazione Modeler Administration Console fornisce un'interfaccia utente di console per monitorare e configurare le installazioni di SPSS Modeler Server ed è disponibile gratuitamente per i clienti esistenti di SPSS Modeler Server. L'applicazione può essere installata solo sui computer Windows; tuttavia, può gestire un server installato su qualsiasi piattaforma supportata.

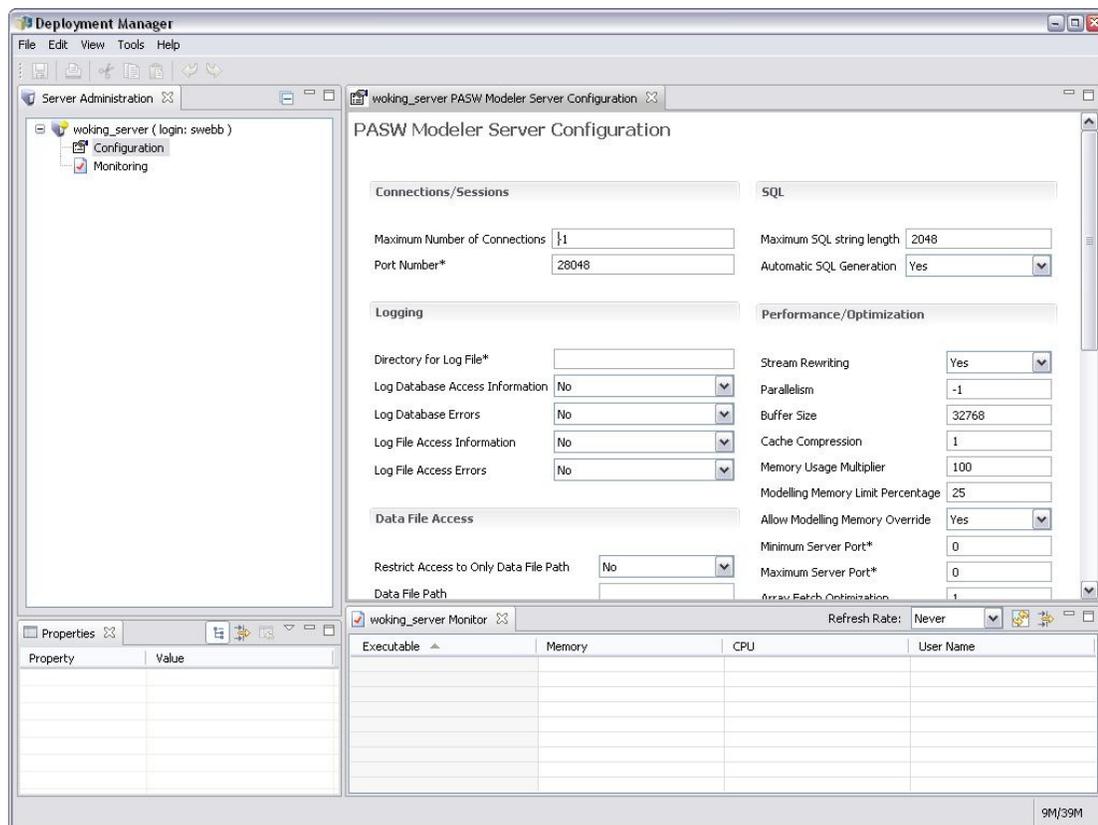
Numerose opzioni disponibili tramite Modeler Administration Console possono essere anche specificate nel file *options.cfg*, che si trova nella directory di installazione di SPSS Modeler Server nella cartella */config*. Tuttavia, l'applicazione Modeler Administration Console fornisce un'interfaccia grafica condivisa che consente all'utente di connettersi, configurare e monitorare più server.

Avvio di Modeler Administration Console

- Dal menu Start di Windows, scegliere [Tutti i] Programmi, IBM SPSS Collaboration and Deployment Services, Deployment Manager e IBM SPSS Modeler Administration Console <nn.n>.

Quando l'applicazione viene eseguita per la prima volta, i riquadri dell'amministratore del server e delle proprietà appariranno vuoti (a meno che non sia già installato Deployment Manager con una connessione server IBM SPSS Collaboration and Deployment Services già configurata). Dopo aver configurato Modeler Administration Console, nel riquadro dell'amministratore del server a sinistra viene visualizzato un nodo per ogni SPSS Modeler Server che si desidera amministrare. Il riquadro di destra mostra le opzioni di configurazione disponibili per il server selezionato. È necessario [impostare una connessione](#) per ciascuno dei server da amministrare.

Figura 4-1
Interfaccia di Modeler Administration Console



Configurazione dell'accesso tramite Modeler Administration Console

L'accesso come amministratore a SPSS Modeler Server tramite Modeler Administration Console è controllato dalla riga `administrators` del file `options.cfg`, che si trova nella directory di installazione di SPSS Modeler Server in `/config`. La riga è impostata come commento per default. Modificare la riga per consentire l'accesso a utenti specifici o utilizzare `*` per consentire l'accesso a tutti gli utenti, come negli esempi seguenti:

```
administrators, "*"
administrators, "jsmith,mjones,achavez"
```

- La riga deve iniziare con `administrators`, e le voci devono essere indicate tra virgolette. Il sistema opera una distinzione tra maiuscole e minuscole.
- Separare più ID utente mediante virgole.
- Per gli account di Windows, non utilizzare nomi di dominio.
- Prestare attenzione nell'uso dell'asterisco, poiché questo carattere consente a chiunque possieda un account utente valido per IBM® SPSS® Modeler Server (ovvero, nella maggior parte dei casi, a tutti gli utenti in rete) di accedere alle opzioni di configurazione e di modificarle.

Connessioni a SPSS Modeler Server

È necessario specificare una connessione per ogni istanza di SPSS Modeler Server da amministrare nella rete. In seguito è necessario accedere a ogni singolo server. Anche se la connessione al server viene ricordata da una sessione all'altra di Modeler Administration Console, ciò non avviene per le credenziali di accesso che devono pertanto essere inserite ogni volta che viene avviato Modeler Administration Console.

Per impostare una connessione al server

- ▶ Verificare che il servizio IBM® SPSS® Modeler Server sia stato avviato.
- ▶ Dal menu File, scegliere Nuova e quindi Connessione al server amministrato.
- ▶ Nella prima pagina della procedura guidata, digitare un nome per la connessione al server. Il nome è per uso personale e deve essere di tipo descrittivo, per esempio *Server di produzione*. Verificare che l'opzione Tipo sia impostata su IBM SPSS Modeler Server amministrato, quindi fare clic su Avanti.
- ▶ Nella seconda pagina, digitare il nome host o l'indirizzo IP del server. Se la porta di default è stata modificata, digitare il numero della porta. Fare clic su Fine. La nuova connessione al server viene visualizzata nel riquadro dell'amministratore del server.

Per eseguire le varie attività di amministrazione è ora necessario accedere al server.

Per accedere al server

- ▶ Nel riquadro dell'amministratore, fare doppio clic sul server a cui si desidera accedere per selezionarlo.
- ▶ Nella finestra di dialogo di login, digitare le proprie credenziali (utilizzare l'account utente per il server di hosting). Fare clic su OK.

Se il login non riesce e appare il messaggio Impossibile ottenere i privilegi di amministratore sul server, molto probabilmente l'accesso dell'amministratore non è stato configurato correttamente. [Per ulteriori informazioni, vedere l'argomento Configurazione dell'accesso tramite Modeler Administration Console a pag. 27.](#)

Se non è possibile eseguire l'accesso e viene visualizzato il messaggio Impossibile connettersi al server '<server>', assicurarsi che l'ID utente e la password siano corretti e verificare che il servizio SPSS Modeler Server sia in esecuzione. Per esempio, in Windows, scegliere Pannello di controllo > Strumenti di amministrazione > Servizi e verificare l'opzione per IBM SPSS Modeler Server. Se la colonna Stato non indica Avviato, selezionare la riga nella schermata e fare clic su Avvia, quindi riprovare a eseguire l'accesso.

Una volta che l'utente ha effettuato l'accesso a SPSS Modeler Server, sotto il nome del server vengono visualizzate due opzioni: [Configurazione](#) e [Monitoraggio](#). Fare doppio clic su una delle opzioni.

Configurazione di SPSS Modeler Server

Il riquadro di configurazione mostra le opzioni di configurazione di SPSS Modeler Server. Utilizzare questo riquadro per modificare le opzioni in base alle esigenze. Per salvare le modifiche, fare clic su Salva nella barra degli strumenti (per rendere effettiva la modifica del numero di porta, è necessario riavviare il server).

Le opzioni sono descritte nelle sezioni che seguono. Per ogni opzione è riportata fra parentesi la riga corrispondente del file *options.cfg*. Le opzioni visibili solo in *options.cfg* sono descritte alla fine di questa sezione.

Connessioni/sessioni

Figura 4-2
Impostazioni per connessioni e sessioni

The screenshot shows a dialog box titled "Connections/Sessions". It contains two input fields: "Maximum Number of Connections" with the value "-1" and "Port Number*" with the value "28047".

Numero massimo di connessioni. (max_sessions) Numero massimo di utenti collegati contemporaneamente. Un valore pari a -1 indica nessun limite.

Numero porta. (port_number) Il numero di porta su cui è in attesa SPSS Modeler Server. Modificarlo se la porta di default (28047) è già utilizzata da un'altra applicazione. Gli utenti finali devono conoscere il numero di porta per poter utilizzare SPSS Modeler Server.

Registrazione

Figura 4-3
Impostazioni di registrazione

The screenshot shows a dialog box titled "Logging". It contains a text field for "Directory for Log File*" and four dropdown menus for "Log Database Access Information", "Log Database Errors", "Log File Access Information", and "Log File Access Errors", all set to "No".

Directory del file di registro. (log_directory) Percorso del file *messages.log*. Una stringa vuota indica [percorso di installazione server]/log. Si noti che è necessario utilizzare le barre (/). Per informazioni sul contenuto di questo file di registro, vedere [Registro del server a pag. 44](#).

Registra informazioni di accesso al database. (log_database_access) Aggiunge al file di registro le informazioni relative all'accesso al database.

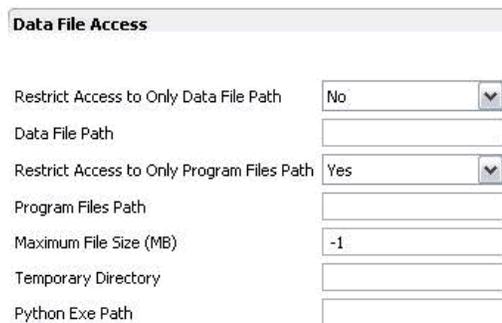
Registra errori del database. (log_database_errors) Aggiunge al file di registro gli errori del database.

Registra informazioni di accesso ai file. (log_file_access) Aggiunge al file di registro le informazioni relative all'accesso ai file.

Registra errori di accesso ai file. (log_file_errors) Aggiunge al file di registro gli errori di accesso ai file.

Accesso ai file di dati

Figura 4-4
Impostazioni per l'accesso ai file di dati



Data File Access	
Restrict Access to Only Data File Path	No
Data File Path	
Restrict Access to Only Program Files Path	Yes
Program Files Path	
Maximum File Size (MB)	-1
Temporary Directory	
Python Exe Path	

Limita l'accesso solo al percorso del file di dati. (data_files_restricted) Quando è impostata su *sì*, questa opzione limita l'accesso ai file di dati alla directory standard dei dati e a quelle elencate dall'opzione Percorso file di dati descritta di seguito.

Percorso file di dati. (data_file_path) Un elenco di directory aggiuntive in cui i client possono accedere ai file di dati in lettura e scrittura. Questa opzione viene ignorata a meno che non sia selezionata l'opzione Limita l'accesso solo al percorso del file di dati. Si noti che per specificare i percorsi è necessario utilizzare le barre (/). Se le directory sono più di una, separarne i nomi con il punto e virgola (per esempio, [percorso di installazione server]/dati;c:/dati;c:/temp).

Limita l'accesso solo al percorso dei file dei programmi. (program_files_restricted) Quando è impostata su *sì*, questa opzione limita l'accesso ai file dei programmi alla directory standard *bin* e a quelle elencate nell'opzione Percorso dei file dei programmi descritta di seguito.

Percorso dei file dei programmi. (program_file_path) Un elenco di directory aggiuntive da cui i client possono eseguire i programmi. Questa opzione viene ignorata a meno che non sia selezionata l'opzione Limita l'accesso solo al percorso dei file dei programmi. Si noti che per specificare i percorsi è necessario utilizzare le barre (/). Se le directory sono più di una, separarle con il punto e virgola.

Dimensione massima file (MB). (max_file_size) Dimensione massima dei file di dati temporanei ed esportati creati durante l'esecuzione dello stream (non valido per i file di dati SAS e SPSS Statistics). Un valore pari a -1 indica nessun limite.

Directory temporanea. (temp_directory) Directory utilizzata per archiviare file di dati temporanei (file della cache). Idealmente, questa directory dovrebbe trovarsi su un'unità o controller ad alta velocità separati, poiché la velocità di accesso a questa directory può influire sensibilmente sulle prestazioni. È possibile specificare più directory temporanee, separandole con la virgola. Le directory devono trovarsi su dischi diversi. La prima è utilizzata più di frequente, mentre quelle aggiuntive vengono utilizzate per archiviare file di lavoro temporanei quando determinate

operazioni di preparazione dei dati (per esempio, di ordinamento) utilizzano il parallelismo durante l'esecuzione. Consentendo a ogni thread di esecuzione di utilizzare dischi separati per l'archiviazione temporanea, è possibile migliorare le prestazioni. Utilizzare le barre (/) in tutte le specifiche dei percorsi. *Nota:* i file temporanei vengono creati in questa directory durante l'avvio di SPSS Modeler Server. Verificare di disporre dei diritti di accesso necessari per la directory (è possibile ad esempio che la directory temporanea sia una cartella di rete condivisa), altrimenti l'avvio di SPSS Modeler Server non verrà eseguito.

Percorso Python Exe. (`python_exe_path`) Percorso completo al file eseguibile di python che include il nome del file eseguibile. Si noti che potrebbe essere necessario impostare l'opzione Limita l'accesso solo al percorso dei file dei programmi su *No* a seconda della posizione in cui è installato python.

Performance/ottimizzazione

Figura 4-5
Impostazioni relative a performance e ottimizzazione

Performance/Optimization	
Stream Rewriting	Yes
Parallelism	-1
Buffer Size	32768
Cache Compression	1
Memory Usage Multiplier	100
Modelling Memory Limit Percentage	25
Allow Modelling Memory Override	Yes
Minimum Server Port	0
Maximum Server Port	0
Array Fetch Optimization	1

Riscrittura degli stream. (`stream_rewriting_enabled`) Consente al server di ottimizzare gli stream riscrivendoli. Per esempio, è possibile che il server spinga le operazioni di riduzione dei dati più vicino al nodo di input allo scopo di ridurre al minimo le dimensioni dell'insieme di dati il più tempestivamente possibile. In genere, si consiglia di disattivare questa opzione solo se l'ottimizzazione provoca un errore oppure risultati imprevisti. Questa impostazione prevale sulla corrispondente impostazione dell'ottimizzazione client.

Parallelismo. (`max_parallelism`) Descrive il numero di thread di lavoro paralleli che SPSS Modeler può utilizzare per l'esecuzione di uno stream. Se si imposta questa opzione su 0 o su un numero negativo, si indica a IBM® SPSS® Modeler che il numero dei thread deve corrispondere al numero dei processori disponibili sul computer; il valore di default di questa opzione è -1. Per disattivare l'elaborazione parallela (per i computer multiprocessore) impostare questa opzione su 1. Per consentire un'elaborazione parallela limitata, impostare l'opzione su un numero inferiore a quello dei processori presenti nel computer. Si noti che un processore hyper-threaded o dual-core viene considerato come due processori.

Dimensioni buffer (byte). (`io_buffer_size`) I file di dati trasferiti dal server al client vengono smistati attraverso un buffer avente il numero di byte specificato.

Compressione cache. (`cache_compression`) Numero intero compreso tra 0 e 9 che controlla la compressione dei file della cache e di altri file nella directory temporanea del server. La compressione riduce lo spazio utilizzato su disco, funzionalità importante quando lo spazio disponibile è limitato, e migliora le prestazioni riducendo il numero di attività su disco necessarie per creare e leggere la cache. La compressione aumenta il tempo del processore, ma questo aumento è quasi sempre compensato dal minor tempo di accesso al disco. Si noti che è possibile comprimere solo determinate cache, ovvero quelle con accesso sequenziale. Questa opzione non si applica a cache ad accesso casuale, come quelle utilizzate dagli algoritmi di addestramento della rete. L'impostazione 0 disattiva completamente la compressione. I valori a partire da 1 forniscono gradi crescenti di compressione, ma con un corrispondente costo in termini di tempo di accesso. Il valore di default è 1. Raramente sono necessari valori superiori, ad eccezione di quando lo spazio su disco è estremamente limitato.

Moltiplicatore utilizzo memoria. (`memory_usage`) Controlla la quantità di memoria fisica assegnata all'ordinamento e altre cache in memoria. L'impostazione di default è 100, corrispondente a circa il 10% di memoria fisica. Aumentare questo valore per migliorare la performance di ordinamento quando è disponibile memoria libera, ma prestare attenzione a non aumentarlo troppo per evitare un paging eccessivo.

Percentuale limite memoria di modellazione. (`modelling_memory_limit_percentage`) Controlla la quantità di memoria fisica assegnata all'addestramento di modelli Rete neurale, Kohonen e *K*-Means. L'impostazione di default è 25%. Aumentare questo valore per migliorare la performance di addestramento quando è disponibile memoria libera, ma prestare attenzione a non aumentarlo troppo per evitare un paging eccessivo quando i dati si riversano su disco.

Consenti sovrascrittura memoria di modellazione. (`allow_modelling_memory_override`) Attiva o disattiva l'opzione Ottimizza per velocità in determinati nodi Modelli. Per default, l'opzione è attiva. Questa opzione consente all'algoritmo di modellazione di richiedere tutta la memoria disponibile, ignorando l'opzione di limite percentuale. Può essere opportuno disattivarla se è necessario condividere risorse di memoria sul computer server.

Porta del server minima e Porta del server massima. (`max_server_port` e `min_server_port`) Specifica l'intervallo di numeri di porta che può essere utilizzato per le connessioni socket aggiuntive tra client e server richieste per i modelli interattivi e per l'esecuzione degli stream. Tali casi richiedono che il server rimanga in ascolto su un'altra porta. Se non si limita l'intervallo, potrebbero verificarsi dei problemi per gli utenti che utilizzano sistemi con firewall. Il valore di default è in entrambi i casi -1, ovvero "nessuna limitazione". Pertanto, per impostare il server affinché resti in ascolto sulla porta 8000 o superiore, si imposterà `min_server_port` su 8000 e `max_server_port` su -1.

Tenere presente che è necessario aprire porte aggiuntive sul server principale per poter aprire o eseguire uno stream e di conseguenza più porte se si desidera aprire o eseguire stream in contemporanea. Questa operazione è necessaria per catturare il feedback dall'esecuzione degli stream.

Per default, SPSS Modeler utilizzerà qualsiasi porta aperta disponibile. Se non ne trova una disponibile, per esempio se sono tutte chiuse da un firewall, quando si esegue lo stream viene visualizzato un errore. Per configurare l'intervallo delle porte, SPSS Modeler avrà bisogno di due porte aperte disponibili (oltre alla porta server principale) per ciascuno stream contemporaneo, più una porta aggiuntiva per ciascun client connesso che abbia una connessione ODBC. Per eseguire

un numero maggiore di stream contemporaneamente, è possibile ampliare l'intervallo di numeri di porta e aprire di conseguenza più porte sul firewall.

Nota: se si modificano questi parametri, è necessario riavviare SPSS Modeler Server affinché le modifiche abbiano effetto.

Ottimizzazione recupero matrice. (`sql_row_array_size`) Controlla il modo in cui SPSS Modeler Server recupera i dati dalla sorgenti dati ODBC. Il valore di default è 1, ovvero viene recuperata una riga per volta. Se si aumenta questo valore, il server leggerà le informazioni in blocchi più estesi, recuperando il numero di righe specificato in una matrice. Con alcune combinazioni sistema operativo/database, ciò può migliorare le performance delle istruzioni SELECT.

