

*IBM SPSS Modeler Server 16 Guida  
della performance e amministrazione*

**IBM**

**Nota**

Prima di utilizzare queste informazioni e il relativo prodotto, leggere le informazioni in "Note" a pagina 71.

**Informazioni sul prodotto**

La presente edizione si applica alla versione 16, release 0, modifica 0 di IBM(r) SPSS(r) Modeler e a tutte le release e le modifiche successive se non diversamente indicato nelle nuove edizioni.

# Indice

<b>Prefazione</b> . . . . .	<b>v</b>
-----------------------------	----------

## **Capitolo 1. Informazioni su IBM SPSS**

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

## **Capitolo 2. Architettura e suggerimenti relativi all'hardware** . . . . .

<b>relativi all'hardware</b> . . . . .	<b>7</b>
Architettura di IBM SPSS Modeler . . . . .	7
Descrizione dell'architettura . . . . .	7
Suggerimenti relativi all'hardware . . . . .	9
Spazio temporaneo su disco e RAM . . . . .	10
Accesso ai dati . . . . .	11
Riferimento a file di dati . . . . .	12
Importazione di file di dati di IBM SPSS Statistics . . . . .	12
Istruzioni di installazione . . . . .	12

## **Capitolo 3. Supporto per IBM SPSS**

<b>Modeler</b> . . . . .	<b>15</b>
Connessione a IBM SPSS Modeler Server . . . . .	15
Configurazione SSO (Single Sign-On) . . . . .	16
Aggiunta e modifica della connessione di IBM SPSS Modeler Server . . . . .	17
Ricerca di server in IBM SPSS Collaboration and Deployment Services . . . . .	18
Dati e file system . . . . .	18
Autenticazione degli utenti . . . . .	19
Creazione di file . . . . .	19
Differenze nei risultati . . . . .	19

## **Capitolo 4. Amministrazione di IBM SPSS Modeler Server** . . . . .

<b>SPSS Modeler Server</b> . . . . .	<b>21</b>
Avvio e arresto di IBM SPSS Modeler Server . . . . .	21
Per avviare, arrestare e verificare lo stato di Clementine Server in Windows . . . . .	21
Per avviare, arrestare e verificare lo stato di Clementine Server in UNIX . . . . .	21
Gestione dei processi di server bloccati (sistemi UNIX) . . . . .	22
Amministrazione . . . . .	22

Utilizzo di IBM SPSS Modeler Administration Console . . . . .	23
Avvio di Modeler Administration Console . . . . .	23
Configurazione dell'accesso tramite Modeler Administration Console . . . . .	23
Connessioni a SPSS Modeler Server . . . . .	24
Configurazione di SPSS Modeler Server . . . . .	24
Monitoraggio di SPSS Modeler Server . . . . .	29
Utilizzo del file options.cfg . . . . .	29
Chiusura delle connessioni di database inutilizzate . . . . .	30
Utilizzo di SSL per il trasferimento sicuro dei dati . . . . .	30
Come funziona il protocollo SSL . . . . .	30
Protezione delle comunicazioni client-server e server-server con SSL . . . . .	31

## **Capitolo 5. Panoramica delle performance** . . . . .

<b>performance</b> . . . . .	<b>35</b>
Impostazioni relative alle performance e all'ottimizzazione del server . . . . .	35
Impostazioni relative alle performance e all'ottimizzazione del client . . . . .	35
Ottimizzazione e utilizzo del database . . . . .	37
Ottimizzazione SQL . . . . .	37

## **Capitolo 6. Ottimizzazione SQL** . . . . .

<b>Ottimizzazione SQL</b> . . . . .	<b>39</b>
Funzionamento della generazione SQL . . . . .	40
Esempio di generazione SQL . . . . .	41
Configurazione dell'ottimizzazione SQL . . . . .	42
Anteprima di SQL generato . . . . .	43
Visualizzazione di SQL per i nugget del modello . . . . .	43
Suggerimenti per l'ottimizzazione della generazione SQL . . . . .	43
Nodo con supporto per generazione SQL . . . . .	44
Espressioni e operatori CLEM con supporto di generazione SQL . . . . .	48
Utilizzo di funzioni SQL in espressioni CLEM . . . . .	51
Scrittura di query SQL . . . . .	51

## **Appendice A. Configurazione di Oracle per piattaforme UNIX** . . . . .

<b>per piattaforme UNIX</b> . . . . .	<b>53</b>
Configurazione di Oracle per l'ottimizzazione SQL . . . . .	53