SQL

Figura 4-6
Impostazioni SQL

The screenshot shows a dialog box titled "SQL". It contains two settings:

- "Maximum SQL string length" is a text input field containing the value "2048".
- "Automatic SQL Generation" is a dropdown menu currently set to "Yes".

Lunghezza massima stringa SQL. (`max_sql_string_length`) Per una stringa importata dal database con SQL, numero massimo di caratteri dei quali viene garantito il passaggio. A seconda del sistema operativo, le stringhe di lunghezza maggiore potrebbero essere troncate a destra senza alcuna segnalazione. L'intervallo valido è compreso tra 1 e 65.535 caratteri.

Generazione automatica SQL. (`sql_generation_enabled`) Consente la generazione SQL automatica di stream con un possibile miglioramento sostanziale delle prestazioni. Per default, l'opzione è attiva. Si consiglia di disattivarla solo se il database non è in grado di supportare query inviate da SPSS Modeler Server. Si noti che questa impostazione prevale sull'impostazione di ottimizzazione client corrispondente. Ai fini del calcolo del punteggio, si tenga anche presente che è necessario attivare la generazione SQL separatamente per ogni nodo Modelli, indipendentemente da questa impostazione.

SSL

Figura 4-7
Impostazioni SSL

The screenshot shows a dialog box titled "SSL". It contains four settings:

- "Enable SSL" is a dropdown menu currently set to "No".
- "Certificate File" is an empty text input field.
- "Private Key File" is an empty text input field.
- "Private Key Password" is an empty text input field.

Attiva SSL. (`ssl_enabled`) Attiva la crittografia SSL per le connessioni tra SPSS Modeler e SPSS Modeler Server.

File di certificato. (ssl_certificate_file) Il file di certificato SSL pubblico da caricare all'avvio del server.

File di chiave privata. (ssl_private_key_file) Il file della chiave SSL privata da caricare all'avvio del server.

Password di chiave privata. (ssl_private_key_password) Chiave password del certificato specificato. Si noti che la chiave deve essere stata firmata da almeno una Autorità di certificazione elencata nel file del certificato configurato.

Configurazione di Coordinator of Processes

Figura 4-8
Impostazioni di configurazione di Coordinator of Processes

Co-ordinator Of Processes Configuration	
Host*	spsscop
Port Number*	8080
Login Name*	
Password*	
Enabled*	No
Server Name*	
Description*	
Update Interval (min)*	2
Weight	1
Service Host*	

Items marked (*) require a server restart to take effect.

Host. (cop_host) Nome host o indirizzo IP del servizio Coordinator of Processes. Il valore di default “spsscop” è un alias che gli amministratori possono decidere di aggiungere in DNS per l'host IBM SPSS Collaboration and Deployment Services.

Numero porta. (cop_port_number) Numero di porta del servizio Coordinator of Processes. L'impostazione di default di IBM SPSS Collaboration and Deployment Services è 8080.

Login name. (cop_user_name) Nome utente per l'autenticazione al servizio Coordinator of Processes. È un nome di login IBM SPSS Collaboration and Deployment Services che potrebbe includere il prefisso di un provider di protezione (per esempio, ad/jsmith).

Password. (cop_password) Password crittografata per l'autenticazione al servizio Coordinator of Processes.

Per creare il file delle password è necessario utilizzare l'apposito programma di utilità, *pwutil*, contenuto nella directory *bin* dell'installazione di SPSS Modeler Server.

Enabled. (cop_enabled) Stabilisce se il server deve tentare di registrarsi con il servizio Coordinator of Processes. L'impostazione di default è quella di *non* registrarsi perché l'amministratore deve scegliere quali servizi sono pubblicizzati tramite Coordinator of Processes.

Server name. (cop_service_name) Nome dell'istanza di SPSS Modeler Server; il valore di default è il nome host.

Descrizione. (cop_service_description) Descrizione dell'istanza.

Update interval (min). (cop_update_interval) Numero di minuti tra i messaggi keep-alive; il valore di default è 2.

Peso. (cop_service_weight) Peso dell'istanza, indicato come numero intero compreso tra 1 e 10. Un peso maggiore attira più connessioni. Il valore di default è 1.

Service host. (cop_service_host) Nome host completo dell'host IBM® SPSS® Modeler Server. Il valore di default del nome host è derivato automaticamente; l'amministratore può ignorarlo nel caso di host multihomed.

Opzioni visibili solo in options.cfg

administrators. Specificare il nome utente degli utenti a cui si desidera concedere l'accesso come amministratore. [Per ulteriori informazioni, vedere l'argomento Configurazione dell'accesso tramite Modeler Administration Console a pag. 27.](#)

allow_config_custom_overrides. Non modificare se non esplicitamente richiesto da un addetto del supporto tecnico.

db_udf_enabled. Se impostato su Y (default), porta l'opzione di generazione SQL a generare SQL con funzioni definite dall'utente (UDF, user-defined function), anziché semplice SQL SPSS Modeler. Solitamente SQL UDF offre prestazioni migliori del semplice SQL.

default_sql_string_length. Specifica la larghezza di default delle colonne di stringhe che saranno create nelle tabelle di cache del database. I campi stringa delle tabelle di cache del database saranno creati con una larghezza di default di 255 in assenza di informazioni relative al tipo a monte. Se i dati hanno larghezze superiori, istanziare un nodo Tipo a monte con i valori corrispondenti o impostare questo parametro su un valore grande abbastanza da contenere questi valori di stringa.

max_transfer_size. Solo per uso interno del sistema; non modificare.

shell. (Solo per server UNIX) Ignora l'impostazione di default della shell UNIX, per esempio shell, "/usr/bin/ksh". Per default, IBM® SPSS® Modeler utilizza la shell definita nel profilo utente dell'utente che si connette a IBM® SPSS® Modeler Server.

start_process_as_login_user. Impostare l'opzione su Y se si utilizza Single Sign-On e si è connessi a un server IBM® SPSS® Collaboration and Deployment Services che non utilizza lo stesso metodo di autenticazione dell'host SPSS Modeler Server. Inoltre, impostare l'opzione su Y se si esegue SPSS Modeler Server come processo non root in UNIX senza un database privato di password.

supported_sql_type_logging. Solo per scopi di debug; non modificare se non esplicitamente richiesto da un addetto del supporto tecnico.

time_ecode_execution_log. Impostare questa opzione su Y se il file *message.log* (nella cartella *log* dell'installazione SPSS Modeler Server) deve includere i singoli tempi di esecuzione di ogni nodo dello stream.

Monitoraggio di SPSS Modeler Server

Il riquadro di monitoraggio di Modeler Administration Console visualizza un'istantanea di tutti i processi in corso sul computer di SPSS Modeler Server, in modo analogo al Task Manager di Windows. Per attivare il riquadro di monitoraggio, fare doppio clic sul nodo Monitoraggio sotto al server desiderato nel riquadro dell'amministratore. Nel riquadro viene così visualizzata un'istantanea aggiornata dei dati del server. I dati vengono aggiornati alla velocità indicata (per default, un minuto). Per aggiornare manualmente i dati, fare clic sul pulsante Aggiorna. Per visualizzare nell'elenco solo i processi di SPSS Modeler Server, fare clic sul pulsante Escludi i processi non SPSS Modeler.

Figura 4-9
Riquadro di monitoraggio di IBM SPSS Modeler Administration Console

Executable	Memory	CPU	User Name
FrameworkService	13,036 K	00	SYSTEM
LocalSch	1,640 K	00	SYSTEM
Mcshield	23,364 K	00	SYSTEM
SoftMon	1,632 K	00	SYSTEM
System	176 K	00	SYSTEM
UpdaterUI	2,156 K	00	cstrahs
VsTskMgr	464 K	00	SYSTEM
clemserv	10,452 K	00	SYSTEM
clemserv	22,312 K	00	SYSTEM

Utilizzo del file *options.cfg*

Il file *options.cfg* è contenuto nella directory *[percorso di installazione server]/config*. Ogni impostazione è rappresentata da una coppia nome-valore separata da virgola, in cui il **nome** è il nome dell'opzione e il **valore** è il valore dell'opzione. Il simbolo del cancelletto (#) indica i commenti.

Nota: La maggior parte delle opzioni di configurazione può essere modificata utilizzando IBM® SPSS® Modeler Administration Console anziché il file di configurazione, con alcune eccezioni. [Per ulteriori informazioni, vedere l'argomento Opzioni visibili solo in options.cfg a pag. 35.](#)

Con l'utilizzo di Modeler Administration Console si evita di riavviare il server per rendere effettiva la modifica delle opzioni (fatta eccezione per il cambio della porta). [Per ulteriori informazioni, vedere l'argomento Utilizzo di IBM SPSS Modeler Administration Console a pag. 26.](#)

Opzioni di configurazione che è possibile aggiungere al file di default

Per default, il caching in-database viene attivato con IBM® SPSS® Modeler Server. È possibile disattivare questa funzione aggiungendo la riga seguente al file *options.cfg*:

```
enable_database_caching, N
```

Questa operazione determina la creazione dei file temporanei sul server anziché nel database.

Per visualizzare o modificare le opzioni di configurazione di IBM SPSS Modeler Server

- ▶ Aprire il file *options.cfg* con un editor di testo.
- ▶ Individuare le opzioni da modificare. Per un elenco completo delle opzioni, vedere [Configurazione di SPSS Modeler Server a pag. 29](#).
- ▶ Modificare i valori nel modo desiderato. Si noti che nei valori relativi ai percorsi è necessario utilizzare come separatore la barra (/) anziché la barra rovesciata.
- ▶ Salvare il file.
- ▶ Arrestare e riavviare SPSS Modeler Server per rendere effettive le modifiche. [Per ulteriori informazioni, vedere l'argomento Avvio e arresto di IBM SPSS Modeler Server a pag. 23](#).

Chiusura delle connessioni di database inutilizzate

Di default, IBM® SPSS® Modeler salva nella cache almeno una connessione a un database dopo che è stato eseguito l'accesso a quella connessione. La sessione del database rimane aperta anche quando non vengono eseguiti stream che richiedono l'accesso al database.

La memorizzazione delle connessioni di database può accelerare i tempi di esecuzione in quanto fa sì che SPSS Modeler non debba riconnettersi al database ogni volta che viene eseguito uno stream. Tuttavia, in alcuni ambienti, è importante che le applicazioni liberino le risorse del database non appena possibile. Se troppe sessioni di SPSS Modeler mantengono connessioni inutilizzate al database, le risorse del database possono esaurirsi.

Per evitare questo problema, è possibile disattivare l'opzione *cache_connection* di SPSS Modeler nel file di configurazione di un database personalizzato. Questa operazione può anche rendere SPSS Modeler meno soggetta a errori di connessione al database (per esempio, i timeout di sessione) che possono verificarsi quando le connessioni vengono utilizzate per un lungo periodo da una sessione di SPSS Modeler.

Per fare in modo che le connessioni di database inutilizzate vengano chiuse:

- ▶ Individuare la directory *[percorso installazione server]/config*.
- ▶ Aggiungere il seguente file (o aprirlo se esiste già):
odbc-custom-properties.cfg
- ▶ Aggiungere la seguente riga al file:
cache_connection, N
- ▶ Salvare e chiudere il file.
- ▶ Riavviare IBM® SPSS® Modeler Server in modo che le modifiche diventino effettive.

IBM SPSS Statistics Posizione della licenza

Per consentire a IBM® SPSS® Modeler di utilizzare i nodi Trasformazioni, Modelli e Output Statistics, è necessario disporre di una copia di IBM® SPSS® Statistics installata e concessa in licenza sul computer dove viene eseguito lo stream. Quando viene eseguito su server, la configurazione della licenza deve essere effettuata sul server.

Windows

- ▶ Dal prompt dei comandi, passare alla directory bin di IBM® SPSS® Modeler Server ed eseguire:

```
statisticsutility -location =<percorso del file licenza di IBM SPSS Statistics Server>
```

in cui <path to SPSS Statistics Server license file> è la directory di installazione di una copia di SPSS Statistics Server dotata di licenza.

UNIX

- ▶ Dal prompt dei comandi, passare alla directory bin di SPSS Modeler Server ed eseguire:

```
./statisticsutility -location =<percorso del file licenza di IBM SPSS Statistics Server>
```

in cui <path to SPSS Statistics Server license file> è la directory di installazione di una copia di SPSS Statistics Server dotata di licenza.

Utilizzo di SSL per il trasferimento sicuro dei dati

SSL (Security Sockets Layer) è un protocollo per la crittografia dei dati trasferiti tra due computer. SSL assicura che la comunicazione tra i computer sia protetta. SSL può crittografare l'autenticazione di un nome utente/password e il contenuto di uno scambio tra un server e un client.

Come funziona il protocollo SSL

SSL si basa sulle chiavi pubbliche e private del server, oltre a un certificato di chiave pubblica che collega l'identità del server alla relativa chiave pubblica.

- ▶ Quando un client si connette a un server, il client autentica il server con il certificato di chiave pubblica.
- ▶ Il client genera quindi un numero casuale, lo crittografa con la chiave pubblica del server e invia il messaggio crittografato al server.
- ▶ Il server decrittografa il numero casuale tramite la chiave privata.
- ▶ Dal numero casuale, sia il server sia il client creano le chiavi di sessione utilizzate per crittografare e decrittografare le informazioni successive.

Il certificato di chiave pubblica viene firmato solitamente da un'autorità di certificazione. Le autorità di certificazione, quali VeriSign e Thawte, sono organizzazioni che emettono, autenticano e gestiscono le credenziali di sicurezza contenute nei certificati di chiave pubblica. In sintesi

L'autorità di certificazione conferma l'identità del server. Solitamente l'autorità di certificazione applica una tariffa per l'emissione di un certificato, ma è possibile generare anche certificati autofirmati.

Protezione delle comunicazioni client-server e server-server con SSL

Le fasi principali per garantire la protezione delle comunicazioni client-server e server-server con SSL sono:

- ▶ Installare OpenSSL sul server.
- ▶ Ottenere e installare il certificato e le chiavi SSL.
- ▶ Abilitare e configurare SSL nell'applicazione di amministrazione del server (IBM® SPSS® Collaboration and Deployment Services Deployment Manager).
- ▶ Se lo si desidera, installare la crittografia "Unlimited strength" sui computer client.
- ▶ Se si utilizza un certificato autofirmato, copiare il certificato sul computer client.
- ▶ Aggiungere il certificato all'archivio chiavi client.
- ▶ Istruire gli utenti finali su come abilitare SSL nelle connessioni al server.

Nota: può capitare talvolta che un server si comporti come un client. Un esempio è IBM® SPSS® Statistics Server che si connette al IBM® SPSS® Collaboration and Deployment Services Repository. In questo caso, SPSS Statistics Server è il *client*.

Installazione di OpenSSL

Se sul server non è già disponibile OpenSSL, è necessario installarlo.

- ▶ Scaricare OpenSSL da <http://www.openssl.org/>. Assicurarsi di utilizzare la versione di OpenSSL adatta alla versione del server:

Tabella 4-1
Versioni OpenSSL

Prodotto server	Versione OpenSSL compatibile
IBM® SPSS® Statistics 20	1.0.0
SPSS Statistics 17-19 (non Linux® su System z®) IBM® SPSS® Modeler 15 (non Linux® su System z®)	0.9.8 e le relative sottoversioni (0.9.8a, 0.9.8b e così via)
SPSS Statistics 19 (Linux® su System z®) SPSS Modeler 15 (Linux® su System z®)	1.0.0

Nota: Quando è configurato per SSL, IBM® SPSS® Modeler Server per UNIX tenta di collegarsi a due librerie con nomi come (il suffisso *.so* può essere specifico di una determinata piattaforma):

- *libssl.so.0.9.8*
- *libcrypto.so.0.9.8*

Quando si utilizza una sottoversione di SSL (0.9.8a, 0.9.8b, ecc.) è necessario creare collegamenti simbolici per queste librerie con le versioni *xxx.0.9.8* richieste da SPSS Modeler Server.

- ▶ Seguire le istruzioni per l'installazione e la configurazione del software. Se si crea autonomamente OpenSSL, attenersi alle seguenti linee guida:

Windows. OpenSSL deve essere creato con DLL (che sono multithread di default).

UNIX. OpenSSL deve supportare i thread multipli (tale supporto non è sempre di default) e le librerie condivise.

- ▶ Assicurarsi che il sistema sia in grado di trovare le librerie SSL:

Windows. Assicurarsi che i moduli OpenSSL siano inclusi nel percorso di sistema.

UNIX. Aggiungere le librerie SSL alla fine del percorso della libreria di caricamento definito in *modelersrv.sh*. Ad esempio, su Solaris, aggiungere la seguente riga allo script di avvio del server:

```
LD_LIBRARY_PATH_64=$LD_LIBRARY_PATH_64:/usr/local/openssl; export LD_LI  
BRARY_PATH_64
```

Ottenere e installare il certificato e le chiavi SSL

- ▶ Ottenere un file di certificato e chiave SSL. Per eseguire questa operazione è possibile procedere in due modi.
 - Acquistarlo da un'autorità di certificazione pubblica, quale VeriSign o Thawte. L'autorità di certificazione pubblica firma il certificato per verificare il server che lo utilizza.
 - Generare i file della chiave e di certificato con un'autorità di certificati autofirmati interna. OpenSSL fornisce un apposito strumento di gestione dei certificati, oppure è possibile cercare su Internet le istruzioni per creare un certificato SSL autofirmato.
- ▶ Copiare il file o i file della chiave e del certificato in una directory locale sul server. Le chiavi pubbliche e private possono essere archiviate in un file. Assicurarsi che le autorizzazioni di accesso neghino la navigazione casuale di questa directory.
- ▶ Installare il certificato e le chiavi SSL sul server applicazioni. Per ulteriori informazioni su come le chiavi e il certificato operano con un server applicazioni specifico, vedere la documentazione del fornitore originale. Tenere presente che potrebbe venire richiesto di aggiungere il certificato all'archivio chiavi Java.

Attivazione e configurazione di SSL in IBM SPSS Collaboration and Deployment Services Deployment Manager

- ▶ Avviare l'applicazione di amministrazione del server (IBM® SPSS® Collaboration and Deployment Services Deployment Manager) e connettersi al server.
- ▶ Nella pagina di configurazione, impostare Secure Sockets Layer su Sì.
- ▶ In File di chiave pubblica SSL, specificare il percorso completo al file della chiave pubblica.
- ▶ In File di chiave privata SSL, specificare il percorso completo al file della chiave privata.

Nota: se le chiavi pubblica e privata sono archiviate in un solo file, specificare lo stesso file in File di chiave pubblica SSL e File di chiave privata SSL.

- ▶ Dai menu, scegliere:
File > Salva
- ▶ Riavviare il servizio o il daemon server. Al riavvio, verrà richiesta la password SSL. In Windows, è possibile selezionare Memorizza password per archiviare la password in modo sicuro. Questa opzione elimina la necessità di inserire la password ogni volta che il server viene avviato.

Configurazione di SSL in IBM SPSS Modeler Server

Il modo in cui si fornisce la password per l'accesso al file di chiave privata SSL, se richiesta, dipende dal sistema operativo.

Sistemi Windows

Nei sistemi Windows è possibile fornire la password in uno dei seguenti modi:

- Specificarla nel file *options.cfg* utilizzando l'opzione `ssl_private_key_password`. [Per ulteriori informazioni, vedere l'argomento SSL a pag. 33.](#)
- Specificarla mediante una finestra del prompt. Se si omette l'opzione `ssl_private_key_password` dal file *options.cfg*, la finestra del prompt viene visualizzata all'avvio.

Sistemi UNIX

Nei sistemi UNIX è possibile fornire la password in uno dei seguenti modi:

- Specificarla nel file *options.cfg* utilizzando l'opzione `ssl_private_key_password`. [Per ulteriori informazioni, vedere l'argomento SSL a pag. 33.](#)
- Aggiungere un'opzione al file *modelersrv.sh* per richiedere la password all'avvio.

Se si utilizza il secondo modo, individuare la riga seguente nel file *modelersrv.sh*:

```
if "$INSTALLEDPATH/$SCLEMDNAME" -server $ARGS; then
```

Aggiungere l'opzione `-request_ssl_password` come segue:

```
if "$INSTALLEDPATH/$SCLEMDNAME" -request_ssl_password -server $ARGS; then
```

Installazione della crittografia "Unlimited strength"

Nel Java Runtime Environment, fornito con il prodotto, è attivata la cosiddetta crittografia "US export-strength". Per una maggiore protezione dei dati, si consiglia di aggiornarla alla crittografia "Unlimited strength".

IBM J9

- ▶ Scaricare i file Java Cryptography Extension (JCE) Unlimited Strength Jurisdiction Policy relativi alla propria versione SDK dal sito Web IBM.com.

- ▶ Estrarre i file di Unlimited Jurisdiction Policy dal file ZIP. Il file ZIP contiene un file *US_export_policy.jar* e un file *local_policy.jar*. All'interno dell'installazione del server delle applicazioni WebSphere, accedere alla directory *\$JAVA_HOME/jre/lib/security* ed eseguire il backup dei file *US_export_policy.jar* e *local_policy.jar*.
- ▶ Sostituire i propri file *US_export_policy.jar* e *local_policy.jar* con i due file scaricati.
- ▶ Attivare la protezione in Administration Console del server delle applicazioni WebSphere. Inizialmente, accertarsi che tutti gli agenti del nodo della cella siano attivi. Per ulteriori informazioni, vedere la documentazione di WebSphere. Affinché la protezione rimanga abilitata dopo il riavvio del server, è necessario selezionare un'area di autenticazione disponibile dall'elenco in Security > Secure administration, applications, and infrastructure e fare clic su Set as current.
- ▶ Scollegarsi da Administration Console.
- ▶ Arrestare il server.
- ▶ Riavviare il server.

Sun Java

- ▶ Scaricare i file Java Cryptography Extension (JCE) Unlimited Strength Jurisdiction Policy relativi alla propria versione SDK dal sito Web Sun Java.
- ▶ Decomprimere il file scaricato.
- ▶ Copiare i due file *.jarlocal_policy.jar* e *US_export_policy.jar* in *<cartella di installazione>/jre/lib/security*, dove *<cartella di installazione>* è la cartella in cui è stato installato il prodotto.

Aggiunta del certificato all'archivio chiavi client (per le connessioni al repository o a IBM SPSS Modeler Server)

Nota: saltare questo passaggio se si sta utilizzando un certificato firmato da un'autorità di certificazione.

Se per connettersi al repository IBM® SPSS® Collaboration and Deployment Services o a IBM® SPSS® Modeler Server si utilizza SSL e si dispone di certificati autofirmati, è necessario aggiungere il certificato all'archivio chiavi Java del client. I passaggi seguenti devono essere eseguiti sul computer *client*.

Si ricordi che un computer server potrebbe fungere anche da client. Un esempio è IBM® SPSS® Statistics Server che si connette al IBM® SPSS® Collaboration and Deployment Services Repository. In questo caso, SPSS Statistics Server è il *client* e pertanto è necessario copiare il certificato del IBM SPSS Collaboration and Deployment Services Repository server in SPSS Statistics Server.

- ▶ Copiare il file del certificato (.cer) dal server SSL alla directory *<dir_installazione>/jre/lib/security* sul computer client.

- ▶ Nella stessa directory, fare clic con il pulsante destro del mouse sul file *cacerts* e selezionare Proprietà.
- ▶ Deselezionare la casella di controllo Sola lettura.
- ▶ Aprire un prompt dei comandi e passare alla directory nel percorso seguente, dove *<dir installazione prodotto>* è la directory in cui il prodotto è stato installato:

```
<dir installazione prodotto>/jre/bin
```

- ▶ Immettere il comando seguente:

```
keytool -import -alias <nome alias> -file <percorso cert> -keystore <percorso archivio chiavi>
```

Dove *<nome alias>* è un nome alias arbitrario per il certificato, *<percorso cert>* è il percorso completo del certificato e *<percorso archivio chiavi>* è il percorso completo dell'archivio chiavi Java, che potrebbe essere *<dir installazione prodotto>/lib/security/jssecacerts* o *<dir installazione prodotto>/lib/security/cacerts*.

- ▶ Quando richiesto, immettere la password dell'archivio chiavi; la password di default è *changeit*.
- ▶ Alla domanda se si considera affidabile il certificato, rispondere *yes*.

Istruzione agli utenti finali per l'attivazione di SSL

Quando gli utenti finali si connettono al server tramite un prodotto client devono attivare SSL nella finestra di dialogo. Assicurarsi di comunicare agli utenti la casella di controllo appropriata da selezionare. Per esempio, in IBM® SPSS® Modeler, gli utenti devono selezionare la casella di controllo *Abilita connessione sicura* nella finestra di dialogo *Repository: Server*, a cui si accede mediante l'opzione *Repository* del menu *Strumenti*.

Configurazione del prefisso dell'URL

Se IBM® SPSS® Collaboration and Deployment Services Repository è impostato per consentire l'accesso SSL, il valore dell'impostazione di configurazione *Prefisso URL* deve essere modificato nel modo seguente:

1. Accedere al repository utilizzando la console basata sul browser.
2. Accedere all'opzione di configurazione *Prefisso URL*.
Configurazione > Installazione > Prefisso URL
3. Impostare il valore del prefisso su *https* anziché su *http* e impostare il valore della porta sul numero di porta SSL. Per esempio:

```
[default]
http://<hostname>:<port>
[SSL-enabled]
https://<hostname>:<SSLport>
```

Protezione di LDAP con SSL

LDAP (Lightweight Directory Access Protocol) è uno standard IETF (Internet Engineering Task Force) per lo scambio di informazioni tra le directory e i database di rete contenenti qualsiasi livello di informazioni. Per i sistemi che richiedono ulteriore sicurezza, i provider LDAP, quali Active Directory di Microsoft, possono operare su un SSL (Secure Socket Layer), purché il server Web o applicazioni supporti LDAP su SSL. Utilizzando SSL insieme a LDAP si può assicurare che le password di accesso, le informazioni sull'applicazione e altri dati sensibili non siano oggetto di furto o compromissioni.

Il seguente esempio mostra come attivare LDAPS tramite Active Directory di Microsoft come provider di protezione. Per informazioni più specifiche sui passaggi o per trovare suggerimenti su una particolare release del provider di protezione, consultare la documentazione del fornitore originale.

1. Verificare che Active Directory e Enterprise Certificate Authority siano installate e funzionanti.
2. Utilizzare l'autorità di certificazione per generare un certificato e importarlo nell'archivio certificati dell'installazione di IBM® SPSS® Collaboration and Deployment Services Deployment Manager. In questo modo è possibile stabilire la connessione LDAPS tra il IBM® SPSS® Collaboration and Deployment Services Repository e il server Active Directory.

Per configurare Deployment Manager per connessioni Active Directory protette, verificare che esista una connessione al repository.

3. Avviare IBM® SPSS® Collaboration and Deployment Services Deployment Manager.
4. Dal menu Strumenti scegliere Server Administration.
5. Accedere a un server amministrato definito precedentemente.
6. Fare doppio clic sull'icona Configurazione per espandere la gerarchia del server.
7. Fare doppio clic sull'icona Provider di protezione per espandere la gerarchia.
8. Fare doppio clic sul provider di protezione Active Directory.
9. Inserire i valori di configurazione per l'istanza di Active Directory con i certificati di sicurezza installati.
10. Selezionare la casella di controllo Utilizza SSL.
11. Annotare il nome nel campo Utente dominio. Gli accessi successivi tramite Active Directory sono autenticati con SSL.

Per ulteriori informazioni sull'installazione, configurazione e implementazione di LDAPS su un server applicazioni particolare, vedere la documentazione del fornitore originale.

Registro del server

IBM® SPSS® Modeler Server conserva le operazioni importanti in un file registro denominato *messages.log*. Per default, questo file si trova nella directory *[percorso di installazione server]/log*, ma la sua posizione è configurabile. Quando viene avviato SPSS Modeler Server,

il file registro viene aggiornato in modo da indicare l'ora di avvio e il numero di porta su cui l'applicazione è in attesa. Da quel punto in poi, il file registro viene aggiornato con il risultato ogni volta che un client effettua un tentativo di connessione al server.

Per default, il server inserisce nel file registro gli eventi seguenti:

- Avvio del server
- Arresto del server
- Connessione accettata
- Connessione respinta
- Sessione avviata
- Sessione terminata
- Errore di sistema
- Login riuscito
- Login non riuscito

Se si desidera registrare le informazioni relative agli eventi di accesso al database, impostare le opzioni Registra informazioni di accesso al database e Registra errori del database su *sì* nella configurazione di SPSS Modeler Server. Gli eventi di accesso al database che vengono registrati sono i seguenti:

- Accesso al database riuscito
- Accesso al database non riuscito
- Disconnessione dal database

Se si desidera registrare le informazioni relative agli eventi di accesso ai file, impostare le opzioni Registra informazioni di accesso ai file e Registra errori di accesso ai file su *sì* nella configurazione di SPSS Modeler Server. Gli eventi di accesso ai file che vengono registrati sono i seguenti:

- Apertura/creazione di file
- Chiusura di file

Di seguito è riportato un esempio di file registro (sono stati aggiunti i numeri delle righe):

```
1. 2006/06/14 08:41:23 [167]: IBM SPSS Modeler Server Version 0.1.0 started
2. 2006/06/14 08:41:23 [167]: Listening on port 28028
3. 2006/06/14 08:49:28 [167]: Accepted connection from localhost:1827
4. 2006/06/14 08:49:28 [167]: Session 105 (localhost:1827) started
5. 2006/06/14 08:49:29 [167-105]: Login failed for user: SPSS\awilson
6. 2006/06/14 09:27:27 [167]: Session 105 ended
7. 2006/06/14 09:53:39 [167]: Accepted connection from localhost:1837
8. 2006/06/14 09:53:39 [167]: Session 63 (localhost:1837) started
9. 2006/06/14 09:53:40 [167-63]: Login succeeded for user: SPSS\awilson
10. 2006/06/14 10:27:31 [167]: Session 63 ended
```

Tutte le voci presenti nel file registro contengono un timestamp. Le righe 1 e 2 vengono inserite nel file registro all'avvio di SPSS Modeler Server. La riga 3 viene aggiunta quando un utente del computer *localhost* sulla porta numero *1827* tenta di connettersi al server. La riga 4 indica che con questo utente è stata avviata una sessione con il server. La riga 5 mostra che l'utente non è riuscito

ad accedere correttamente. Poco dopo, l'utente ha terminato la sessione. Mezzora dopo viene effettuato un altro tentativo di connessione da *localhost* e viene avviata una sessione. Questa volta l'accesso riesce. Dopo altri trenta minuti, la sessione termina e la connessione viene chiusa.

Panoramica delle performance

Le performance effettive durante l'analisi dei dati sono influenzate da numerosi fattori, dalla configurazione del server e del database all'ordine dei singoli nodi all'interno di uno stream. In generale, le performance sono migliori quando si adottano i seguenti accorgimenti:

- Archiviare i dati in un DBMS e utilizzare la generazione e l'ottimizzazione SQL quando possibile.
- Utilizzare componenti hardware di livello pari o superiore a quello consigliato in [il capitolo 2](#).
- Verificare che le impostazioni relative alle performance e all'ottimizzazione di client e server siano configurate correttamente. Si noti che quando SPSS Modeler è collegato a un'installazione SPSS Modeler Server, le impostazioni di performance e ottimizzazione del server sovrascrivono le impostazioni equivalenti del client.
- Progettare gli stream per ottenere il massimo delle prestazioni. [Per ulteriori informazioni, vedere l'argomento Considerazioni sulle performance degli stream e dei nodi in il capitolo 13 in Manuale dell'utente di IBM SPSS Modeler 15.](#)

Ulteriori informazioni sui singoli fattori che incidono sulle prestazioni sono riportate nelle sezioni che seguono.

Impostazioni relative alle performance e all'ottimizzazione del server

Per ottimizzare le prestazioni è possibile configurare alcune impostazioni di IBM® SPSS® Modeler Server regolando con l'applicazione IBM® SPSS® Modeler Administration Console, fornita con SPSS Modeler Server. [Per ulteriori informazioni, vedere l'argomento Utilizzo di IBM SPSS Modeler Administration Console in il capitolo 4 a pag. 26.](#)

Le impostazioni sono riunite sotto l'intestazione Performance e ottimizzazione nella finestra di configurazione di Modeler Administration Console e sono già configurate per ottenere prestazioni ottimali nella maggior parte delle installazioni. Tuttavia, può essere necessario modificarle in funzione dell'hardware utilizzato, delle dimensioni degli insiemi di dati e del contenuto degli stream. [Per ulteriori informazioni, vedere l'argomento Performance/ottimizzazione in il capitolo 4 a pag. 31.](#)

Impostazioni relative alle performance e all'ottimizzazione del client

Le impostazioni relative alle performance e all'ottimizzazione del client sono disponibili nella scheda Opzioni della finestra di dialogo Proprietà stream. Per visualizzarle, scegliere le seguenti opzioni dal menu del client.

Strumenti > Proprietà stream > Opzioni > Ottimizzazione

Per ottimizzare le prestazioni degli stream sono disponibili opzioni di ottimizzazione. Si noti che le impostazioni relative alle prestazioni e all'ottimizzazione in IBM® SPSS® Modeler Server (se utilizzato) prevalgono sulle impostazioni equivalenti nel client.

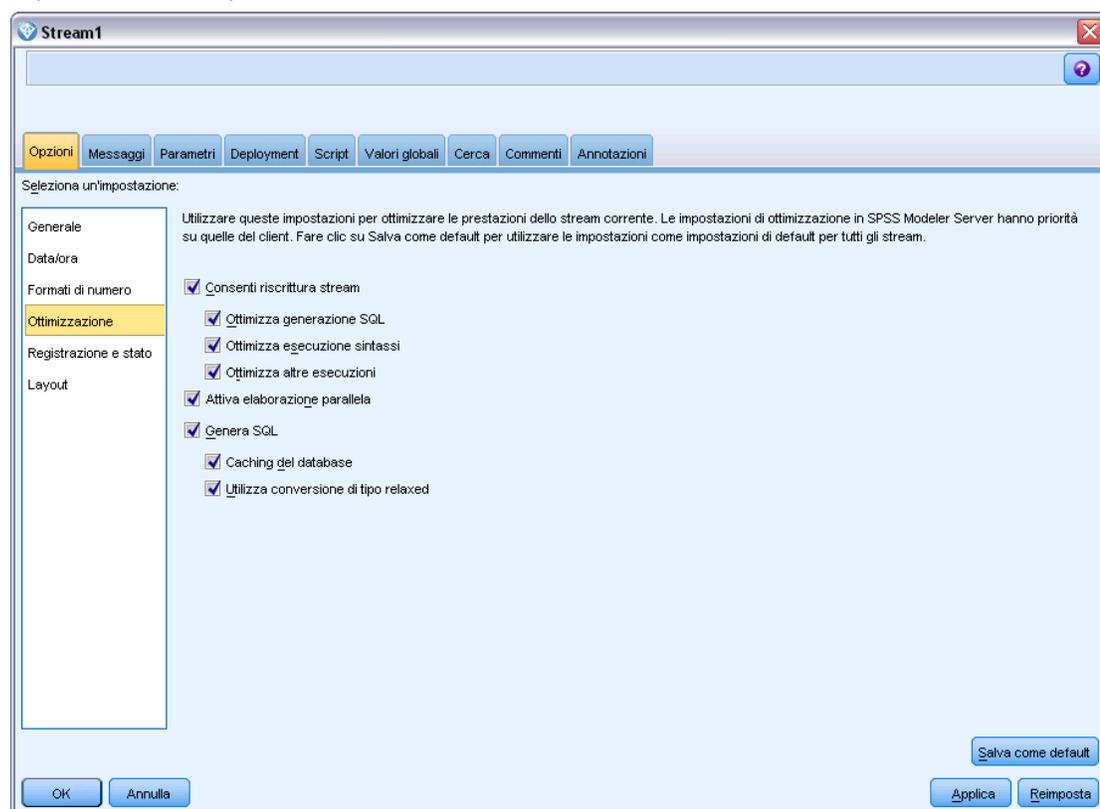
Nota: le funzionalità di modellazione in-database e ottimizzazione SQL richiedono che sul computer IBM® SPSS® Modeler sia attivata la connettività SPSS Modeler Server. Con questa impostazione attivata, è possibile accedere agli algoritmi di database, restituire codice SQL direttamente da SPSS Modeler e accedere a SPSS Modeler Server. Per verificare lo stato attuale della licenza, scegliere le seguenti opzioni dal menu SPSS Modeler.

Guida > Informazioni su > Ulteriori dettagli

Se la connettività è abilitata, l'opzione Abilitazione server viene visualizzata nella scheda Stato della licenza.

Per ulteriori informazioni, vedere l'argomento [Connessione a IBM SPSS Modeler Server](#) in il capitolo 3 in *Manuale dell'utente di IBM SPSS Modeler 15*.

Figura 5-1
Impostazione delle opzioni di ottimizzazione dello stream



Nota: il supporto del push back e dell'ottimizzazione SQL dipende dal tipo di database in uso. Per le ultime informazioni sui database e i driver ODBC che sono supportati e testati per l'uso con IBM® SPSS® Modeler 15, vedere il sito aziendale dedicato al supporto tecnico all'indirizzo <http://www.ibm.com/support>.

Consenti riscrittura stream. Selezionare questa opzione per consentire la riscrittura degli stream in SPSS Modeler. Sono disponibili due tipi di riscrittura ed è possibile selezionarne uno o entrambi. La riscrittura degli stream riordina in background i nodi in uno stream per migliorarne il funzionamento, senza alterare la semantica degli stream.

- **Ottimizza generazione SQL.** Questa opzione consente di riordinare i nodi nello stream in modo da permettere di eseguire il push back di un maggior numero di operazioni tramite la generazione SQL per l'esecuzione nel database. Quando trova un nodo che non può essere reso in codice SQL, l'ottimizzatore controlla più avanti l'eventuale presenza di nodi a valle che è possibile rendere in SQL e spostare in modo sicuro davanti al nodo del problema senza compromettere la semantica dello stream. Non solo il database può eseguire le operazioni in modo più efficiente di SPSS Modeler, ma questi pushback o "rinvii" hanno anche la possibilità di ridurre la dimensione dell'insieme di dati restituito a SPSS Modeler per l'elaborazione. Di conseguenza, si riduce il traffico di rete e si sveltiscono le operazioni stream. Si noti che è necessario selezionare la casella di controllo Genera SQL per consentire l'ottimizzazione SQL. [Per ulteriori informazioni, vedere l'argomento Ottimizzazione SQL in il capitolo 6 a pag. 52.](#)
- **Ottimizza esecuzione sintassi.** Questo metodo di riscrittura degli stream aumenta l'efficienza delle operazioni che incorporano più di un nodo contenente la sintassi IBM® SPSS® Statistics. L'ottimizzazione si ottiene combinando i comandi di sintassi all'interno di un'unica operazione, invece di eseguirli singolarmente come operazioni separate.
- **Ottimizza altre esecuzioni.** Questo metodo di riscrittura degli stream aumenta l'efficienza delle operazioni (quelle che non possono essere delegate al database). L'ottimizzazione viene raggiunta riducendo la quantità dei dati nello stream non appena possibile. Pur mantenendo l'integrità dei dati, lo stream viene riscritto per spingere le operazioni più vicino alla sorgente dati, riducendo in tal modo i dati a valle per operazioni dispendiose quali le unioni.

Attiva elaborazione parallela. Quando l'esecuzione avviene su un computer con più processori, questa opzione consente al sistema di distribuire il carico di lavoro su più processori, con un conseguente miglioramento delle prestazioni. L'elaborazione parallela è utile in combinazione con i seguenti nodi, utilizzati in combinazione o singolarmente: C5.0, Unione (per chiave), Ordina, Discretizza (metodi di discretizzazione per rango e per blocco) e Aggregazione (utilizzando uno o più campi chiave).

Genera SQL. Selezionare questa opzione per attivare la generazione SQL, consentendo di rinviare operazioni dello stream al database tramite codice SQL per generare processi di esecuzione in grado di migliorare le prestazioni. Per migliorare ulteriormente le prestazioni, è inoltre possibile selezionare l'opzione Ottimizza generazione SQL per aumentare al massimo il numero di operazioni rinviate al database. Dopo che le operazioni per un nodo sono state rinviate al database, il nodo viene evidenziato dal colore porpora durante l'esecuzione dello stream.