## **Appendice B. Configurazione degli script di avvio di UNIX** . . . . .

<b>script di avvio di UNIX</b> . . . . .	<b>55</b>
Introduzione . . . . .	55
Script . . . . .	55
Avvio e arresto automatico di IBM SPSS Modeler Server . . . . .	55
Avvio e arresto manuale di IBM SPSS Modeler Server . . . . .	56
Modifica degli script . . . . .	56
Verifica delle autorizzazioni per la creazione di file IBM SPSS Modeler Server e Data Access Pack . . . . .	56

Risoluzione dei problemi di una configurazione ODBC . . . . .	58
Percorsi delle librerie . . . . .	61
Configurazione di un driver per il nodo Enterprise View . . . . .	61

**Appendice C. Configurazione ed esecuzione di SPSS Modeler Server come processo non root in UNIX . . . . 63**

Introduzione . . . . .	63
Configurazione non root senza un database privato di password . . . . .	63
Configurazione non root con un database di password privato . . . . .	63
Esecuzione di SPSS Modeler Server come utente non root . . . . .	65
Risoluzione dei problemi di mancata autenticazione degli utenti . . . . .	65

**Appendice D. Configurazione ed esecuzione di SPSS Modeler Server utilizzando un file di password privato su Windows . . . . . 67**

Introduzione . . . . .	67
Configurazione utilizzando un database di password privato . . . . .	67

**Appendice E. Bilanciamento del carico con cluster di server. . . . . 69**

<b>Note . . . . .</b>	<b>71</b>
Marchi . . . . .	72

**Indice analitico . . . . . 75**

---

## 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 visuale di SPSS Modeler favorisce l'applicazione di una competenza di business 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 di business. 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 di business. Le aziende, gli enti governativi e le università di tutto il mondo si affidano alla tecnologia IBM SPSS ottenendo un vantaggio competitivo in termini di attrazione, mantenimento e ampliamento dei clienti, riducendo al contempo frodi e 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 di business 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 accordo di manutenzione.



---

## Capitolo 1. 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 di business e di eseguirne la distribuzione nelle operazioni di business per migliorare i processi decisionali. Progettato secondo il modello CRISP-DM conforme agli standard di settore, IBM SPSS Modeler supporta l'intero processo di data mining, dai dati a risultati di business migliori.

IBM 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, consultare <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
- Adattatori IBM SPSS Modeler Server 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 tendenze, 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, consultare l'argomento "Edizioni di IBM SPSS Modeler" a pagina 2.

### 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 nel 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 attività 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 un flusso 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 i flussi pubblicati.

## **Adattatori IBM SPSS Modeler Server 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, un flusso 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 supplementare 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.

**IBM SPSS Modeler Text Analytics** utilizza tecnologie linguistiche avanzate e di NLP (Natural Language Processing) per elaborare rapidamente una grande varietà di dati di testo non strutturati, estrarre ed organizzare i concetti chiave e raggruppare tali 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 IBM 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=swg27038316>.

La documentazione in entrambi i formati è inoltre disponibile presso il Centro informazioni SPSS Modeler all'indirizzo <http://publib.boulder.ibm.com/infocenter/spssmodl/v16r0m0/>.

## Documentazione di SPSS Modeler Professional

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

- **IBM SPSS Modeler - Guida per l'utente.** Introduzione generale all'utilizzo di SPSS Modeler che illustra come creare flussi di dati, gestire valori mancanti, generare espressioni CLEM, utilizzare progetti e report e assemblare flussi per la distribuzione tramite IBM SPSS Collaboration and Deployment Services, le applicazioni predittive o IBM SPSS Modeler Advantage.
- **IBM SPSS Modeler - Nodi origine, elaborazione e output.** 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 di modellazione.** Descrizioni di tutti i nodi utilizzati per creare modelli data mining. IBM SPSS Modeler offre numerosi metodi di modellazione ricavati dall'apprendimento automatico, dall'intelligenza artificiale e dalla statistica.
- **IBM SPSS Modeler Algorithms Guide.** Descrizione dei fondamenti di matematica per i metodi di modellazione utilizzati in IBM SPSS Modeler. Questa guida è disponibile solo in formato PDF.
- **IBM SPSS Modeler Applications Guide.** 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, consultare l'argomento "Esempi di applicazioni" a pagina 5.
- **IBM SPSS Modeler - Guida per script e automazione.** Informazioni sulle modalità di automazione del sistema tramite script, incluse le proprietà che è possibile utilizzare per manipolare nodi e flussi.
- **IBM SPSS Modeler - Guida alla distribuzione.** Informazioni sull'esecuzione di flussi e scenari IBM 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 come routine di elaborazione dei dati o algoritmi di modellazione come nodi in IBM SPSS Modeler.
- **IBM SPSS Modeler - Guida al mining nel 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.
- **IBM SPSS Modeler Administration Console - Guida per l'utente.** Informazioni sull'installazione e l'utilizzo dell'interfaccia utente della console per il monitoraggio e la configurazione di IBM SPSS Modeler Server. La console viene implementata come plug-in dell'applicazione Deployment Manager.
- **IBM SPSS Modeler - Guida CRISP-DM.** Guida passo a passo al data mining tramite la metodologia CRISP-DM con SPSS Modeler.
- **IBM SPSS Modeler Batch - Guida per l'utente.** 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.

- **IBM SPSS Modeler Entity Analytics User Guide.** 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.
- **IBM SPSS Modeler Social Network Analysis User Guide.** Guida che spiega come eseguire l'analisi dei social network con SPSS Modeler; comprende l'analisi di gruppo e l'analisi di diffusione.
- **SPSS Modeler Text Analytics - Guida per l'utente.** 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.
- **IBM SPSS Modeler Text Analytics Administration Console - Guida per l'utente.** 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 SPSS Modeler Text Analytics. 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 di business 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 i flussi di esempio sono installati nella cartella *Demos* nella directory di installazione del prodotto. Per ulteriori informazioni, consultare l'argomento "Cartella Demos".

**Esempi di modellazione del database.** Vedere gli esempi nella *IBM SPSS Modeler Guida al mining nel database*.

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

---

## Cartella Demos

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



---

## Capitolo 2. 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 IBM 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 IBM SPSS Modeler Server (rispetto a IBM SPSS Modeler come applicazione autonoma) sono numerosi, soprattutto quando occorre gestire insiemi di dati voluminosi:

- Oltre che con Windows, IBM SPSS Modeler Server può essere eseguito su sistemi UNIX, consentendo maggiore flessibilità nella scelta di dove installarlo. Su qualsiasi piattaforma è possibile destinare un computer server più veloce e potente ai processi di data mining.
- IBM SPSS Modeler Server è ottimizzato per garantire prestazioni veloci. Quando non è possibile inviare le operazioni al database, IBM 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, IBM 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 di business.
- Con l'utilizzo di strumenti di amministrazione quali IBM SPSS Modeler Administration Console (fornito con IBM 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 illustrati nel grafico "IBM SPSS Modeler Server Architecture".

- **IBM 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 IBM SPSS Modeler il cui motore di esecuzione è però inattivo quando questo è connesso a IBM SPSS Modeler Server per l'analisi distribuita. IBM SPSS Modeler funziona solo su sistemi operativi Windows.
- **IBM SPSS Modeler Server.** Il software del server installato su un computer server, con connettività di rete sia su IBM SPSS Modeler(s) sia sul database. IBM SPSS Modeler Server viene eseguito come servizio (in Windows) o come processo daemon (in UNIX), ess (on UNIX), in attesa delle connessioni dei client. L'applicazione gestisce l'esecuzione dei flussi e degli script creati mediante IBM SPSS Modeler.
- **Server di database.** Il server di database può essere un data warehouse in tempo reale (per esempio Oracle su un server UNIX di grandi dimensioni) o, per ridurre l'impatto su altri sistemi operativi, un data mart su un server locale/di reparto (per esempio, SQL Server in Windows).

## Architettura IBM SPSS Modeler Server

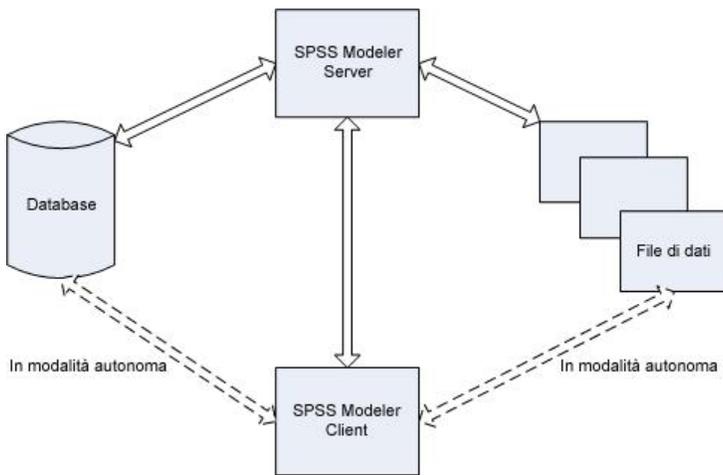


Figura 1. Architettura IBM SPSS Modeler Server

Con l'architettura distribuita, gran parte dell'elaborazione avviene sul server. Quando l'utente finale esegue un flusso, IBM SPSS Modeler invia al server una descrizione del flusso 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, IBM SPSS Modeler Server può eseguire tutte le operazioni di IBM 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, IBM SPSS Modeler Server risulta totalmente compatibile con i file flat.