- **Caching del database.** Per gli stream che generano SQL da eseguire nel database, i dati possono essere inseriti nella cache midstream in una tabella temporanea nel database anziché nel file system. Insieme all'ottimizzazione SQL, ciò consente di ottenere significativi vantaggi in termini di prestazioni. Per esempio, l'output di uno stream che unisce più tabelle per creare una visualizzazione data mining può essere inserito nella cache e riutilizzato in un secondo tempo in base alle esigenze. Con l'inserimento nella cache del database attivato, è sufficiente fare clic con il pulsante destro del mouse su qualsiasi nodo non terminale per inserire i dati nella cache in quel punto e la cache viene creata automaticamente direttamente nel database alla successiva esecuzione dello stream. Ciò consente la generazione di SQL

per i nodi a valle, migliorando ulteriormente le prestazioni. In alternativa, se lo si desidera, è possibile disattivare questa opzione, per esempio se i criteri o le autorizzazioni precludono la scrittura dei dati nel database. Se l'inserimento nella cache del database o l'ottimizzazione SQL non sono attivati, la cache verrà scritta nel file system. [Per ulteriori informazioni, vedere l'argomento Opzioni di copia dei dati nella cache per i nodi in il capitolo 5 in Manuale dell'utente di IBM SPSS Modeler 15.](#)

- **Utilizza conversione di tipo relaxed.** Questa opzione consente la conversione dei dati da stringhe a numeri o da numeri a stringhe, purché siano archiviati in un formato adatto. Ad esempio, se i dati sono conservati nel database sotto forma di stringa, ma contengono un numero significativo, possono essere convertiti per essere utilizzati quando si esegue il push back.

Nota: a causa di piccole differenze di implementazione del codice SQL, gli stream eseguiti in un database potrebbero restituire risultati leggermente diversi da quelli che restituiscono quando vengono eseguiti in SPSS Modeler. Per ragioni analoghe, tali differenze possono variare anche in base al fornitore del database.

Ottimizzazione e utilizzo del database

Server del database. Se possibile, creare un'istanza del database dedicata per il data mining al fine di evitare che le query eseguite da IBM® SPSS® Modeler influiscano sulle prestazioni del server di produzione. Le istruzioni SQL generate da SPSS Modeler possono assorbire molte risorse: è possibile infatti che più operazioni in corso sul computer di IBM® SPSS® Modeler Server eseguano contemporaneamente istruzioni SQL nello stesso database.

Mining in-database. Molti produttori di database forniscono estensioni per il data mining per i loro prodotti. Tali estensioni consentono di eseguire attività di data mining (quali la creazione di modelli o il calcolo di punteggi) sul server del database o su un server dedicato separato. Le funzionalità di mining in-database di SPSS Modeler ne completano e ne ampliano le funzioni di generazione SQL, offrendo un sistema per controllare le estensioni database proprietarie fornite dai produttori. In alcuni casi, questo approccio consente di evitare overhead di trasferimento dei dati potenzialmente dispendiosi fra SPSS Modeler e il database. Il caching del database può garantire ulteriori vantaggi. Per maggiori informazioni, vedere il file *DatabaseMiningGuide.pdf*, disponibile nella cartella *\Documentation* del disco di installazione di IBM® SPSS® Modeler.

Ottimizzazione SQL

Per ottenere prestazioni ottimali, è necessario cercare sempre di ottimizzare la quantità di istruzioni SQL generate per sfruttare le prestazioni e la scalabilità del database. Solo le parti dello stream che non possono essere compilate in SQL devono essere eseguite in IBM® SPSS® Modeler Server. Per ulteriori informazioni, vedere [il capitolo 6](#).

Caricamento di dati contenuti in file

I dati che non sono archiviati in un database non possono usufruire dell'ottimizzazione SQL. Se i dati da analizzare non si trovano già in un database, è possibile caricarli mediante il nodo di output Database. Questo nodo può essere utilizzato anche per memorizzare insieme di dati intermedi derivanti dalla preparazione dei dati e i risultati del deployment. [Per ulteriori informazioni, vedere](#)

[l'argomento Esportazione database - Opzioni avanzate in il capitolo 7 in IBM SPSS Modeler 15 Nodi di input, elaborazione e output.](#)

IBM® SPSS® Modeler può interfacciarsi con i caricatori esterni di molti dei più diffusi sistemi di database. Il software è corredato di numerosi script che sono disponibili insieme alla documentazione nella sottodirectory */scripts* della cartella di installazione di SPSS Modeler.

La tabella che segue mostra i potenziali vantaggi del caricamento di massa in termini di performance. I numeri mostrano il tempo impiegato per esportare 250.000 record e 21 campi in un database di Oracle. Il caricatore esterno utilizzato è l'utilità *sqlldr* di Oracle.

Tabella 5-1

Vantaggi del caricamento di massa in termini di performance

Opzione di esportazione	Tempo (in secondi)
Default (ODBC)	409
Caricamento di massa tramite ODBC	52
Caricamento di massa tramite caricatore esterno	33

Ottimizzazione SQL

Una delle funzionalità più potenti di IBM® SPSS® Modeler è la capacità di eseguire numerose operazioni di preparazione dei dati e di mining direttamente nel database. Tramite la generazione di codice SQL che è possibile rinviare al database per l'esecuzione, molte operazioni, quali campionamento, ordinamento, derivazione di nuovi campi e alcuni tipi di rappresentazione grafica, possono essere eseguite nel database anziché nel computer IBM® SPSS® Modeler o IBM® SPSS® Modeler Server. Quando si lavora su insiemi di dati di grandi dimensioni, i **pushback**, o rinvii, possono migliorare significativamente la performance in diversi modi:

- Riducendo le dimensioni dei risultati impostati per essere trasformati da DBMS a SPSS Modeler. Quando gli insiemi di risultati di grandi dimensioni vengono letti tramite un driver ODBC, potrebbero verificarsi inefficienze a livello di driver o I/O di rete. Per questo motivo, le operazioni che beneficiano maggiormente dell'ottimizzazione SQL sono la selezione e l'aggregazione di righe e colonne (nodi Seleziona, Campione e Aggregazione), che in genere riducono le dimensioni dell'insieme di dati da trasferire. I dati possono anche essere inseriti nella cache in una tabella temporanea nel database, in punti critici dello stream (per esempio dopo un nodo Unione o Seleziona) per migliorare ulteriormente le prestazioni.
- Sfruttando le prestazioni e la scalabilità del database. L'efficienza viene incrementata in quanto un DBMS può sfruttare l'elaborazione parallela, hardware più potente, una gestione più sofisticata dell'archiviazione su disco e la presenza di indici.

Considerati questi vantaggi, SPSS Modeler è progettato per aumentare al massimo la quantità di SQL generati da ogni stream in modo che solo le operazioni che non possono essere compilate in SQL vengano eseguite da SPSS Modeler Server. A causa dei limiti di espressione dell'SQL standard (SQL-92), tuttavia, è possibile che alcune operazioni non siano supportate. [Per ulteriori informazioni, vedere l'argomento Suggerimenti per l'ottimizzazione della generazione SQL a pag. 58.](#)

Nota: A causa di piccole differenze di implementazione del codice SQL, gli stream eseguiti in un database potrebbero restituire risultati leggermente diversi quando vengono eseguiti in SPSS Modeler. Per ragioni analoghe, tali differenze possono variare anche in base al fornitore del database. Per esempio, in base alla configurazione del database per la distinzione tra maiuscole e minuscole nel confronto e nella collazione di stringhe, gli stream di SPSS Modeler eseguiti utilizzando il push back SQL potrebbero produrre risultati diversi da quelli eseguiti senza push back SQL. Per consigli sulla configurazione del database, rivolgersi all'amministratore di database. Per ottimizzare la compatibilità con SPSS Modeler, i confronti tra stringhe del database devono fare distinzione tra maiuscole e minuscole.

Nota: le funzionalità di modellazione in-database e ottimizzazione SQL richiedono che sul computer SPSS Modeler sia attivata la connettività SPSS Modeler Server. Con questa impostazione attivata, è possibile accedere agli algoritmi di database, restituire codice SQL direttamente da SPSS Modeler e accedere a SPSS Modeler Server. Per verificare lo stato attuale della licenza, scegliere le seguenti opzioni dal menu SPSS Modeler.

Guida > Informazioni su > Ulteriori dettagli

Se la connettività è abilitata, l'opzione *Abilitazione server* viene visualizzata nella scheda *Stato della licenza*.

Per ulteriori informazioni, vedere l'argomento *Connessione a IBM SPSS Modeler Server* in il capitolo 3 in *Manuale dell'utente di IBM SPSS Modeler 15*.

Nota: Quando gli stream vengono eseguiti in un database, i dettagli di data e ora sono prelevati dal database. Questi valori potrebbero non corrispondere alla data e ora locali o di SPSS Modeler Server se, ad esempio, il database risiede su un computer dislocato in un paese o fuso orario diverso.

Requisiti del database

Per le ultime informazioni sui database e i driver ODBC che sono supportati e testati per l'uso con SPSS Modeler, vedere le matrici di compatibilità del prodotto sul sito aziendale dedicato al supporto tecnico all'indirizzo <http://www.ibm.com/support>.

Si noti che è possibile ottenere ulteriori miglioramenti in termini di performance utilizzando la modellazione di database. Per ulteriori informazioni, vedere l'argomento *Panoramica sulla modellazione di database* in il capitolo 2 in *IBM SPSS Modeler 15 Guida alla modellazione in-database*.

Funzionamento della generazione SQL

I frammenti iniziali di uno stream dai nodi di input del database sono gli obiettivi principali della generazione SQL. Quando si incontra un nodo che non può essere compilato in SQL, i dati vengono estratti dal database e la successiva elaborazione viene eseguita da IBM® SPSS® Modeler Server.

Durante la preparazione degli stream e prima dell'esecuzione, il processo di generazione SQL si verifica come descritto di seguito:

- Il server riordina gli stream per spostare i nodi a valle nella “zona SQL” dove l'operazione può essere eseguita in sicurezza. Questa funzione può essere disattivata nel server.
- Procedendo dai nodi di input verso i nodi terminali, le espressioni SQL vengono costruite in modo incrementale. Questa fase si interrompe quando viene riscontrato un nodo che non è possibile convertire in SQL oppure il nodo terminale (per esempio un nodo *Tabella* o *Grafico*) viene convertito in SQL. Alla fine della fase, ogni nodo viene etichettato con un'istruzione SQL se il nodo e i relativi predecessori hanno un equivalente SQL.
- Procedendo dai nodi con gli equivalenti SQL più complicati verso i nodi di input, viene verificata la validità dell'SQL. L'SQL che è stato convalidato correttamente viene scelto per l'esecuzione.
- I nodi per i quali tutte le operazioni hanno generato SQL vengono evidenziati in colore porpora nell'area di disegno dello stream. In base ai risultati, è possibile riorganizzare ulteriormente lo stream, dove appropriato, per beneficiare al massimo dell'esecuzione del database. Per ulteriori informazioni, vedere l'argomento *Suggerimenti per l'ottimizzazione della generazione SQL* a pag. 58.

Punti di miglioramento

L'ottimizzazione SQL migliora le performance in numerose operazioni sui dati:

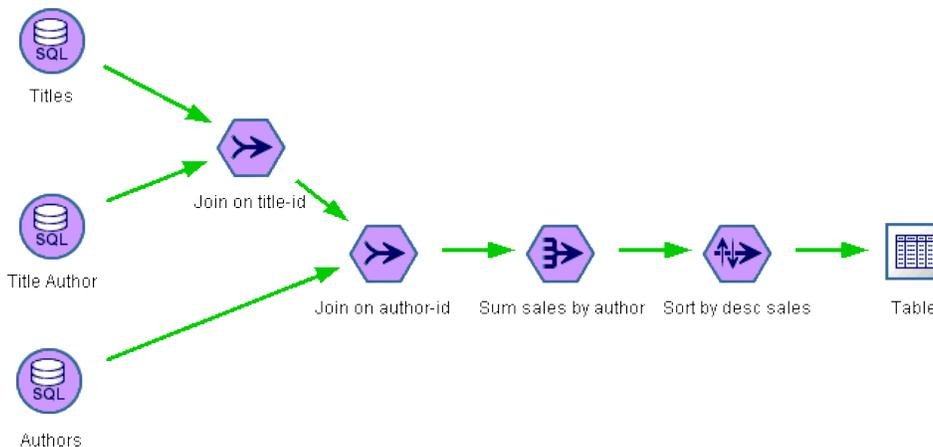
- **Join (unione per chiave).** Le operazioni di join possono migliorare l'ottimizzazione nei database.
- **Aggregazione.** Per produrre i loro risultati, i nodi Aggregazione, Distribuzione e Web utilizzano l'aggregazione. I dati riepilogati utilizzano la larghezza di banda in modo molto più ridotto rispetto ai dati originali.
- **Selezione.** La scelta dei record in base a determinati criteri riduce la quantità di record.
- **Ordinamento.** L'ordinamento dei record è un'attività che utilizza molte risorse e viene eseguita in modo più efficiente in un database.
- **Derivazione dei campi.** In un database, i nuovi campi vengono generati in modo più efficiente.
- **Proiezione dei campi.** SPSS Modeler Server estrae dal database solo i campi necessari per la successiva elaborazione, riducendo al minimo i requisiti in termini di larghezza di banda e memoria. Lo stesso vale per i campi superflui nei file piatti: sebbene il server debba leggere i campi superflui, non assegna loro alcuna archiviazione.
- **Calcolo dei punteggi.** La generazione di un SQL può derivare da alberi decisionali, insiemi di regole, regressione lineare e modelli generati da fattori.

Esempio di generazione SQL

Lo stream seguente unisce tre tabelle di database per operazioni chiave ed esegue un'aggregazione e un ordinamento.

Figura 6-1

Stream ottimizzato con nodi porpora che indicano push back SQL (operazioni eseguite nel database)



SQL generato

Di seguito è riportato l'SQL generato per questo stream:

```
SELECT
  T2. au_lname AS C0,
```

```

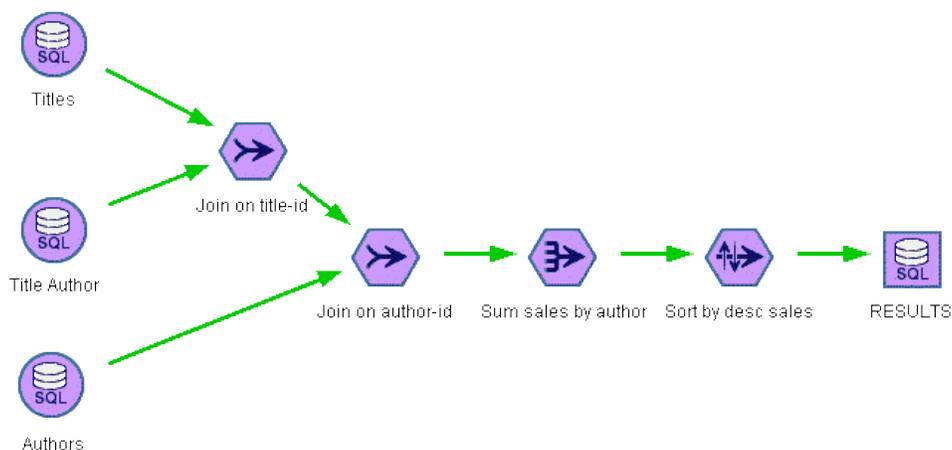
T2.au_fname AS C1,
SUM({fn CONVERT(T0.ytd_sales ,SQL_BIGINT)}) AS C2
FROM
dbo.titles T0,
dbo.titleauthor T1,
dbo.authors T2
WHERE
(T0.title_id = T1.title_id )
AND (T1.au_id = T2.au_id )
GROUP BY T2.au_lname ,T2.au_fname
ORDER BY 3 DESC

```

Esecuzione degli stream

Quando lo stream termina con un nodo Database, è possibile che l'intero stream venga eseguito nel database.

Figura 6-2
Stream intero eseguito nel database



Configurazione dell'ottimizzazione SQL

- ▶ Installare un driver ODBC e configurare una sorgente dati per il database che si desidera utilizzare. [Per ulteriori informazioni, vedere l'argomento Accesso ai dati in il capitolo 2 a pag. 13.](#)
- ▶ Creare uno stream che utilizza un nodo origine per estrarre i dati dal database.
- ▶ Assicurarsi che la generazione SQL sia attivata sul client e sul server, se applicabile. Per default, è attivata su entrambi.

Attivazione dell'ottimizzazione SQL sul client

- ▶ Scegliere Opzioni utente dal menu Strumenti.
- ▶ Fare clic sulla scheda Ottimizzazione. Selezionare Genera SQL per attivare l'ottimizzazione SQL. Se lo si desidera, è possibile selezionare altre impostazioni per migliorare le prestazioni.

Per ulteriori informazioni, vedere l'argomento [Impostazioni relative alle performance e all'ottimizzazione del client in il capitolo 5 a pag. 47](#).

Attivazione dell'ottimizzazione SQL sul server

Poiché le impostazioni del server sovrascrivono qualsiasi specifica eseguita nel client, le impostazioni di configurazione del server Riscrittura stream e Generazione automatica SQL devono essere entrambe attivate. Per ulteriori informazioni sulla modifica delle impostazioni di IBM® SPSS® Modeler Server, vedere la sezione [Performance/ottimizzazione nel il capitolo 4](#).

Attivazione dell'ottimizzazione durante il calcolo del punteggio dei modelli

Ai fini del calcolo del punteggio, si tenga presente che è necessario attivare la generazione SQL separatamente per ogni nodo di modellazione, indipendentemente dalle impostazioni a livello di server o client. Questo perché alcuni modelli generano espressioni SQL estremamente complesse che potrebbero non essere valutate correttamente all'interno del database. Il database può segnalare errori quando si tenta di eseguire l'espressione SQL generata, a causa delle sue dimensioni o della sua complessità.

Per determinare se la generazione SQL migliora le prestazioni di un determinato modello può essere necessario procedere in modo empirico. Questa operazione viene eseguita sulla scheda Impostazioni dopo aver aggiunto un modello generato a uno stream. [Per ulteriori informazioni, vedere l'argomento Utilizzo degli insiemi di modelli negli stream in il capitolo 3 in IBM SPSS Modeler 15 Nodi Modelli](#).

Anteprima di SQL generato

È possibile visualizzare l'anteprima di SQL generato nel registro del messaggio prima di eseguirlo nel database. Potrebbe essere utile agli scopi di debug e consente di esportare l'SQL generato per modificarlo o eseguirlo nel database in futuro. Inoltre, indica quali nodi verranno rinviati al database, informazione che può essere utile per stabilire se è possibile riordinare lo stream per migliorare le performance.

- ▶ Assicurarsi che le opzioni Visualizza SQL nel registro messaggi durante l'esecuzione dello stream e Visualizza dettagli generazione SQL nel registro messaggi durante la preparazione dello stream siano selezionate nella finestra di dialogo Opzioni utente. [Per ulteriori informazioni, vedere l'argomento Impostazioni relative alle performance e all'ottimizzazione del client in il capitolo 5 a pag. 47](#).
- ▶ Nell'area di disegno dello stream, selezionare il nodo o lo stream che si desidera visualizzare in anteprima.

- Fare clic sul pulsante Anteprima SQL nella barra degli strumenti.