È inoltre disponibile il bilanciamento del carico utilizzando un cluster di server per l'elaborazione. Il clustering è disponibile a partire da IBM SPSS Collaboration and Deployment Services 3.5 tramite il plug-in Coordinator of Processes. Per ulteriori informazioni, consultare l'argomento Appendice E, "Bilanciamento del carico con cluster di server", a pagina 69. È possibile connettersi a un server o a un cluster gestiti in Coordinator of Processes direttamente tramite la finestra di dialogo Login server di IBM SPSS Modeler. Per ulteriori informazioni, consultare l'argomento "Connessione a IBM SPSS Modeler Server" a pagina 15.

### Client standalone

IBM SPSS Modeler può essere configurato anche per essere eseguito come applicazione desktop autonoma, come illustrato nel seguente grafico. Per ulteriori informazioni, consultare Capitolo 3, "Supporto per IBM SPSS Modeler", a pagina 15.

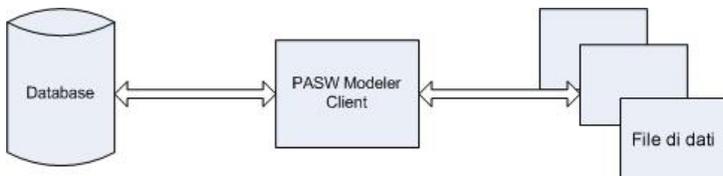


Figura 2. IBM SPSS Modeler standalone

---

## Suggerimenti relativi all'hardware

Nel pianificare l'installazione di IBM SPSS Modeler Server è opportuno prendere in considerazione l'hardware che sarà utilizzato. Benché IBM 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 IBM 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 IBM SPSS Modeler Server effettua la connessione. In particolare, le operazioni di creazione del modello assorbono molte risorse e la loro esecuzione è molto più veloce quando non devono competere con altre applicazioni.

*Note:* sebbene l'installazione di IBM 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 IBM SPSS Modeler Server può accettare connessioni da più client (utenti), e ogni connessione client può avviare l'esecuzione di diversi flussi. Pertanto, in un server possono essere in esecuzione diverse attività 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 del modello 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 IBM 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 IBM 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 flat 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 i flussi 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 Raccolta
- 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.
- Qualsiasi funzione @, per esempio @MIN, @MAX e @AVE, in cui viene calcolato il parametro dell'offset.

### 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, IBM 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 la possibilità di utilizzare un file system distinto per i file temporanei di IBM SPSS Modeler, creati 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 Modelli viene utilizzata la RAM. I nodi Rete neurale, Kohonen e Medie K

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  
(numero\_di\_record \* numero\_di\_celle\_per\_record) \* numero\_di\_byte\_per\_cella

dove numero\_di\_celle\_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 include una serie di driver ODBC che è possibile utilizzare per questo scopo, disponibili nel DVD IBM SPSS Data Access Pack oppure dal sito di download. Per domande su come creare o impostare le autorizzazioni per le origini dati ODBC, contattare l'amministratore di database.

### Driver ODBC supportati

Per le ultime informazioni sui database e i driver ODBC che sono supportati e testati per l'uso con IBM SPSS Modeler 16, 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 IBM 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 IBM SPSS Modeler Server. Per IBM 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 IBM SPSS Modeler che da IBM SPSS Modeler Server, i driver ODBC devono essere installati su entrambi i computer.
- Se IBM SPSS Modeler è in esecuzione su Terminal Services, i driver ODBC devono essere installati sul server Terminal Services in cui è stato installato IBM SPSS Modeler.

**Importante:** se si sta utilizzando IBM SPSS Modeler Server su UNIX per accedere ad un database Teradata e il driver ODBC fornito da Teradata, è necessario utilizzare ODBC Driver Manager installato con tale driver. (**Nota:** non è necessario apportare queste modifiche se si utilizza il driver SDAP Teradata). Per poter apportare questa modifica a IBM 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). È necessario riavviare IBM SPSS Modeler Server per applicare la modifica. Per i dettagli relativi alle piattaforme IBM 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 IBM SPSS Modeler Server nei sistemi UNIX. Per configurare UNIX per il caricamento del gestore dei driver DataDirect, immettere i comandi seguenti:

```
cd modeler_server_install_directory/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 IBM SPSS Modeler imposta la directory di lavoro per gli oggetti *locali* dei client (per esempio i flussi), ma non modifica la directory di lavoro del server.

**UNIX.** Per informazioni su come configurare IBM SPSS Modeler Server su sistemi UNIX per il funzionamento con la tecnologia IBM SPSS Statistics Data Access, vedere Appendice B, "Configurazione degli script di avvio di UNIX", a pagina 55.

## 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, *\\mydataserver\ServerData\Sales 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 IBM SPSS Modeler Server. Il fatto che un file di testo utilizzi il formato UNIX o DOS non è importante, 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 IBM 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 IBM 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 IBM SPSS Modeler IBM SPSS Statistics possono avviare automaticamente il client IBM SPSS Statistics ma gli utenti devono prima verificare che il client IBM SPSS Statistics in esecuzione nella stessa modalità di IBM SPSS Modeler.

---

## Istruzioni di installazione

Per informazioni sull'installazione di IBM SPSS Modeler Server, consultare le istruzioni contenute nella cartella *\documentation\installation\* sul CD 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.



---

## Capitolo 3. 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 IBM SPSS Modeler Server
- Differenze nei risultati riscontrati dagli utenti nel passaggio da IBM 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 IBM 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 IBM 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 connettersi a un server

1. 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 IBM SPSS Modeler.
2. 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. Consultare l'argomento "Aggiunta e modifica della connessione di IBM SPSS Modeler Server" a pagina 17 per ulteriori informazioni.
  - Fare clic su **Cerca** per accedere al server o al cluster di server in Coordinator of Processes. Consultare l'argomento "Ricerca di server in IBM SPSS Collaboration and Deployment Services" a pagina 18 per ulteriori informazioni.

**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 (SSO)**, 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 (SSO) non è disponibile oppure se si seleziona la casella per disabilitare Single Sign-On (SSO) (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.

3. Fare clic su **OK** per completare la connessione.

Per disconnettersi da un server

1. 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 IBM SPSS Modeler.
2. Nella finestra di dialogo, selezionare il Server locale e fare clic su **OK**.

## Configurazione SSO (Single Sign-On)

È possibile connettersi a un IBM SPSS Modeler Server in esecuzione su una qualsiasi piattaforma che utilizza l'SSO (Single Sign-On). Per connettersi utilizzando l'SSO (Single Sign-On) è necessario prima configurare il proprio server IBM SPSS Modeler e le macchine client.

Se si utilizza un SSO (Single Sign-On) per connettersi a IBM SPSS Modeler Server e IBM SPSS Collaboration and Deployment Services, è necessario connettersi a IBM SPSS Collaboration and Deployment Services prima di connettersi a IBM SPSS Modeler.

**Nota:** Prima di configurare il proprio server IBM SPSS Modeler e le macchine client per l'SSO (Single Sign-On), è necessario essere certi che le macchine abbiano accesso al server controller del dominio.

## Configurazione di SSO (Single Sign-On) su Windows

Per configurare IBM SPSS Modeler Server

1. Assicurarsi che la macchina del server Windows sia un membro del dominio AD (Active Directory).
2. Nell'ubicazione di installazione IBM SPSS Modeler Server, individuare la cartella denominata config.
3. Nella cartella config, creare una cartella secondaria denominata sso.
4. Nella cartella sso, creare un file krb5.conf. Le istruzioni su come creare il file krb5.conf si trovano all'indirizzo [http://web.mit.edu/kerberos/krb5-current/doc/admin/conf\\_files/krb5\\_conf.html](http://web.mit.edu/kerberos/krb5-current/doc/admin/conf_files/krb5_conf.html). Di seguito è riportato un esempio di un file krb5.conf:

```
[libdefaults]
    default_realm = MODELERSSO.COM
    dns_lookup_kdc = true
    dns_lookup_realm = false

[realms]
    MODELERSSO.COM = {
        kdc = <AD server IP>:88
        admin_server = <AD server IP>:749
        default_domain = MODELERSSO.COM
    }