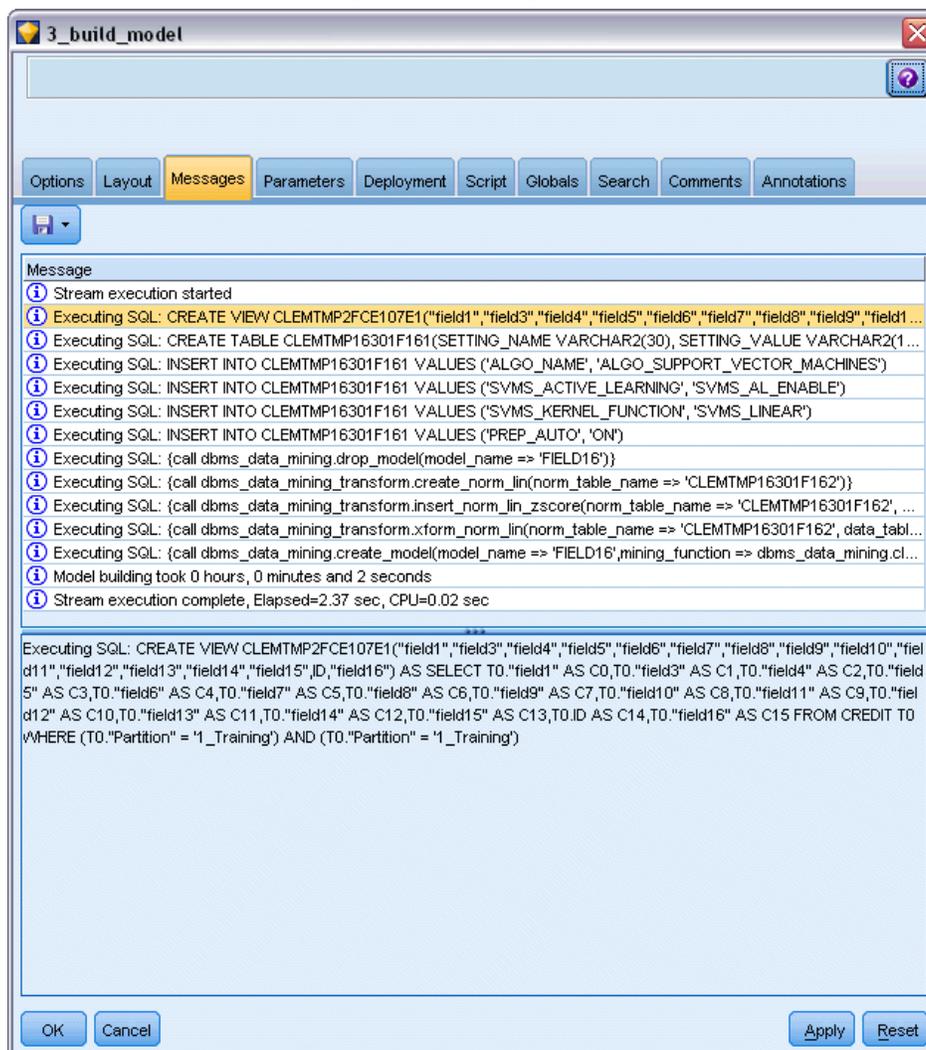
Tutti i nodi per i quali viene generato SQL (e che verranno rinviati al database quando lo stream sarà eseguito) sono colorati in porpora nell'area di disegno dello stream.

Figura 6-3
Pulsante Anteprima SQL



- Per vedere l'anteprima dell'SQL generato, dai menu scegliere:
Strumenti > Proprietà stream > Messaggi...

Figura 6-4
SQL generato visualizzato nel registro dei messaggi



Visualizzazione di SQL per gli insiemi di modelli

Per alcuni modelli, è possibile generare SQL per l'insieme di modelli, rinviando la fase di calcolo del punteggio del modello al database. L'utilizzo principale di questa funzione non è migliorare le prestazioni, ma consentire il rinvio del loro SQL completo agli stream contenenti questi insiemi di modelli. [Per ulteriori informazioni, vedere l'argomento Nodo con supporto per generazione SQL a pag. 59.](#)

Per visualizzare SQL per un insieme di modelli che supporta la generazione SQL:

- ▶ Selezionare la scheda Impostazioni nell'insieme di modelli.
- ▶ Scegliere una delle opzioni Genera con (senza) supporto per valori mancanti o Genera SQL per questo modello in base alle esigenze.
- ▶ Nel menu dell'insieme di modelli, scegliere:
File > Esporta in SQL
- ▶ Salvare il file.
- ▶ Aprire il file per visualizzare l'SQL.

Suggerimenti per l'ottimizzazione della generazione SQL

Per ottenere un incremento significativo delle prestazioni dall'ottimizzazione SQL, tenere presente quanto segue.

Ordine stream. La generazione SQL può essere interrotta se la funzione del nodo non ha equivalente semantico in SQL, perché la funzionalità di data mining di IBM® SPSS® Modeler è più ricca delle tradizionali operazioni di elaborazione dati sopportate dal codice SQL standard. Quando si verifica tale situazione, la generazione SQL viene eliminata anche per tutti i nodi a valle. Pertanto, è possibile migliorare in modo significativo le prestazioni riordinando i nodi in modo da mettere le operazioni che interrompono SQL quanto più possibile a valle. L'ottimizzatore SQL può eseguire automaticamente una certa quantità di operazioni di riordino (è sufficiente assicurarsi che la funzione di riscrittura stream sia attivata), ma sono possibili ulteriori miglioramenti. Un buon candidato a questo scopo è il nodo Seleziona, che spesso può essere spostato in avanti. [Per ulteriori informazioni, vedere l'argomento Nodo con supporto per generazione SQL a pag. 59.](#)

CLEM espressioni. Se uno stream non può essere riordinato, è possibile modificare le opzioni dei nodi e le espressioni CLEM oppure ricomporre diversamente il modo in cui viene eseguita l'operazione, in modo che la generazione SQL non venga più inibita. I nodi Nuovo campo, Seleziona e simili possono essere resi in codice SQL, purché tutti gli operatori delle espressioni CLEM abbiano equivalenti SQL. La maggior parte degli operatori può essere resa in codice SQL, ma vi sono numerosi operatori che inibiscono la generazione SQL (in particolare, le funzioni di sequenza ["funzioni @"]). Talvolta, la generazione viene interrotta perché la query generata è diventata troppo complessa per essere gestita dal database. [Per ulteriori informazioni, vedere l'argomento Espressioni e operatori CLEM con supporto di generazione SQL a pag. 63.](#)

Nodi di input multipli. Qualora uno stream abbia più nodi di input del database, la generazione SQL viene applicata ad ogni ramo di input separatamente. Se la generazione viene interrotta su un ramo, può proseguire su un altro. Se due rami si uniscono, e se entrambi possono essere espressi in codice SQL fino all'unione, l'unione stessa può essere sostituita con un join del database e la generazione può proseguire a valle.

Algoritmi del database. La stima del modello viene sempre eseguita su IBM® SPSS® Modeler Server anziché nel database, eccetto nei casi di utilizzo di algoritmi nativi del database da Microsoft, IBM, o Oracle. [Per ulteriori informazioni, vedere l'argomento Panoramica sulla modellazione di database in il capitolo 2 in IBM SPSS Modeler 15 Guida alla modellazione in-database.](#)

Calcolo del punteggio dei modelli. Per alcuni modelli, per il calcolo del punteggio in-database il modello generato viene reso in codice SQL. Tuttavia, alcuni modelli generano espressioni SQL estremamente complesse che non sempre vengono valutate efficacemente all'interno del database. Per questo motivo, è necessario attivare separatamente la generazione SQL per ciascun nodo del modello. Se viene rilevato che un nodo del modello inibisce la generazione SQL, andare alla scheda Impostazioni della finestra di dialogo del nodo e selezionare Genera SQL per questo modello (con alcuni modelli sono disponibili ulteriori opzioni per il controllo della generazione). Eseguire dei test per confermare l'utilità dell'opzione per l'applicazione in uso. [Per ulteriori informazioni, vedere l'argomento Nodo con supporto per generazione SQL a pag. 59.](#)

Quando si eseguono test sui nodi Modelli per verificare se la generazione SQL per i modelli funziona correttamente, si consiglia di salvare prima tutti gli stream da IBM® SPSS® Modeler. Alcuni sistemi di database possono bloccarsi tentando di elaborare il codice SQL generato, potenzialmente complesso, e richiedono quindi la chiusura di SPSS Modeler dal Task Manager di Windows.

Caching del database. Se si sta utilizzando una cache del nodo per salvare i dati in punti critici dello stream (per esempio, dopo un nodo Unione o Aggregazione), assicurarsi che siano attivate sia la cache del database che l'ottimizzazione SQL. In questo modo, nella maggior parte dei casi i dati possono essere inseriti nella cache in una tabella temporanea nel database anziché nel file system. [Per ulteriori informazioni, vedere l'argomento Configurazione dell'ottimizzazione SQL a pag. 55.](#) Per ulteriori informazioni, vedere anche [Opzioni di copia dei dati nella cache per i nodi.](#)

SQL specifici dei fornitori. La maggior parte del codice SQL generato è conforme agli standard (SQL-92), ma talvolta, se utili, vengono sfruttate alcune funzioni non standard specifiche dei fornitori. Il livello di ottimizzazione SQL può variare a seconda della sorgente del database.

Nodo con supporto per generazione SQL

Nelle tabelle che seguono vengono riportati i nodi che rappresentano le operazioni di data mining che supportano la generazione SQL. Ad eccezione dei nodi Modelli database, i nodi che non appaiono in queste tabelle non supportano la generazione SQL.

È possibile visualizzare l'anteprima dell'SQL generato prima di eseguirlo. [Per ulteriori informazioni, vedere l'argomento Anteprima di SQL generato a pag. 56.](#)

Sorgenti

Nodo con supporto per generazione SQL	Note
Database	Questo nodo viene utilizzato per specificare tabelle e viste da utilizzare in analisi più approfondite. Questo nodo consente l'immissione di query SQL. Evita gli insiemi di risultati con nomi di colonna duplicati. Per ulteriori informazioni, vedere l'argomento Scrittura di query SQL a pag. 66.

Operazioni su record

Nodo con supporto per generazione SQL	Note
Select	Supporta la generazione SQL solo se consente l'SQL generato per l'espressione nuovo campo (vedere espressioni di seguito). Se uno qualsiasi dei campi ha dei valori nulli, la generazione SQL non fornisce gli stessi risultati di scarto che si ottengono in IBM® SPSS® Modeler nativo.
Campione	Il campionamento semplice supporta la generazione SQL. Vedere Tabella 6-1 a pag. 60. Il campionamento complesso supporta la generazione SQL per i database Netezza.
Aggrega	Il supporto della generazione SQL dell'aggregazione dipende dal tipo di archiviazione dei dati. Vedere Tabella 6-2 a pag. 61.
Aggregazione RFM	Supporta la generazione, tranne nel caso in cui si effettui il salvataggio della data della seconda o terza transazione più recente, oppure con l'inclusione delle sole transazioni recenti. Tuttavia, l'inclusione delle transazioni recenti funziona se la funzione <code>datetime_date(YEAR,MONTH,DAY)</code> viene rinviata.
Ordina	
Unisci	SQL non generato per l'unione per ordine. L'unione per chiave con join esterno completo o parziale è supportata solo se è supportata dal database/driver. I campi di input non corrispondenti si possono rinominare tramite un nodo Filtro o dalla scheda Filtro di un nodo di input. Supporta la generazione SQL per l'unione in base alla condizione. Per tutti i tipi di unione, <code>SQL_SP_EXISTS</code> non è supportato se gli input hanno origine in database diversi.
Accoda	Supporta la generazione se gli input non sono ordinati.
Elimina dupl.	

Tabella 6-1

Supporto della generazione SQL nel nodo Campione per il campionamento semplice

Modalità	Campione	Dimensione massima	Seme	DB2 OS/Z	DB2 OS/400	DB2 Windows/UNIX	Netezza	Oracle	SQL Server	Teradata
Include	Primo	n/d		Y	Y	Y	Y	Y	Y	Y
	l-ogni-n	off		Y	Y	Y	Y	Y		Y
		max		Y	Y	Y	Y	Y		Y
	% casuale	off	off			Y	Y	Y		Y
			on			Y		Y		
	max	off			Y	Y	Y		Y	

Modalità	Campione	Dimensione massima	Seme	DB2 OS/Z	DB2 OS/400	DB2 Windows/UNIX	Netezza	Oracle	SQL Server	Teradata
			on			Y		Y		
Discard	Primo	off					Y	Y		
		max					Y	Y		
	l-ogni-n	off		Y	Y	Y	Y	Y		Y
		max		Y	Y	Y	Y	Y		Y
	% casuale	off	off			Y	Y	Y		Y
			on			Y		Y		
		max	off			Y	Y	Y		Y
			on			Y		Y		

Tabella 6-2
Supporto per generazione SQL nel nodo Aggregazione

Archiviazione	Sum	Mean	Min	Max	SDev	Median	Count	Variance	Percentile
Intero	Y	Y	Y	Y	Y		Y	Y	
Reale	Y	Y	Y	Y	Y		Y	Y	
Data			Y	Y			Y		
Ora			Y	Y			Y		
Timestamp			Y	Y			Y		
Stringa			Y	Y			Y		

Operazioni su campi

Nodo con supporto per generazione SQL	Note
Type	Supporta la generazione SQL se il nodo Tipo viene istanziato e nessuna verifica di tipo ABORT o WARN viene specificata.
Filtro	
Nuovo campo	Supporta la generazione SQL se consente l'SQL generato per l'espressione nuovo campo (vedere espressioni di seguito).
Risultati classificatore binario	Supporta la generazione SQL per gli obiettivi continui. Per gli altri obiettivi, supporta la generazione solo se viene utilizzato il metodo classificatore binario "Confidenza più elevata vince".
Riempimento	Supporta la generazione SQL se consente l'SQL generato per l'espressione nuovo campo (vedere espressioni di seguito).
Anonimizza	Supporta la generazione SQL per gli obiettivi continui e la generazione SQL parziale per gli obiettivi nominali e flag.
Ricodifica	
Categorizzazione	Supporta la generazione SQL se si utilizza il metodo di discretizzazione "Blocchi (conteggio uguale)" con selezionata l'opzione "Leggi dalla scheda Valori bin se disponibile".
Analisi RFM	Supporta la generazione SQL se è selezionata l'opzione "Leggi dalla scheda Valori bin se disponibile", ma i nodi a valle non la supportano.
Partition	Supporta l'assegnazione di record alle partizioni da parte della generazione SQL.
Crea flag	
Ristruttura	

Grafici

Nodo con supporto per generazione SQL	Note
Lavagna grafica	La generazione SQL è supportata per i seguenti tipi di grafici: Area, Area 3D, Barra, Barra 3D, Barra dei conteggi, Mappa termica, Torta, Torta 3D, Torta dei conteggi. Per gli istogrammi, la generazione SQL è supportata solo per i dati categoriali.
Distribuzione	
Web	
Valutazione	

Insiemi di modelli

Per alcuni modelli, è possibile generare SQL per l'insieme di modelli, rinviando la fase di calcolo del punteggio del modello al database. L'utilizzo principale di questa funzione non è migliorare le prestazioni, ma consentire il rinvio del loro SQL completo agli stream contenenti questi insiemi di modelli. [Per ulteriori informazioni, vedere l'argomento Visualizzazione di SQL per gli insiemi di modelli a pag. 58.](#)

Insieme di modelli con supporto per generazione SQL	Note
C&R Tree	Supporta la generazione SQL per l'opzione albero singolo, ma non per le opzioni di boosting, bagging o insiemi di dati di grandi dimensioni.
QUEST	
CHAID	
C5.0	
Elenco decisionale	
Lineare	Supporta la generazione SQL per l'opzione modello standard, ma non per le opzioni di boosting, bagging o insiemi di dati di grandi dimensioni.
Rete neurale	Supporta la generazione SQL per l'opzione modello standard (solo MLP), ma non per le opzioni di boosting, bagging o insiemi di dati di grandi dimensioni.
Fattoriale/PCA	
Logistico	Supporta la generazione SQL per la procedura multinomiale ma non binomiale. Per il modello multinomiale, la generazione non è supportata se sono selezionate le confidenze, ad eccezione degli obiettivi di tipo flag.
Insiemi di regole generati	

Output

Nodo con supporto per generazione SQL	Note
Tabella	Supporta la generazione se consente la generazione SQL per l'espressione evidenza (vedi espressioni di seguito).
Matrice	Supporta la generazione, tranne nel caso in cui nelle opzioni dei campi sia selezionato "Tutti i valori numerici".

Nodo con supporto per generazione SQL	Note
Analisi	Supporta la generazione, in base alle opzioni selezionate.
Trasforma	
Statistiche	Supporta la generazione se non viene utilizzata l'opzione Correla.
Report	
Calcola globali	

Esporta

Nodo con supporto per generazione SQL	Note
Database	
Publisher	Lo stream pubblicato contiene SQL generato.

Espressioni e operatori CLEM con supporto di generazione SQL

Nelle tabelle che seguono vengono riportate le operazioni e le espressioni matematiche che supportano la generazione SQL e spesso vengono utilizzate durante il data mining. Le operazioni non presenti in queste tabelle non supportano la generazione SQL nella versione corrente.

Operatori

Operazione con supporto di generazione SQL	Note
+	
-	
/	
*	
><	Utilizzato per concatenare le stringhe.

Operatori relazionali

Operazione con supporto di generazione SQL	Note
=	
/=	Utilizzato per specificare "non uguale."
>	
>=	
<	
<=	

Funzioni

Operazione con supporto di generazione SQL	Note
abs	
allbutfirst	

Operazione con supporto di generazione SQL	Note
allbutlast	
and	
arccos	
arcsin	
arctan	
arctanh	
cos	
div	
exp	
fracof	
hasstartstring	
hassubstring	
integer	
intof	
isalphacode	
islowercode	
isnumbercode	
isstartstring	
issubstring	
isuppercode	
last	
length	
locchar	
log	
log10	
lowertoupper	
max	
member	
min	
negate	
not	
number	
or	
pi	
real	
rem	
round	
sign	
sin	
sqrt	
string	
strmember	
subscrs	
substring	

Operazione con supporto di generazione SQL	Note
substring_between	
uppertolower	
to_string	

Funzioni speciali

Operazione con supporto di generazione SQL	Note
@NULL	
@GLOBAL_AVE	Le funzioni globali speciali vengono utilizzate per recuperare i valori globali calcolati dal nodo Globali.
@GLOBAL_SUM	
@GLOBAL_MAX	
@GLOBAL_MEAN	
@GLOBAL_MIN	
@GLOBALSDEV	

Funzioni di aggregazione

Operazione con supporto di generazione SQL	Note
Sum	
Mean	
Min	
Max	
Count	
SDev	

Utilizzo di funzioni SQL in espressioni CLEM

La funzione @SQLFNsi possono utilizzare per aggiungere funzioni SQL con nome in espressioni CLEM al solo scopo dell'esecuzione del database. Questa funzionalità può rivelarsi utile in casi particolari in cui è necessario utilizzare codice SQL proprietario o altre personalizzazioni specifiche per fornitore.

L'uso di questa funzionalità non è contemplato nel contratto di assistenza standard di IBM® SPSS® Modeler, poiché l'esecuzione dipende da componenti di database esterni al di fuori del controllo di IBM Corp., tuttavia può essere implementata in casi speciali, generalmente nell'ambito di un contratto di servizio. Per ulteriori informazioni, rivolgersi a <http://www.ibm.com/software/analytics/spss/services/>.

Scrittura di query SQL

Quando si utilizza il nodo Database, è consigliabile prestare attenzione alle query SQL che possono generare un insieme di dati con nomi di colonna duplicati. Questi nomi duplicati spesso impediscono l'ottimizzazione SQL per qualsiasi nodo a valle.

IBM® SPSS® Modeler utilizza istruzioni **SELECT** nidificate per rinviare l'SQL negli stream che utilizzano una query SQL nel nodo di input Database. In altre parole, lo stream nidifica la query specificata nel nodo di input Database all'interno di una o più istruzioni **SELECT** generate durante l'ottimizzazione dei nodi a valle. Quindi, se l'insieme di risultati di una query contiene nomi di colonna duplicati, non è possibile nidificare un'istruzione tramite RDBMS. Spesso si possono verificare difficoltà di nidificazione durante il join di una tabella in cui una colonna con lo stesso nome viene selezionata in più di una tabella unita. Per esempio, se si considera questa query nel nodo di input:

```
SELECT e.ID, e.LAST_NAME, d.*  
FROM EMP e RIGHT OUTER JOIN  
DEPT d ON e.ID = d.ID;
```

La query impedisce la successiva ottimizzazione SQL, in quanto l'istruzione **SELECT** produce un insieme di dati con due colonne denominate **ID**.

Per consentire l'ottimizzazione SQL completa, è richiesta una forma più esplicita nella scrittura delle query SQL ed è necessario specificare alias di colonna quando una situazione con nomi di colonna duplicati. L'istruzione di seguito illustra una query più esplicita:

```
SELECT e.ID AS ID1, e.LAST_NAME, d.*  
FROM EMP e RIGHT OUTER JOIN  
DEPT d ON e.ID = d.ID;
```

Configurazione di Oracle per piattaforme UNIX

Configurazione di Oracle per l'ottimizzazione SQL

Quando IBM® SPSS® Modeler Server è in esecuzione sulle piattaforme UNIX e legge da un database Oracle, è consigliabile seguire i suggerimenti seguenti per essere certi che l'SQL generato sia ottimizzato interamente all'interno del database.