[domain_realm]
    .modelersso.com = MODELERSSO.COM
```

Per configurare IBM SPSS Modeler Client

1. Assicurarsi che la macchina Windows locale su cui viene eseguito IBM SPSS Modeler sia un membro del dominio AD.
2. Aggiungere l'utente del dominio come amministratore sulla macchina locale.
3. Abilitare Windows al fine di accedere alla chiave di sessione TGT:
  - a. Dal menu di **Avvio**, fare clic su **Esegui**.
  - b. Immettere `regedit` e fare clic su **OK** per aprire Editor di registro.
  - c. Passare all'ubicazione di registro appropriata per il sistema operativo della macchina locale:

- Su Windows XP: My Computer\HKEY\_LOCAL\_MACHINE\System\CurrentControlSet\Control\Lsa\Kerberos
  - Su Windows Vista o Windows 7: My Computer\HKEY\_LOCAL\_MACHINE\System\CurrentControlSet\Control\Lsa\Kerberos\Parameters
- d. Fare clic con il tasto destro del mouse sulla cartella e selezionare **New > DWORD**. Il nome del nuovo valore deve essere `allowtgtsessionkey`.
  - e. Impostare il valore `allowtgtsessionkey` sul valore esadecimale 1, ovvero `0x0000001`.
  - f. Chiudere l'Editor di registro.
  - g. Se il sistema operativo del computer locale è Windows Vista o Windows 7, è necessario eseguire `kinit.exe`, che è possibile trovare nella posizione di installazione `<IBM SPSS Modeler >\jre\bin`.
4. Nella cartella `config` dell'ubicazione di installazione IBM SPSS Modeler, creare una cartella denominata `sso`.
  5. Copiare il file `krb5.conf` dal server nella cartella `sso`.
  6. Riavviare la macchina locale e la macchina del server.

### Configurazione di SSO (Single Sign-On) su UNIX

Per configurare l'SSO (Single Sign-On) per le macchine server UNIX, è possibile aggiungere la macchina UNIX al dominio AD di Windows e poi seguire le istruzioni per la configurazione dell'SSO (Single Sign-On) su Windows. Per ulteriori informazioni, consultare l'argomento "Configurazione di SSO (Single Sign-On) su Windows" a pagina 16. In alternativa, è possibile eseguire le seguenti operazioni:

1. Creare un account utente del dominio per la macchina UNIX.
2. Modificare il nome host. Se si sta utilizzando RedHat Linux, aprire il file `/etc/sysconfig/network` e modificare `HOSTNAME` nel formato `<nome>.<realm>`. Ciò abilita l'AD per individuare le credenziali del server.
3. Per abilitare il server DNS al fine di individuare la macchina UNIX, effettuare una delle seguenti operazioni:
  - Aprire il file `%windows%/system32/drivers/etc/hosts` e aggiungere il mapping IP/host, ad esempio:
 

```
192.168.1.102 test.modelersso.com test
```

 Oppure
  - Aggiungere una nuova voce dell'area di ricerca inversa. In questo modo verrà aggiunto un mapping IP/host al server DNS.

Se la voce DNS della macchina UNIX non è corretta, è possibile aggiungere manualmente la voce di ricerca inversa sul server DNS.

4. Seguire le istruzioni per la configurazione di IBM SPSS Modeler Client, che si trovano nell'argomento "Configurazione di SSO (Single Sign-On) su Windows" a pagina 16.

### 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.

Per aggiungere connessioni server

1. Nel menu **Strumenti**, fare clic su **Login server**. Viene visualizzata la finestra di dialogo Login server.
2. Nella finestra di dialogo, fare clic su **Aggiungi**. Viene visualizzata la finestra di dialogo Login server: **Aggiungi/Modifica server**.

3. 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 di default non funziona, chiedere il numero di porta corretto all'amministratore di sistema.
  - **Descrizione.** Immettere una descrizione facoltativa della connessione al server.
  - **Verificare la 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

1. Nel menu Strumenti, fare clic su **Login server**. Viene visualizzata la finestra di dialogo Login server.
2. 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.
3. 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.

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. Consultare l'argomento "Aggiunta e modifica della connessione di IBM SPSS Modeler Server" a pagina 17 per ulteriori informazioni.

Per cercare server e cluster

1. Nel menu Strumenti, fare clic su **Login server**. Viene visualizzata la finestra di dialogo Login server.
2. 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.
3. Selezionare un server o un cluster di server dall'elenco.
4. 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:

- **Informazioni origine 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 IBM 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 IBM SPSS Modeler Server, tutte le operazioni effettuate per conto dell'utente vengono eseguite nel contesto di sicurezza di quell'utente. L'accesso alle tabelle di 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 IBM SPSS Modeler Server dal client con l'autenticazione di default di Windows.

**UNIX.** Si presuppone per default che in UNIX IBM 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 IBM 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, consultare l'argomento "Configurazione non root con un database di password privato" a pagina 63.

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

## 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, consultare l'argomento "Spazio temporaneo su disco e RAM" a pagina 10.

**UNIX.** Le versioni UNIX di IBM 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, consultare l'argomento "Verifica delle autorizzazioni per la creazione di file" a pagina 56.

---

## 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 un flusso non ordini esplicitamente i record, l'ordine in cui vengono presentati i record nei flussi 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 IBM 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.** IBM SPSS Modeler in esecuzione in modalità locale, utilizza un formato interno per la memorizzazione dei valori a virgola mobile diverso da IBM SPSS Modeler Server. A causa delle differenze di arrotondamento, i risultati ottenuti con le due versioni possono variare leggermente.

---

## Capitolo 4. 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 IBM 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 o come processo daemon in UNIX.

*Note di pianificazione:* l'arresto di IBM SPSS Modeler Server scollega gli utenti finali e ne termina le sessioni, per cui è opportuno 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).

1. **Windows XP.** Aprire il menu Start di Windows. 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 di Windows. Scegliere **Pannello di controllo**, **Strumenti di amministrazione**, **Servizi**.
2. 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 IBM SPSS Modeler Server.

1. Passare alla directory di installazione di IBM 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 IBM SPSS Modeler Server.
2. Per avviare il server, al prompt dei comandi digitare  