Specifiche delle opzioni internazionali

Se SPSS Modeler Server viene eseguito con opzioni internazionali diverse da quelle contenute nei driver Connect ODBC, è possibile riconfigurare il computer per incrementare l'ottimizzazione SQL. I driver Connect ODBC vengono distribuiti solo con i file delle opzioni internazionali *en_US*. Di conseguenza, se SPSS Modeler Server viene eseguito con opzioni internazionali diverse o se queste ultime non sono state definite completamente nella shell in cui SPSS Modeler Server è stato avviato, non è possibile ottimizzare completamente l'SQL generato in Oracle. I motivi sono i seguenti:

- SPSS Modeler Server utilizza i file delle opzioni internazionali ODBC corrispondenti a quelle in cui viene eseguito per tradurre i codici restituiti dal database in stringhe di testo. Quindi le utilizza per determinare a quale database connettersi.
- Se le opzioni internazionali (così come vengono restituite a SPSS Modeler Server dalla query `$LANG` del sistema) non sono *en_US*, IBM® SPSS® Modeler non può tradurre in testo i codici che riceve dal driver ODBC. In altre parole, all'avvio di una connessione database, a SPSS Modeler Server viene restituito un codice non tradotto, anziché la stringa *Oracle*. Ciò significa che SPSS Modeler non è in grado di ottimizzare stream per Oracle.

Per controllare e reimpostare le specifiche delle opzioni internazionali:

- ▶ In una shell UNIX, eseguire:

```
#locale
```

Vengono restituite le informazioni relative alle opzioni internazionali per la shell. Per esempio:

```
$ locale
LANG=en_US.ISO8859-15
LC_CTYPE="en_US.ISO8859-15"
LC_NUMERIC="en_US.ISO8859-15"
LC_TIME="en_US.ISO8859-15"
LC_COLLATE="en_US.ISO8859-15"
LC_MONETARY="en_US.ISO8859-15"
LC_MESSAGES="en_US.ISO8859-15"
LC_ALL=en_US.ISO8859-15
```

- ▶ Passare alla directory Connect ODBC/locale. Qui si trova una directory singola, *en_US*.
- ▶ Creare un collegamento di tipo soft alla directory *en_US*, specificando il nome delle opzioni internazionali nella shell. Di seguito è riportato un esempio:

```
#ln -s en_US en_US.IS08859-15
```

Per opzioni internazionali non inglesi, quali *fr_FR.IS08859-1*, è possibile creare il collegamento di tipo soft come segue:

```
#ln -s en_US fr_FR.IS08859-1
```

- ▶ Una volta creato il collegamento, riavviare SPSS Modeler Server dalla stessa shell. (SPSS Modeler Server riceve le informazioni sulle impostazioni internazionali dalla shell dalla quale viene avviato.

Note

Durante l'ottimizzazione di un computer UNIX per i rinvii SQL ad Oracle, tenere presente quanto segue:

- È necessario specificare le opzioni internazionali in modo completo. Nell'esempio riportato in alto, è necessario creare il collegamento nella forma *language_territory.code-page*. Quindi la directory esistente delle opzioni internazionali *en_US* non è sufficiente.
- Per ottimizzare al meglio il mining in-database, è necessario definire sia *LANG* sia *LC_ALL* nella shell usata per avviare SPSS Modeler Server. È possibile non definire *LANG* nella shell come per altre variabili di ambiente prima di riavviare SPSS Modeler Server. Per esempio, vedere la definizione seguente:

```
#LANG=en_US.IS08859-15; export LANG
```

- A ogni avvio di SPSS Modeler Server, è necessario verificare che le informazioni relative alle opzioni internazionali della shell siano state specificate completamente e che il collegamento di tipo soft appropriato esista nella directory ODBC/locale.

Configurazione degli script di avvio di UNIX

Introduzione

Questa appendice descrive alcuni degli script forniti con le versioni UNIX di IBM® SPSS® Modeler Server e illustra come configurarli. Gli script vengono utilizzati per:

- Configurare SPSS Modeler Server per avviarsi automaticamente al riavvio del computer server.
- Arrestare e riavviare manualmente SPSS Modeler Server.
- Modificare le autorizzazioni relative ai file creati da SPSS Modeler Server.
- Configurare SPSS Modeler Server per funzionare con i driver ODBC Connect forniti con il programma SPSS Modeler Server. [Per ulteriori informazioni, vedere l'argomento IBM SPSS Modeler Server e Data Access Pack a pag. 71.](#)
- Configurare SPSS Modeler Server per utilizzare IBM® SPSS® Collaboration and Deployment Services Enterprise View Driver durante la lettura di dati con un nodo Enterprise View. [Per ulteriori informazioni, vedere l'argomento Configurazione di un driver per il nodo Enterprise View a pag. 78.](#)

Script

IBM® SPSS® Modeler Server utilizza numerosi script, fra cui:

- **modelersrv.sh.** Lo script di avvio manuale di SPSS Modeler Server è contenuto nella directory di installazione di SPSS Modeler Server. Esso configura l'ambiente del server quando il processo daemon del server viene avviato *manualmente*. Eseguirlo quando si desidera avviare e arrestare manualmente il server. Modificarlo quando è necessario cambiare la configurazione per l'avvio manuale.
- **auto.sh.** Questo script configura il sistema per avviare automaticamente il processo daemon del server all'avvio. Eseguirlo una sola volta per configurare il sistema per l'avvio automatico. Non è necessario modificarlo. Lo script è contenuto nella directory di installazione di SPSS Modeler Server.
- **rc.modeler.** Quando si esegue lo script *auto.sh*, il sistema crea questo script in una posizione che dipende dal sistema operativo del server in uso. Esso configura l'ambiente del server quando questo viene avviato *automaticamente*. Modificarlo quando è necessario cambiare la configurazione per l'avvio automatico.

Tabella B-1

Posizione dello script *rc.modeler* a seconda del sistema operativo

Sistema operativo	Posizione
AIX	<i>/etc/rc.modeler</i>
HP-UX	<i>/sbin/init.d/rc.modeler</i>
Solaris	<i>/etc/init.d/rc.modeler</i>

Avvio e arresto automatico di IBM SPSS Modeler Server

IBM® SPSS® Modeler Server deve essere avviato come processo daemon. Il programma di installazione comprende uno script (*auto.sh*) che è possibile eseguire per configurare il sistema per l'arresto e il riavvio automatico di SPSS Modeler Server.

Per configurare il sistema per l'avvio e l'arresto automatico

- ▶ Accedere come utente root.
- ▶ Passare alla directory di installazione di SPSS Modeler Server.
- ▶ Eseguire lo script. Al prompt di UNIX digitare:

```
./auto.sh
```

Nella posizione riportata nella tabella sopra viene creato uno script di avvio automatico, *rc.modeler*. Il sistema operativo utilizzerà *rc.modeler* per avviare il processo daemon di SPSS Modeler Server ogni volta che viene riavviato il computer server. Lo script *rc.modeler* sarà utilizzato anche per arrestare il daemon ogni volta che viene arrestato il sistema.

Avvio e arresto manuale di IBM SPSS Modeler Server

È possibile avviare e arrestare manualmente IBM® SPSS® Modeler Server eseguendo lo script *modelersrv.sh*.

Per avviare e arrestare manualmente IBM SPSS Modeler Server

- ▶ Passare alla directory di installazione di SPSS Modeler Server.
- ▶ Per avviare il server, al prompt dei comandi di UNIX digitare:

```
./modelersrv.sh start
```

- ▶ Per arrestare il server, al prompt dei comandi di UNIX digitare:

```
./modelersrv.sh stop
```

Modifica degli script

Se si utilizza sia l'avvio manuale che quello automatico, apportare modifiche parallele sia allo script *modelersrv.sh*, sia allo script *rc.modeler*. Se si utilizza solo l'avvio manuale, apportare le modifiche a *modelersrv.sh*. Se si utilizza solo l'avvio automatico, apportare le modifiche a *rc.modeler*.

Per modificare gli script

- ▶ Arrestare IBM® SPSS® Modeler Server. (Per ulteriori informazioni, vedere l'argomento [Avvio e arresto manuale di IBM SPSS Modeler Server](#) a pag. 70.)
- ▶ Individuare la posizione dello script desiderato. (Per ulteriori informazioni, vedere l'argomento [Script](#) a pag. 69.)
- ▶ Aprire lo script in un editor di testo, apportare le modifiche e salvare il file.
- ▶ Avviare SPSS Modeler Server, automaticamente (riavviando il computer server) o manualmente.

Verifica delle autorizzazioni per la creazione di file

IBM® SPSS® Modeler Server crea file temporanei dotati di autorizzazione in lettura, scrittura ed esecuzione per tutti gli utenti. Questa impostazione di default può essere ignorata modificando l'impostazione UMASK dello script di avvio in *modelersrv.sh*, in *rc.modeler* o in entrambi gli script. (per ulteriori informazioni, vedere [Modifica degli script](#) sopra). Si consiglia di utilizzare 077 come impostazione UMASK più restrittiva. Impostazioni più restrittive potrebbero causare problemi di autorizzazioni per SPSS Modeler Server.

IBM SPSS Modeler Server e Data Access Pack

Per utilizzare i driver ODBC con IBM® SPSS® Modeler Server, è necessario che l'ambiente ODBC sia configurato mediante *odbc.sh* quando viene avviato il processo SPSS Modeler Server. È possibile eseguire questa operazione modificando lo script di avvio di IBM® SPSS® Modeler appropriato: *modelersrv.sh*, *rc.modeler* o entrambi. (Per ulteriori informazioni, vedere l'argomento [Modifica degli script](#) a pag. 71.)

Per ulteriori informazioni, vedere il sito Web del supporto tecnico all'indirizzo <http://www.ibm.com/support>. Per domande sulla creazione o l'impostazione di autorizzazioni per sorgenti dati ODBC, rivolgersi all'amministratore del database.

Per configurare ODBC per l'avvio insieme a IBM SPSS Modeler Server

- ▶ Arrestare l'host SPSS Modeler Server, se è in esecuzione.
- ▶ Scaricare l'archivio tar compresso adatto alla piattaforma su cui SPSS Modeler Server è installato. Assicurarsi di scaricare i driver per la versione di SPSS Modeler Server installata. Copiare il file nel percorso in cui si desidera installare i driver ODBC (per esempio, */usr/spss/odbc*).
- ▶ Estrarre il file.

- ▶ Eseguire lo script *setodbcpath.sh* estratto dal file.
- ▶ Modificare lo script *odbc.sh* per aggiungere la definizione di ODBCINI nella parte inferiore dello script ed esportarlo, per esempio:

```
ODBCINI=/usr/spss/odbc/odbc.ini; export ODBCINI
```

ODBCINI deve puntare al percorso completo del file *odbc.ini* che si desidera venga letto da SPSS Modeler per ottenere l'elenco delle sorgenti dati ODBC definite (un file *odbc.ini* di default viene installato con i driver).

- ▶ Salvare *odbc.sh*.
- ▶ (solo per le installazioni a 64 bit di SPSS Modeler Server; per le altre installazioni, procedere con il passaggio successivo) Definire ed esportare LD_LIBRARY_PATH_64 in *odbc.sh*:

```
if [ "$LD_LIBRARY_PATH_64" = "" ]; then
    LD_LIBRARY_PATH_64=<library_path>
else
    LD_LIBRARY_PATH_64=<library_path>:$LD_LIBRARY_PATH_64
fi
export LD_LIBRARY_PATH_64
```

dove *library_path* è lo stesso percorso utilizzato per la definizione LD_LIBRARY_PATH già presente nello script inizializzato con il percorso di installazione (per esempio, */usr/spss/odbc/lib*). Il metodo più semplice consiste nel copiare le istruzioni if ed export per LD_LIBRARY_PATH nel file *odbc.sh*, aggiungerle alla fine del file e sostituire le stringhe "LD_LIBRARY_PATH" nelle nuove istruzioni if ed export aggiunte con "LD_LIBRARY_PATH_64".

Il file *odbc.sh* risultante in un'installazione di SPSS Modeler Server a 64 bit sarà:

```
if [ "$LD_LIBRARY_PATH" = "" ]; then
    LD_LIBRARY_PATH=/usr/spss/odbc/lib
else
    LD_LIBRARY_PATH=/usr/spss/odbc/lib:$LD_LIBRARY_PATH
fi
export LD_LIBRARY_PATH
if [ "$LD_LIBRARY_PATH_64" = "" ]; then
    LD_LIBRARY_PATH_64=/usr/spss/odbc/lib
else
    LD_LIBRARY_PATH_64=/usr/spss/odbc/lib:$LD_LIBRARY_PATH_64
fi
export LD_LIBRARY_PATH_64
ODBCINI=/usr/spss/odbc/odbc.ini; export ODBCINI
```

Ricordarsi di esportare LD_LIBRARY_PATH_64 e di definirlo con il ciclo if.

- ▶ Modificare il file *odbc.ini* definito in precedenza utilizzando \$ODBCINI. Definire i nomi delle sorgenti dati richieste (i nomi variano a seconda del database a cui si ha accesso).
- ▶ Salvare il file *odbc.ini*.

- Configurare SPSS Modeler Server per l'utilizzo dei driver. A questo scopo, modificare *modelersrv.sh* e aggiungere la riga seguente immediatamente sotto la riga che definisce SCLEMDNAME:

```
. <odbc.sh_path>
```

dove *odbc.sh_path* è il percorso completo del file *odbc.sh* modificato nella fase iniziale di questa procedura, per esempio:

```
./usr/spss/odbc/odbc.sh
```

Nota: è importante rispettare la sintassi. Assicurarsi di lasciare uno spazio tra il primo punto e il percorso del file.

- Salvare *modelersrv.sh*.

Per verificare la connessione

- Riavviare SPSS Modeler Server.
- Connettersi a SPSS Modeler Server da un client.
- Sul client, aggiungere un nodo di input Database nell'area di disegno.
- Aprire il nodo e verificare la presenza dei nomi delle sorgenti dati definiti nel file *odbc.ini* nella fase precedente della procedura di configurazione.

Se il risultato non è quello previsto o si verificano errori quando si tenta di connettersi a una sorgente dati definita, attenersi alla procedura di risoluzione dei problemi. [Per ulteriori informazioni, vedere l'argomento Risoluzione dei problemi di una configurazione ODBC a pag. 74.](#)

Per configurare ODBC per l'avvio insieme a IBM SPSS Modeler Solution Publisher Runtime

Quando la connessione al database può essere stabilita da SPSS Modeler Server, sarà possibile configurare un'installazione di IBM® SPSS® Modeler Solution Publisher Runtime sullo stesso server facendo riferimento allo stesso script *odbc.sh* dallo script di avvio di SPSS Modeler Solution Publisher Runtime.

- Modificare lo script *modelerrun* in SPSS Modeler Solution Publisher Runtime per aggiungere la riga seguente immediatamente sopra l'ultima riga dello script:

```
. <odbc.sh_path>
```

dove *odbc.sh_path* è il percorso completo del file *odbc.sh* modificato nella fase iniziale di questa procedura, per esempio:

```
./usr/spss/odbc/odbc.sh
```

Nota: è importante rispettare la sintassi. Assicurarsi di lasciare uno spazio tra il primo punto e il percorso del file.

- Salvare il file di script *modelerrun*.

Per configurare ODBC per l'avvio insieme a IBM SPSS Modeler Batch

Non è necessario eseguire alcuna configurazione dello script IBM® SPSS® Modeler Batch per ODBC. La connessione a SPSS Modeler Server viene infatti effettuata da SPSS Modeler Batch per eseguire gli stream. Verificare che la configurazione ODBC di SPSS Modeler Server sia stata eseguita e funzioni correttamente, come descritto in questa sezione.

Per aggiungere o modificare il nome di una sorgente dati

- ▶ Modificare il file *odbc.ini* per inserire il nome nuovo o modificato.
- ▶ Verificare la connessione come descritto precedentemente in questa sezione.

Se la connessione con SPSS Modeler Server funziona correttamente, anche la sorgente dati nuova o modificata dovrebbe funzionare correttamente con SPSS Modeler Solution Publisher Runtime e SPSS Modeler Batch.

Risoluzione dei problemi di una configurazione ODBC**Nessuna sorgente dati elencata o visualizzazione di un testo casuale**

Se si apre un nodo di input Database e l'elenco delle sorgenti dati disponibili è vuoto o contiene immissioni inattese, il problema può essere dovuto allo script di avvio.

- ▶ Verificare che \$ODBCINI sia definito all'interno di *modelersrv.sh* esplicitamente nello script oppure nello script *odbc.sh* cui viene fatto riferimento in *modelersrv.sh*.
- ▶ Nel secondo caso, assicurarsi che ODBCINI punti al percorso completo del file *odbc.ini* utilizzato per la definizione delle sorgenti dati ODBC.
- ▶ Se la specifica del percorso in ODBCINI è corretta, verificare il valore di \$ODBCINI utilizzato nell'ambiente di IBM® SPSS® Modeler Server ripetendo la variabile dall'interno di *modelersrv.sh*. A questo scopo, aggiungere la riga seguente in *modelersrv.sh* dopo il punto in cui viene definito ODBCINI:

```
echo $ODBCINI
```

- ▶ Salvare ed eseguire *modelersrv.sh*. Il valore di \$ODBCINI impostato nell'ambiente di SPSS Modeler Server viene scritto in *stdout* per la verifica.
- ▶ Se non viene restituito alcun valore in *stdout* e la definizione di \$ODBCINI viene eseguita nello script *odbc.sh* cui viene fatto riferimento da *modelersrv.sh*, verificare che la sintassi del riferimento sia corretta. La sintassi dovrebbe essere:

```
. <odbc.sh_path>
```

dove *odbc.sh_path* è il percorso completo del file *odbc.sh* modificato nella fase iniziale di questa procedura, per esempio:

```
./usr/spss/odbc/odbc.sh
```

Nota: è importante rispettare la sintassi. Assicurarsi di lasciare uno spazio tra il primo punto e il percorso del file.

Quando il valore corretto viene ripetuto in *stdout* nel file *modelersrv.sh* in esecuzione, dovrebbe essere possibile visualizzare i nomi delle sorgenti dati nel nodo di input Database quando si riavvia SPSS Modeler Server e si stabilisce la connessione dal client.

Il client IBM SPSS Modeler si arresta quando si fa clic su Connetti nella finestra di dialogo Connessioni database

Questo comportamento può essere causato da un'impostazione errata del percorso della libreria che non include il percorso alle librerie ODBC. Il percorso della libreria è definito da `$LD_LIBRARY_PATH` (e `$LD_LIBRARY_PATH_64` per le versioni a 64 bit).

Per verificare il valore del percorso della libreria nell'ambiente daemon SPSS Modeler Server, ripetere il valore della variabile di ambiente appropriata dall'interno di *modelersrv.sh* dopo la riga in cui viene aggiunto il percorso delle librerie ODBC al percorso della libreria ed eseguire lo script. Il valore del percorso della libreria verrà ripetuto sul terminale alla successiva esecuzione dello script.

Se viene fatto riferimento a *odbc.sh* da *modelersrv.sh* per impostare l'ambiente ODBC SPSS Modeler Server, ripetere il percorso della libreria dalla riga dopo quella in cui viene fatto riferimento allo script *odbc.sh*. Per ripetere il valore, aggiungere la riga seguente nello script, quindi salvare ed eseguire il file di script:

```
echo $<library_path_variable>
```

dove *<library_path_variable>* è la variabile del percorso di libreria appropriata per il sistema operativo del server.

Il valore restituito del percorso della libreria deve includere il percorso alla sottodirectory *lib* dell'installazione ODBC. In caso contrario, aggiungere il percorso al file.

Se si esegue la versione a 64 bit di SPSS Modeler Server, `$LD_LIBRARY_PATH_64` avrà priorità su `$LD_LIBRARY_PATH`, se impostata. Se il problema si verifica in una di queste piattaforme a 64 bit, ripetere `LD_LIBRARY_PATH_64` e `$LD_LIBRARY_PATH` da *modelersrv.sh* e, se necessario, impostare `$LD_LIBRARY_PATH_64` in modo che includa la sottodirectory *lib* dell'installazione ODBC ed esportare la definizione.