```
./modelersrv.sh start
```
3. Per arrestare il server, al prompt dei comandi digitare  

```
./modelersrv.sh stop
```
4. Per verificare lo stato di IBM 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 IBM SPSS Modeler Server, mentre gli altri processi sono le sessioni di IBM SPSS Modeler.

Il programma di installazione di IBM 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, consultare l'argomento "Avvio e arresto automatico di IBM SPSS Modeler Server" a pagina 55.

## Limiti del kernel UNIX

È necessario verificare che i limiti del kernel sul sistema siano sufficienti a consentire il funzionamento di IBM 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 IBM SPSS Modeler Server. A questo scopo, eseguire le operazioni seguenti:

1. Aggiungere i seguenti comandi a *modelersrv.sh*:

```
ulimit -d unlimited
ulimit -m unlimited
ulimit -f unlimited
```

2. Riavviare IBM 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 IBM 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 server inutilizzabili (processi zombie), si consiglia di arrestare e riavviare IBM SPSS Modeler Server a intervalli regolari, utilizzando la seguente sequenza di comandi:

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

I processi di IBM 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 IBM SPSS Modeler Server a intervalli regolari, si consiglia anche di rimuovere tutti i file rimanenti nella directory temporanea di IBM SPSS Modeler.

Dopo che sono stati chiusi tutti i processi server e che sono stati rimossi tutti i file temporanei, è possibile riavviare IBM 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, consultare l'argomento "Utilizzo di IBM SPSS Modeler Administration Console" a pagina 23.

- Utilizzare il file di testo *options.cfg*, contenuto nella directory [percorso di installazione server]/*config*. Per ulteriori informazioni, consultare l'argomento "Utilizzo del file *options.cfg*" a pagina 29.

Si consiglia di installare IBM SPSS 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 IBM SPSS Modeler Server, mentre IBM SPSS Modeler Administration Console consente di autorizzare chiunque sia provvisto di un account utente alla modifica delle opzioni configurabili. Inoltre, IBM SPSS 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 IBM SPSS Modeler Server.

Ulteriori informazioni sull'utilizzo di IBM SPSS 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.

## 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*. Come valore predefinito, questa riga è impostata come commento per cui è necessario modificarla per consentire l'accesso a utenti specifici oppure 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 caratteri maiuscoli/minuscoli.
- 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.

### Impostazione della connessione al server

1. Verificare che il servizio IBM SPSS Modeler Server sia stato avviato.
2. Dal menu File, scegliere **Nuova** e quindi **Connessione al server amministrato**.
3. 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 **Amministrato IBM SPSS Modeler Server**, poi fare clic su **Avanti**.
4. 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

1. Nel riquadro dell'amministratore, fare doppio clic sul server a cui si desidera accedere per selezionarlo.
2. 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, consultare l'argomento "Configurazione dell'accesso tramite Modeler Administration Console" a pagina 23.

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 IBM SPSS Modeler Server sia in esecuzione. Per esempio, in Windows, scegliere Pannello di controllo > Strumenti di amministrazione > Servizi e selezionare 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 IBM 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

**Numero massimo di connessioni.** (*max\_sessions*) Numero massimo di utenti collegati contemporaneamente. Un valore di -1 indica che non c'è alcun limite.

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

## Connessione Analytic Server

**URL Analytic Server.** (as\_url) L'URL di Analytic Server, incluso l'indirizzo IP del server, la porta e il tenant che è membro dell'installazione SPSS Modeler Server.

**Prompt per la password.** (as\_prompt\_for\_password) Specificare N se SPSS Modeler Server è configurato con lo stesso sistema di autenticazione per gli utenti e password di quello utilizzato in Analytic Server; altrimenti specificare Y.

## Accesso ai file di dati

**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 del file di programma.** (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 del file di programma.** (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 del flusso (non valido per i file di dati SAS e SPSS Statistics). Un valore di -1 indica che non c'è alcun limite.

**Directory temporanea.** (temp\_directory) La directory utilizzata per memorizzare i 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 accedere alla directory (ad esempio, se la directory temporanea è 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

**Riscrittura dei flussi.** (stream\_rewriting\_enabled) Consente al server di ottimizzare i flussi riscrivendoli. Per esempio, è possibile che il server spinga le operazioni di riduzione dei dati più vicino al nodo origine

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 imprevedibili. 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 un flusso. L'impostazione su 0 o su un numero negativo comporta che IBM SPSS Modeler faccia corrispondere il numero di thread al numero di processori disponibili sul computer; il valore predefinito per questa opzione è -1. Per disattivare l'elaborazione parallela (per i computer con 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.