Nome della sorgente dati non trovato e nessun driver di default specificato

Se si verifica questo errore quando si fa clic su Connetti nella finestra di dialogo Connessioni database, significa in genere che il file *odbc.ini* non è definito correttamente. Verificare che il nome della sorgente dati (DSN) definito nella sezione [ODBC Data Sources] all'inizio del file corrisponda alla stringa specificata tra le parentesi quadre più avanti in *odbc.ini* per definire il DSN. Se i nomi sono diversi, si verificherà questo tipo di errore quando si tenta di connettersi

utilizzando il DSN dall'interno di IBM® SPSS® Modeler. Di seguito è riportato un esempio di specifica *non corretta*:

```
[ODBC Data Sources]
Oracle=Oracle Wire Protocol

....
....
[Oracle Driver]
Driver=/usr/ODBC/lib/XEora22.so
Description=SPSS 5.2 Oracle Wire Protocol
AlternateServers=
....
```

Sarà necessario modificare una delle due stringhe in grassetto in modo che corrispondano esattamente. In questo modo il problema si dovrebbe risolvere.

Non è possibile caricare il driver specificato

Anche questo tipo di errore indica che il file *odbc.ini* non è definito correttamente. È possibile che il parametro Driver all'interno del paragrafo dei driver non sia impostato correttamente, ad esempio:

```
[ODBC Data Sources]
Oracle=Oracle Wire Protocol

....
....
[Oracle]
Driver=/nosuchpath/ODBC/lib/XEora22.so
Description=SPSS 5.2 Oracle Wire Protocol
AlternateServers=
```

- ▶ Verificare che l'oggetto condiviso specificato dal parametro Driver esista.
- ▶ Se errato, correggere il percorso dell'oggetto condiviso.
- ▶ Se il parametro Driver viene specificato nel seguente formato:

```
Driver=ODBCHOME/lib/XEora22.so
```

significa che gli script ODBC non sono stati inizializzati. Eseguire lo script *setodbcpath.sh* installato con i driver. [Per ulteriori informazioni, vedere l'argomento IBM SPSS Modeler Server e Data Access Pack a pag. 71](#). Dopo aver eseguito lo script, la stringa "ODBCHOME" dovrebbe essere stata sostituita dal percorso all'installazione ODBC. In questo modo il problema si dovrebbe risolvere.

È anche possibile che l'errore sia causato da un problema della libreria del driver. Utilizzare lo strumento *ivtestlib* fornito con ODBC per verificare che sia effettivamente impossibile caricare il driver. Per Connect64, utilizzare lo strumento *ddtestlib*. Risolvere il problema impostando la variabile del percorso delle librerie nello script di avvio.

Per esempio, se è impossibile caricare il driver di Oracle per un'installazione a 32 bit, effettuare le seguenti operazioni:

- Utilizzare `ivtestlib` per verificare che sia impossibile caricare il driver. Per esempio, al prompt dei comandi di UNIX, digitare:

```
sh
cd ODBC DIR
. odbclnt.sh
./bin/ivtestlib MFor815
```

dove `ODBC DIR` deve essere sostituito dal percorso della directory di installazione di ODBC.

- Leggere il messaggio per individuare l'eventuale errore. Per esempio, il messaggio:
Load of MFor815.so failed: ld.so.1: bin/ivtestlib: fatal: libclntsh.so: open failed: No such file or directory indica che la libreria client di Oracle `libclntsh.so` manca o non si trova nel percorso delle librerie (per esempio, non si trova in `LD_LIBRARY_PATH` in Solaris).
- Verificare che la libreria esista. Se non esiste, reinstallare il client Oracle. Se la libreria è presente, digitare la sequenza di comandi riportata di seguito dal prompt dei comandi di UNIX:

```
LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/bigdisk/oracle/product/8.1.6/lib
export LD_LIBRARY_PATH
./bin/ivtestlib Mfor815
```

dove `/bigdisk/oracle/product/8.1.6/lib` deve essere sostituito dal percorso di `libclntsh.so` e `LD_LIBRARY_PATH` è la variabile del percorso delle librerie per il sistema operativo in uso.

Si noti che se si sta eseguendo SPSS Modeler a 64 bit in Linux o Solaris, la variabile del percorso delle librerie contiene il suffisso `_64`. Le prime due righe dell'esempio precedente pertanto diventeranno:

```
LD_LIBRARY_PATH_64=$LD_LIBRARY_PATH_64:/bigdisk/oracle/product/8.1.6/lib
export LD_LIBRARY_PATH_64
```

- Leggere il messaggio per verificare che sia ora possibile caricare il driver. Per esempio, il messaggio:
Load of MFor815.so successful, qehandle is 0xFF3A1BE4
indica che la libreria client di Oracle può essere caricata.
- Correggere il percorso della libreria nello script di avvio di SPSS Modeler.
- Riavviare SPSS Modeler Server con lo script di avvio modificato in precedenza (`modelersrv.sh` o `rc.modeler`).

Percorsi delle librerie

Il nome della variabile del percorso delle librerie varia a seconda del sistema operativo. La tabella che segue contiene indicazioni utili per effettuare le sostituzioni appropriate quando si eseguono operazioni di configurazione o di risoluzione dei problemi sul proprio sistema.

Tabella B-2
Percorso delle librerie in base al sistema operativo

Sistema operativo	Nome variabile del percorso delle librerie
AIX	LIBPATH
HP-UX	SHLIB_PATH
Solaris o Linux a 32 bit	LD_LIBRARY_PATH
Solaris o Linux a 64 bit	LD_LIBRARY_PATH_64

Configurazione di un driver per il nodo Enterprise View

Il nodo Enterprise View consente di creare e di mantenere una connessione tra una sessione di IBM® SPSS® Modeler e una IBM SPSS Collaboration and Deployment Services Enterprise View in un IBM® SPSS® Collaboration and Deployment Services Repository condiviso. [Per ulteriori informazioni, vedere l'argomento Archiviazione e deployment di oggetti di repository in il capitolo 9 in Manuale dell'utente di IBM SPSS Modeler 15.](#)

Per utilizzare il nodo Enterprise View con IBM® SPSS® Modeler Server, è necessario installare e configurare un IBM® SPSS® Collaboration and Deployment Services Enterprise View Driver. Per dettagli sull'installazione del driver, rivolgersi all'amministratore locale. Al termine dell'installazione, è necessario aggiungere un riferimento allo script *pev.sh* nello script di avvio, in *modelersrv.sh*, *rc.modeler* o in entrambi, attenendosi al seguente formato:

```
./usr/odbc/pev.sh
```



Configurazione ed esecuzione di SPSS Modeler Server come processo non root in UNIX

Introduzione

Le istruzioni riportate di seguito contengono informazioni sull'esecuzione di IBM® SPSS® Modeler Server come processo non root su sistemi UNIX.

Esecuzione come root. L'installazione di default di SPSS Modeler Server presuppone che il processo daemon del server venga eseguito come root. L'esecuzione come root consente a IBM® SPSS® Modeler di autenticare in modo affidabile l'accesso di ogni utente e di avviare ogni sessione utente sull'account utente UNIX corrispondente. In questo modo gli utenti possono accedere solo ai file e alle directory di loro competenza. Tuttavia, è possibile eseguire il processo daemon su un account non root seguendo le istruzioni riportate di seguito.

Esecuzione come non root. L'esecuzione di SPSS Modeler Server come processo non root consente di impostare gli ID utente effettivi del processo daemon del server su un account prescelto. Tutte le sessioni utente avviate da SPSS Modeler Server utilizzeranno il medesimo account UNIX, il che significa che tutti i file di dati letti o scritti da SPSS Modeler saranno condivisi da tutti gli utenti di SPSS Modeler. Questo non influisce sull'accesso ai dati del database, poiché gli utenti devono autenticarsi in modo indipendente per tutte le sorgenti dati del database che utilizzano. Senza i privilegi root, SPSS Modeler opera in uno dei due modi seguenti:

- **senza un database privato di password.** Con questo metodo, SPSS Modeler utilizza il database di password UNIX, NIS o server LDAP normalmente utilizzato per l'autenticazione degli utenti sul sistema UNIX. [Per ulteriori informazioni, vedere l'argomento Configurazione non root senza un database privato di password a pag. 79.](#)
- **con un database privato di password.** Con questo metodo, SPSS Modeler esegue l'autenticazione degli utenti facendo riferimento a un database privato di password diverso dal database di password UNIX, NIS o server LDAP normalmente utilizzato per l'autenticazione degli utenti su UNIX. [Per ulteriori informazioni, vedere l'argomento Configurazione non root con un database di password privato a pag. 80.](#)

Configurazione non root senza un database privato di password

Per configurare IBM® SPSS® Modeler Server per l'esecuzione su un account non root senza un database di password privato, seguire la procedura descritta di seguito:

1. Aprire il file *options.cfg* di SPSS Modeler Server per la modifica.
2. Impostare l'opzione `start_process_as_login_user` su Y.

3. Salvare e chiudere il file *options.cfg*.

Attenzione: non abilitare l'impostazione `start_process_as_login_user` e avviare SPSS Modeler Server come *root*. Così facendo i processi di tutti gli utenti connessi al server verrebbero eseguiti come *root*; questa configurazione rappresenta un rischio per la sicurezza del sistema.

Configurazione non root con un database di password privato

Se si sceglie l'autenticazione degli utenti mediante un database di password privato, tutte le sessioni utente verranno avviate sullo stesso account utente non root. Per avviare e arrestare il server e per eseguire operazioni generiche di amministrazione, è necessario comunque accedere come *root*, o utilizzare il comando su *root*.

Per configurare IBM® SPSS® Modeler Server per l'esecuzione su un account non root, effettuare le seguenti operazioni:

1. Creare un gruppo per contenere tutti gli utenti. Al gruppo è possibile assegnare un nome qualunque; in questo esempio sarà utilizzato il nome *modelerusers*.
2. Creare l'account utente su cui eseguire SPSS Modeler Server. L'utilizzo di questo account è riservato al processo daemon di SPSS Modeler Server: in altre parole, l'account non può essere utilizzato per avviare e arrestare il server o per eseguire operazioni di amministrazione generiche in IBM® SPSS® Modeler. In questo esempio, all'account viene assegnato il nome *modelerserv*.

Quando si crea l'account, tenere presente quanto segue:

- L'account non necessita di funzionalità di login.
 - Il gruppo principale deve essere il gruppo `<modelerusers>` creato in precedenza.
 - La home directory può essere la directory di installazione di SPSS Modeler o qualsiasi altra directory di default comoda da utilizzare (è opportuno utilizzare una directory diversa da quella di installazione se si desidera che l'account resti valido anche dopo un aggiornamento del programma).
3. Configurare quindi gli script di avvio per avviare SPSS Modeler Server utilizzando l'account appena creato. Individuare lo script di avvio appropriato e aprirlo in un editor di testo. [Per ulteriori informazioni, vedere l'argomento Script in l'appendice B a pag. 69.](#)
 - ▶ All'interno dello script, modificare la riga che imposta la variabile `CLEMUSER` in modo che faccia riferimento al nome dell'account utente creato per l'esecuzione di SPSS Modeler Server. Per esempio:

```
CLEMUSER=modelerserv
```
 - ▶ Modificare l'impostazione `umask` per consentire al gruppo almeno l'accesso in lettura ai file creati:

```
umask 027
```
 4. Modificare il file delle opzioni del server *config/options.cfg* per impostare l'autenticazione in base al database privato di password accodando la riga:

```
authentication_methods, "pasw_modeler"
```

5. In seguito sarà necessario creare un database privato di password archiviato nel file *config/passwords.cfg*. Il file delle password definisce le combinazioni di nome utente/password a cui è consentito accedere a SPSS Modeler. *Nota:* queste password sono riservate a SPSS Modeler e non hanno alcun rapporto con i nomi utente e le password utilizzati per accedere a UNIX. Per comodità è possibile utilizzare gli stessi nomi utente, ma non le stesse password.

Per creare il file delle password è necessario utilizzare l'apposito programma di utilità, *pwutil*, contenuto nella directory *bin* dell'installazione di SPSS Modeler Server. La sinossi di questo programma è la seguente:

```
pwutil [ nome utente [ password ] ]
```

Il programma prende un nome utente e una password di solo testo e scrive il nome utente e la password cifrata sull'output standard in un formato adatto per essere incluso nel file delle password. Per esempio, per definire un utente *modeler* con la password "data mining" occorre digitare:

```
bin/pwutil modeler "data mining" > config/passwords.cfg
```

Nella maggior parte dei casi è sufficiente definire un unico nome utente, per consentire a tutti gli utenti di accedere con lo stesso nome e la stessa password. È possibile tuttavia creare utenti aggiuntivi accodando i singoli utenti al file mediante l'operatore >>, per esempio:

```
bin/pwutil modeler "data miner2" >> config/passwords.cfg
```

Nota: se viene utilizzato un solo operatore >, il contenuto del file *passwords.cfg* sarà sovrascritto ogni volta, con la sostituzione di tutti gli utenti impostati in precedenza. Ricordare che, indipendentemente da questo, tutti gli utenti condividono lo stesso account utente UNIX.

6. Modificare in modo ricorsivo i proprietari della directory di installazione di SPSS Modeler e dei relativi contenuti impostandoli come utente <*modelerserv*> e gruppo <*modelerusers*>, dove i nomi a cui si fa riferimento sono quelli creati in precedenza. Per esempio:

```
chown -R -h modelerserv:modelerusers .
```

7. Nella directory dei dati può essere utile creare delle sottodirectory per gli utenti di SPSS Modeler, in modo che questi ultimi possano disporre di cartelle dove archiviare i dati di lavoro senza interferenze. Queste directory devono essere di proprietà del gruppo <*modelerusers*> e disporre di autorizzazioni di gruppo per la lettura, scrittura e ricerca. Per esempio, per creare una directory di lavoro per l'utente *bob*:

```
mkdir data/bob
chown bob:modelerusers data/bob
chmod ug=rwx,o= data/bob
```

È possibile inoltre impostare il bit set-group-ID sulla directory in modo che tutti i file di dati copiati al suo interno diventino automaticamente di proprietà del gruppo <*modelerusers*>:

```
chmod g+s data/bob
```

Esecuzione di SPSS Modeler Server come utente non root

Per eseguire SPSS Modeler Server come utente non root, seguire la procedura descritta di seguito:

1. Accedere usando l'account utente non root creato in precedenza.
2. Se si esegue il sistema con l'opzione `start_process_as_login_user` del file di configurazione abilitata, è possibile avviare, arrestare e verificare lo stato di SPSS Modeler Server. [Per ulteriori informazioni, vedere l'argomento Per avviare, arrestare e verificare lo stato di Clementine Server in UNIX in il capitolo 4 a pag. 23.](#)

Gli utenti finali si connettono a SPSS Modeler Server eseguendo l'accesso dal software client. Gli utenti finali devono ricevere le informazioni necessarie per la connessione, compreso l'indirizzo IP o il nome host del server.

Esecuzione di IBM SPSS Modeler Server come utente non root su AIX

Quando si avvia IBM® SPSS® Modeler Server come utente non root su AIX, esiste la possibilità che l'errore `lsattr not found` venga visualizzato nella console. Questo perché `lsattr` si trova nella directory `/usr/sbin`, ma quando SPSS Modeler Server consente l'accesso a un nuovo utente (non root) reimposta il percorso in modo da includervi solo `/usr/bin` e `/bin`.

La soluzione a questo errore consiste nel realizzare il seguente collegamento in `/usr/bin`:

```
su
cd /usr/bin
ln -s /usr/sbin/lsattr lsattr
```

Dopo aver realizzato il collegamento, riavviare SPSS Modeler Server.

Risoluzione dei problemi di mancata autenticazione degli utenti

A seconda del tipo di configurazione di autenticazione definita sul sistema operativo, con una configurazione non root potrebbero verificarsi dei problemi di accesso a SPSS Modeler Server. Ad esempio, un problema di questo tipo potrebbe verificarsi se il sistema operativo è configurato (mediante il file `/etc/nsswitch.conf` o uno analogo) per la consultazione del file locale delle shadow password piuttosto che per l'utilizzo di NIS o LDAP. Questa circostanza si verifica perché SPSS Modeler Server richiede l'accesso in lettura ai file utilizzati per eseguire l'autenticazione, compreso il file `/etc/shadow` o un suo equivalente (ad esempio, `/etc/security/passwd` su AIX), che registra le informazioni protette degli account utente. Tuttavia, le autorizzazioni per i file del sistema operativo sono generalmente definite in modo che `/etc/shadow` sia accessibile solo per l'utente root. In questo scenario, un processo non root non può leggere `/etc/shadow` per convalidare le password degli utenti, con il risultato che l'autenticazione viene negata.

Il problema può essere risolto in vari modi:

- Chiedere all'amministratore di sistema di configurare sul sistema operativo l'autenticazione tramite NIS o LDAP.

- Cambiare le autorizzazioni per i file protetti, ad esempio concedendo l'accesso in lettura al file */etc/shadow*, in modo che l'account utente locale utilizzato per eseguire SPSS Modeler Server possa accedere al file. Nonostante questa soluzione potrebbe essere considerata inadatta agli ambienti di produzione, rappresenta una soluzione temporanea valida per un ambiente di prova al fine di valutare se il problema di autorizzazione è legato alla configurazione del sistema operativo.
- Definire un elenco di controllo dell'accesso (ACL) per il file */etc/shadow* (vedere l'esempio più avanti in questa sezione).
- Eseguire SPSS Modeler Server come root, per abilitare la lettura del file */etc/shadow* da parte dei processi server. *Attenzione:* In questo caso, accertarsi che il file *options.cfg* per SPSS Modeler Server contenga l'opzione *start_process_as_login_user, N* per evitare il problema di protezione descritto in precedenza.

Esempio di elenco di controllo dell'accesso (ACL)

Sui sistemi operativi Solaris si può usare il comando *setfacl* per creare un elenco di controllo dell'accesso (rivolgersi all'amministratore per sapere se questo comando è disponibile sul proprio sistema). L'esempio che segue illustra come creare un elenco di controllo dell'accesso di questo tipo per l'utente *modserv*:

```
# vi /tmp/shadow.acl
user::r--
user:modserv:r--
group:---
mask:r--
other:---

# setfacl -f /tmp/shadow.acl /etc/shadow

# getfacl /etc/shadow

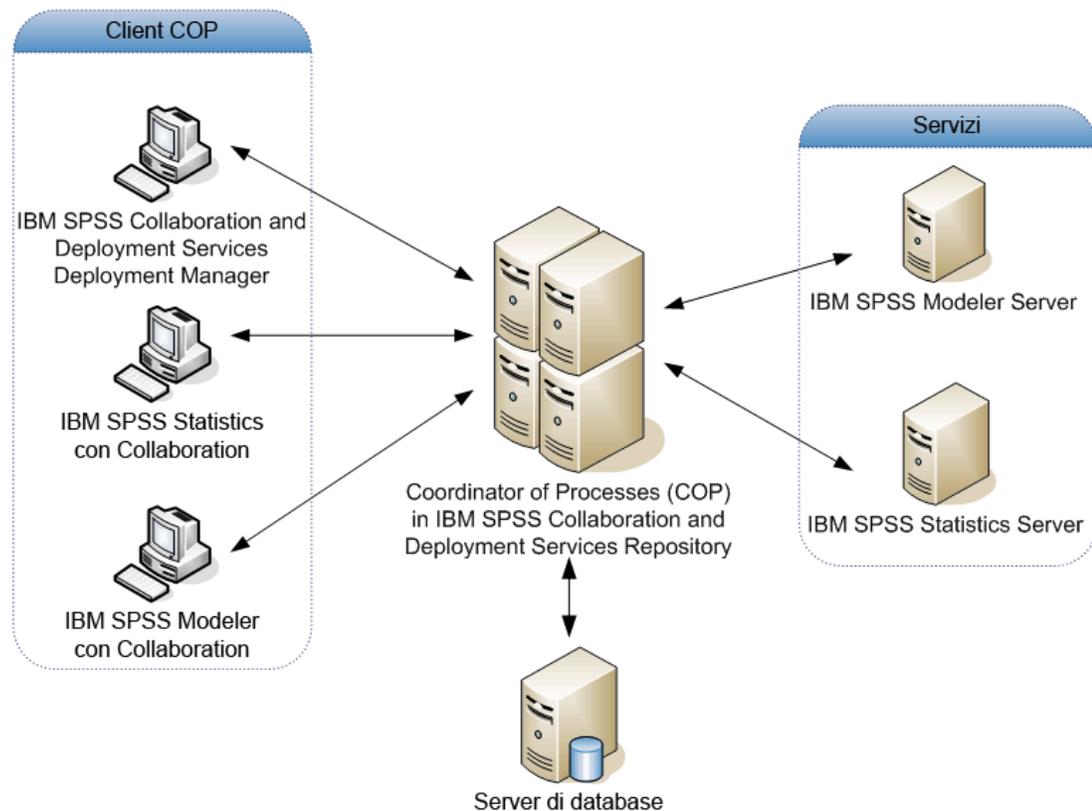
# file: /etc/shadow
# owner: root
# group: sys
user::r--
user:modserv:r--      #effective:r--
group:---             #effective:---
mask:r--
other:---
```

Bilanciamento del carico con cluster di server

Con IBM® SPSS® Collaboration and Deployment Services, per gestire i servizi della rete è possibile utilizzare un plug-in denominato Coordinator of Processes. Coordinator of Processes fornisce funzionalità di gestione server progettate per ottimizzare la comunicazione e l'elaborazione client-server.

I servizi da gestire, quali IBM® SPSS® Statistics Server o IBM® SPSS® Modeler Server, si registrano con Coordinator of Processes all'avvio e inviano periodicamente messaggi di stato aggiornati. I servizi possono inoltre archiviare eventuali file di configurazione necessari in IBM® SPSS® Collaboration and Deployment Services Repository e recuperarli al momento dell'inizializzazione.

Figura D-1
Architettura di Coordinator of Processes



L'esecuzione di stream IBM® SPSS® Modeler su un server può migliorare le prestazioni. Talvolta è possibile che la scelta sia limitata a uno o due server, mentre altre volte i server disponibili potrebbero essere molti poiché esiste una differenza sostanziale tra di loro, per esempio

in termini di proprietario, diritti di accesso, dati server, server di prova e server di produzione e così via. Inoltre, se si dispone di Coordinator of Processes nella propria rete, è possibile che venga offerto un cluster di server.

Un cluster di server è un gruppo di server intercambiabili in termini di configurazione e risorse. Coordinator of Processes stabilisce quello più adatto a rispondere a una richiesta di elaborazione utilizzando un algoritmo che bilancia il carico in base a numerosi criteri, inclusi i pesi dei server, le priorità utente e i carichi di elaborazione correnti. Per ulteriori informazioni, vedere il documento *Coordinator of Processes Service Developer's Guide*, disponibile su richiesta presso IBM Corp..

Ogni volta che ci si connette a un server o a un cluster di server in SPSS Modeler, è possibile immettere un server manualmente o cercare un server o un cluster con Coordinator of Processes. [Per ulteriori informazioni, vedere l'argomento Connessione a IBM SPSS Modeler Server in il capitolo 3 in *Manuale dell'utente di IBM SPSS Modeler 15*.](#)

Note

Queste informazioni sono state preparate per prodotti e servizi offerti in tutto il mondo.

IBM potrebbe non offrire i prodotti, i servizi o le funzionalità di cui si tratta nel presente documento in altri paesi. Contattare il rappresentante IBM locale per informazioni sui prodotti e i servizi attualmente disponibili nella propria zona. Qualsiasi riferimento a un prodotto, programma o servizio IBM non intende dichiarare o implicare che sia possibile utilizzare esclusivamente tale prodotto, programma o servizio IBM. Potrà invece essere utilizzato qualsiasi prodotto, programma o servizio con funzionalità equivalente e che non violi i diritti di proprietà intellettuale di IBM. Tuttavia, è responsabilità dell'utente valutare e verificare il funzionamento di qualsiasi prodotto, programma o servizio non IBM.

IBM può essere titolare di brevetti o domande di brevetto relativi alla materia oggetto del presente documento. La consegna del presente documento non conferisce alcuna licenza rispetto a questi brevetti. Rivolgere per iscritto i quesiti sulle licenze a:

IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785, U.S.A.

Per richieste di informazioni sulle licenze riguardanti il set di caratteri a byte doppio (DBCS), contattare l'Intellectual Property Department di IBM del proprio paese, oppure inviare le richieste in forma scritta all'indirizzo:

Intellectual Property Licensing, Legal and Intellectual Property Law, IBM Japan Ltd., 1623-14, Shimotsuruma, Yamato-shi, Kanagawa 242-8502 Giappone.

Il seguente paragrafo non si applica per il Regno Unito o altri paesi in cui le presenti disposizioni non sono conformi alle leggi locali: INTERNATIONAL BUSINESS MACHINES FORNISCE QUESTA PUBBLICAZIONE “COSÌ COM'È” SENZA GARANZIA DI ALCUN TIPO, SIA ESSA ESPRESSA O IMPLICITA, INCLUSE, MA NON LIMITATE A, LE GARANZIE IMPLICITE DI NON VIOLAZIONE, COMMERCIALIZZABILITÀ O IDONEITÀ A UNO SCOPO SPECIFICO. Alcuni stati non consentono limitazioni di garanzie espresse o implicite in determinate transazioni, pertanto quanto sopra potrebbe non essere applicabile.

Le presenti informazioni possono includere imprecisioni tecniche o errori tipografici. Le modifiche periodiche apportate alle informazioni contenute in questa pubblicazione verranno inserite nelle nuove edizioni della pubblicazione. IBM può apportare miglioramenti e/o modifiche al/ai prodotto/i e/o al/ai programma/i descritti nella presente pubblicazione in qualsiasi momento senza preavviso.

Qualsiasi riferimento nelle presenti informazioni a siti Web non IBM viene fornito esclusivamente per facilitare la consultazione e non rappresenta in alcun modo un'approvazione o sostegno da parte nostra di tali siti Web. I materiali contenuti in tali siti Web non fanno parte dei materiali di questo prodotto IBM e il loro utilizzo è esclusivamente a rischio dell'utente.

IBM può utilizzare o distribuire eventuali informazioni fornite dall'utente nei modi che ritiene appropriati senza incorrere in alcun obbligo nei confronti dell'utente.

I licenziatari del programma che desiderassero informazioni su di esso allo scopo di abilitare: (i) lo scambio di informazioni tra programmi creati indipendentemente e altri programmi (questo compreso) e (ii) l'utilizzo in comune delle informazioni scambiate, dovranno rivolgersi a:

IBM Software Group, All'attenzione di: Licensing, 233 S. Wacker Dr., Chicago, IL 60606, USA.

Tali informazioni saranno fornite in conformità ai termini e alle condizioni in vigore e, in alcuni casi, dietro pagamento.

Il programma concesso in licenza descritto nel presente documento e tutto il materiale correlato disponibile sono forniti da IBM in base ai termini del contratto di licenza cliente IBM, del contratto di licenza internazionale IBM o del contratto equivalente esistente tra le parti.

Tutti i dati sulle prestazioni qui contenuti sono stati elaborati in ambiente controllato. Di conseguenza, i risultati ottenuti con sistemi operativi diversi possono variare in modo significativo. Alcune misurazioni potrebbero essere state effettuate su sistemi in corso di sviluppo e non c'è garanzia che tali misurazioni coincidano con quelle effettuate sui sistemi comunemente disponibili. Inoltre, alcune misurazioni potrebbero essere stime elaborate tramite l'estrapolazione. I risultati effettivi potrebbero variare. Gli utenti di questo documento devono verificare i dati relativi al proprio ambiente specifico.

Le informazioni relative a prodotti non IBM sono state ottenute dai fornitori di tali prodotti, da loro annunci pubblicati e da altre fonti disponibili al pubblico. IBM non ha verificato tali prodotti e non può confermare l'accuratezza delle prestazioni, la compatibilità o qualsiasi altra dichiarazione relativa a prodotti non IBM. Eventuali domande in merito alle funzionalità dei prodotti non IBM vanno indirizzate ai fornitori di tali prodotti.

Qualsiasi affermazione relativa agli obiettivi e alla direzione futura di IBM è soggetta a modifica o revoca senza preavviso e concerne esclusivamente gli scopi dell'azienda.

Le presenti informazioni includono esempi di dati e report utilizzati in operazioni aziendali quotidiane. Per fornire una descrizione il più possibile esaustiva, gli esempi includono nomi di persone, società, marchi e prodotti. Tutti questi nomi sono fittizi e ogni somiglianza a nomi e indirizzi utilizzati da aziende reali è puramente casuale.

Per chi visualizza queste informazioni a video: le fotografie e le illustrazioni a colori potrebbero non essere disponibili.

Marchi

IBM, il logo IBM, ibm.com e SPSS sono marchi di IBM Corporation, registrati in numerose giurisdizioni nel mondo. Un elenco aggiornato dei marchi IBM è disponibile sul Web all'indirizzo <http://www.ibm.com/legal/copytrade.shtml>.

Intel, il logo Intel, Intel Inside, il logo Intel Inside, Intel Centrino, il logo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium e Pentium sono marchi o marchi registrati di Intel Corporation o delle sue consociate negli Stati Uniti e in altri paesi.

Linux è un marchio registrato di Linus Torvalds negli Stati Uniti e/o negli altri paesi.

Microsoft, Windows, Windows NT e il logo Windows sono marchi di Microsoft Corporation negli Stati Uniti e/o negli altri paesi.

UNIX è un marchio registrato di The Open Group negli Stati Uniti e in altri paesi.

Java e tutti i marchi e i logo basati su Java sono marchi di Sun Microsystems, Inc. negli Stati Uniti e/o negli altri paesi.

Altri nomi di prodotti e servizi possono essere marchi commerciali di IBM o di altre aziende.



- accesso ai dati, 13
- accesso come amministratore
 - per IBM SPSS Modeler Server, 27
- account utente
 - IBM SPSS Modeler Server, 21
- aggiunta di connessioni di IBM SPSS Modeler Server, 18–19
- allow_modelling_memory_override
 - file *options.cfg*, 31
- amministratore
 - della posizione della licenza di IBM SPSS Statistics, 38
 - di IBM SPSS Modeler Server, 25
- anteprima
 - Generazione SQL, 56
- architettura
 - lato server, 8
- arresto di IBM SPSS Modeler Server
 - in UNIX, 23
 - in Windows, 23
- autenticazione, 21
- autenticazione degli utenti, 21
- auto.sh* (UNIX)
 - posizione di, 69
- autorizzazioni per i file
 - configurazione in UNIX, 71
 - in IBM SPSS Modeler Server, 21
- avvio automatico del server
 - configurazione in UNIX, 70
- avvio di IBM SPSS Modeler Server
 - in UNIX, 23
 - in Windows, 23

- cache per nodo
 - Generazione SQL, 59
 - scrittura sul database, 59
- cache_compression
 - file *options.cfg*, 31
- caching del database
 - controllo da *options.cfg*, 36
 - Generazione SQL, 59
- caching in-database, 36
- caching, in-database, 36
- chemsrv.sh* (UNIX)
 - posizione di, 69
- compressione della cache, 31
- configurazione di Coordinator of Processes
 - per IBM SPSS Modeler Server, 34
- configurazione di COP
 - per IBM SPSS Modeler Server, 34
- connessioni
 - a IBM SPSS Modeler Server, 16, 18–19
 - cluster di server, 19
- connessioni database
 - chiusura, 37
- Coordinator of Processes, 19
 - bilanciamento del carico, 84
 - cluster di server, 84
- COP, 19
 - bilanciamento del carico, 84
 - cluster di server, 84
- cop_enabled
 - file *options.cfg*, 34
- cop_host
 - file *options.cfg*, 34
- cop_password
 - file *options.cfg*, 34
- cop_port_number
 - file *options.cfg*, 34
- cop_service_description
 - file *options.cfg*, 34
- cop_service_host
 - file *options.cfg*, 34
- cop_service_name
 - file *options.cfg*, 34
- cop_service_weight
 - file *options.cfg*, 34
- cop_update_interval
 - file *options.cfg*, 34
- cop_user_name
 - file *options.cfg*, 34
- crittografia
 - SSL, 38
- crittografia dei dati con SSL
 - attivazione per IBM SPSS Modeler Server, 33

- data_file_path
 - file *options.cfg*, 30
- data_files_restricted
 - file *options.cfg*, 30
- database
 - accesso, 13
 - livelli di supporto, 13
- DB2
 - ottimizzazione SQL, 52–53
- directory temp
 - per IBM SPSS Modeler Server, 30
- documentazione, 4

- elaborazione parallela
 - attivazione, 47
 - controllo, 31
- errore durante l'esecuzione di stream, 32
- esecuzione di più stream, 32
- esempi
 - cenni generali, 6
 - Guida alle applicazioni, 4
 - esempi di applicazioni, 4
 - espressioni CLEM
 - Generazione SQL, 63

- file di dati
 - IBM SPSS Statistics, 15
 - importazione ed esportazione, 15
- file di registro
 - per IBM SPSS Modeler Server, 29, 44
 - visualizzazione SQL generato, 56
- file *messages.log*, 44
- file *options.cfg*, 36
- file temporanei, 11
 - autorizzazioni per (IBM SPSS Modeler Server), 21
- funzione @SQLFN, 65

- Generazione SQL, 47, 52–53, 59
 - anteprima, 56
 - attivazione, 55
 - attivazione per IBM SPSS Modeler Server, 33
 - espressioni CLEM, 58, 63
 - registrazione, 56
 - risrittura degli stream, 58
 - suggerimenti, 58
 - visualizzazione per gli insiemi di modelli, 58
- gestione della memoria
 - opzioni di amministrazione, 31

- IBM SPSS Modeler, 1
 - documentazione, 4
- IBM SPSS Modeler Administration Console, 26
 - accesso come amministratore, 27
- IBM SPSS Modeler Server
 - accesso come amministratore, 27
 - account utente, 21
 - amministrazione di, 26
 - autenticazione degli utenti, 21
 - configurazione di Coordinator of Processes, 34
 - configurazione di COP, 34
 - creazione di file, 21
 - directory temp, 30
 - file di registro, 29
 - ID utente, 16
 - informazioni per gli utenti finali, 20
 - monitoraggio dell'utilizzo, 36
 - nome dominio (Windows), 16
 - nome host, 16, 18
 - numero porta, 16, 18, 29
 - opzioni di amministrazione, 25
 - opzioni di configurazione, 29
 - password, 16
 - processi bloccati, 24
 - processi server, 36
 - risultati diversi rispetto al client, 21
- IBM SPSS Modeler Server, accesso, 16
- IBM SPSS Statistics file di dati
 - importazione ed esportazione, 15
- ID utente
 - IBM SPSS Modeler Server, 16
- impostazioni del firewall
 - file *options.cfg*, 32
- impostazioni della porta
 - file *options.cfg*, 31
- insiemi di modelli
 - visualizzazione SQL per, 58
- io_buffer_size
 - file *options.cfg*, 31

- LDAP, 44
 - protezione, 44
- limiti del kernel in UNIX, 24
- limiti del kernel UNIX, 24
- livelli, supporto database, 13
- log_database_access
 - file *options.cfg*, 29
- log_database_errors
 - file *options.cfg*, 29
- log_directory
 - file *options.cfg*, 29
- log_file_access
 - file *options.cfg*, 29
- log_file_errors
 - file *options.cfg*, 29

- marchi, 87
- max_file_size
 - file *options.cfg*, 30
- max_login_attempts
 - file *options.cfg*, 29
- max_parallelism
 - file *options.cfg*, 31
- max_sessions
 - file *options.cfg*, 29
- max_sql_string_length
 - file *options.cfg*, 33
- memoria, 12
- memory_usage
 - file *options.cfg*, 31
- messaggi
 - visualizzazione SQL generato, 56
- Microsoft SQL Server
 - ottimizzazione SQL, 52–53
- mining in-database, 50
- modellazione
 - gestione della memoria, 31
- modelling_memory_limit_percentage
 - file *options.cfg*, 31

- nodi
 - con supporto per generazione SQL, 59
- nodi porpora, 47
 - ottimizzazione SQL, 54
- nome dominio (Windows)
 - IBM SPSS Modeler Server, 16
- nome host
 - IBM SPSS Modeler Server, 16, 18

- nomi di file
 - UNIX, 14
 - Windows, 14
- nomi di file UNC, 14
- note legali, 86
- numero porta
 - IBM SPSS Modeler Server, 16, 18, 29
- ODBC
 - configurazione in UNIX, 71
- Operatori
 - Generazione SQL, 63
- opzione `cache_connection`, 37
- opzioni di configurazione
 - accesso a file di dati, 30
 - cenni generali, 29
 - connessioni e sessioni, 29
 - coordinator of processes, 34
 - COP, 34
 - crittografia dei dati con SSL, 33
 - di IBM SPSS Modeler Server, 25
 - directory temp, 30
 - elaborazione parallela, 31
 - generazione SQL automatica, 33
 - gestione della memoria, 31
 - lunghezza delle stringhe SQL, 33
 - numero porta, 29
 - performance e ottimizzazione, 31
 - registrazione, 29
 - risrittura degli stream, 31
 - tentativi di accesso, 29
- Oracle
 - ottimizzazione SQL, 52–53, 67
- ottimizzazione, 47
 - Generazione SQL, 52–53, 55
- pacchetto di accesso ai dati
 - configurazione di UNIX per, 71
 - e percorsi delle librerie in UNIX, 77
 - ODBC, configurazione in UNIX, 71
 - risoluzione dei problemi di ODBC in UNIX, 74
- password
 - IBM SPSS Modeler Server, 16
- percorsi, 14
- performance
 - di IBM SPSS Modeler Server, 47
- `port_number`
 - file `options.cfg`, 29
- posizione della licenza IBM SPSS Statistics, 38
- Prefisso URL, 43
- processi zombie, IBM SPSS Modeler Server, 24
- processi, bloccati, 24
- processori, 10
 - multipli, 31
- `program_file_path`
 - file `options.cfg`, 30
- `program_files_restricted`
 - file `options.cfg`, 30
- protezione
 - LDAP, 44
- push back, 47, 52–53, 55, 59
 - espressioni CLEM, 63
- push back SQL. *Vedere anche* Generazione SQL, 52
- RAM, 12
- `rc.modeler` (UNIX)
 - posizione di, 69
- ricerca di connessioni in COP, 19
- risrittura degli stream, 58
 - attivazione, 47
- risultati
 - arrotondamento di, 22
 - differenze tra Client e Server, 21
 - ordine dei record, 21
- script UNIX
 - `auto.sh`, 69
 - `modelersrv.sh`, 69
 - modifica, 71
 - `rc.modeler`, 69
- Secure Sockets Layer, 38
- server
 - accesso a, 16
 - aggiunta di connessioni, 18
 - ricerca di server in COP, 19
- server del database, 50
- server, impostazioni della porta
 - file `options.cfg`, 31
- sicurezza
 - configurazione della creazione di file in UNIX, 71
 - creazione di file, 21
 - SSL, 38
- single sign-on, 18
- sistemi operativi
 - a 64-bit, 11
- sistemi operativi a 64-bit, 11
- sorgenti dati ODBC
 - e UNIX, 71
 - ODBC e script UNIX, 71
- spazio su disco, 11
 - calcolo, 12
- SPSS Modeler Server, 2
- SQL
 - anteprima generata, 56
 - nomi di colonna duplicati, 66
 - ottimizzazione di Oracle, 67
 - query, 66
 - visualizzazione per gli insiemi di modelli, 58
- SQL Server
 - ottimizzazione SQL, 52–53
- `sql_generation_enabled`
 - file `options.cfg`, 33

SSL, 38

cenni generali, 38

protezione delle comunicazioni, 39

ssl_certificate_file

file *options.cfg*, 33

ssl_enabled

file *options.cfg*, 33

ssl_private_key_file

file *options.cfg*, 33

ssl_private_key_password

file *options.cfg*, 33

stato

di IBM SPSS Modeler Server in UNIX, 23

di IBM SPSS Modeler Server in Windows, 23

stream_rewriting_enabled

file *options.cfg*, 31

suggerimenti relativi all'hardware

per IBM SPSS Modeler Server, 10

Tecnologia IBM SPSS Statistics Data Access, 13

temp_directory

file *options.cfg*, 30

unità disco rigido, 12

UNIX

autenticazione degli utenti, 21

configurazione delle autorizzazioni per i file, 71

percorsi delle librerie, 77