**Dimensione del 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 della cache.** (`cache_compression`) Un compreso 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 proporzione 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 le prestazioni 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 proporzione di memoria fisica assegnata all'addestramento di Rete neurale, Kohonen e modelli *k*-means. Il valore predefinito è il 25%. Aumentare questo valore per migliorare le prestazioni 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 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. È possibile disattivarla se è necessario condividere risorse di memoria sul computer server.

**Porta del server minima e 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 del flusso. 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 predefinito è in entrambi i casi -1, ovvero "nessuna limitazione". Così, per esempio, 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 un flusso e di conseguenza più porte se si desidera aprire o eseguire flussi in contemporanea. Questa operazione è necessaria per catturare il feedback dall'esecuzione del flusso.

Per default, IBM 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 il flusso viene visualizzato un errore. Per configurare l'intervallo delle porte, IBM SPSS Modeler avrà bisogno di due porte aperte disponibili (oltre alla porta server principale) per ciascun flusso contemporaneo, più una porta aggiuntiva per ciascun client connesso che abbia una connessione ODBC. Per eseguire un numero maggiore di flussi 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 array.** (`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 un array. Con alcune combinazioni sistema operativo/database, ciò può migliorare le performance delle istruzioni SELECT.

## SQL

**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, il valori stringa di lunghezza maggiore potrebbero essere troncati a destra senza alcuna segnalazione. L'intervallo valido è compreso tra 1 e 65.535 caratteri. Questa proprietà si applica anche al nodo di esportazione del database.

**Generazione automatica SQL.** (`sql_generation_enabled`) Consente la generazione SQL automatica di flussi 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 modellazione, indipendentemente da questa impostazione.

**Lunghezza della stringa SQL predefinita.** (`default_sql_string_length`). Specifica la larghezza predefinite delle colonne di stringhe che saranno create nelle tabelle della 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 stringa.

**Registrazione tipo SQL supportato**(`supported_sql_type_logging`). Solo per scopi di debug; non modificare se non esplicitamente richiesto da un addetto del supporto tecnico.

**Abilita UDF del database.** (`db_udf_enabled`). Se impostato su Y (valore predefinito), comporta che l'opzione di generazione SQL genera SQL UDF (user-defined function) invece di semplice SQL SPSS Modeler. Solitamente SQL UDF offre prestazioni migliori del semplice SQL.

## SSL

**Abilita 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 di password per il 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

**Host.** (`cop_host`) Nome host o indirizzo IP del servizio Coordinator of Processes. Il valore predefinito "spssc" 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.

**Root di contesto.** (`cop_context_root`) L'URL del servizio Coordinator of Processes.

**Nome login.** (`cop_user_name`) Nome utente per l'autenticazione al servizio Coordinator of Processes. Questo è un nome di login IBM SPSS Collaboration and Deployment Services che potrebbe includere un prefisso di provider di sicurezza (per esempio, `ad/jsmith`).

**Password.** (`cop_password`) La password per l'autenticazione al servizio Coordinator of Processes.

**Abilitato.** (`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.

**Nome server.** (`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.

**Intervallo di aggiornamento (min).** (`cop_update_interval`) Numero di minuti tra i messaggi keep-alive; il valore di default è 2.

**Peso.** (`cop_service_weight`) Il peso di questa istanza, specificato come numero intero compreso tra 1 e 10. Un peso maggiore attira più connessioni. L'impostazione di default è 1.

**Host servizio.** (`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 in options.cfg

amministratori. Specifica i nomi degli utenti a cui si desidera concedere l'accesso come amministratore. Per ulteriori informazioni, consultare l'argomento "Configurazione dell'accesso tramite Modeler Administration Console" a pagina 23.

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

`fips_encryption`. Abilita la crittografia conforme a FIPS; l'impostazione predefinita è N.

`max_transfer_size`. Solo per uso interno del sistema, non modificare.

shell. (solo server UNIX) Ignora l'impostazione predefinita per la shell UNIX, ad 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 su Y se si utilizza SSO (single sign-on) e se si è connessi ad un server IBM SPSS Collaboration and Deployment Services che non utilizza lo stesso metodo di autenticazione dell'host IBM SPSS Modeler Server. Impostare anche questo su Y se si esegue SPSS Modeler Server con un database di password privato.

time\_encode\_execution\_log. Impostare su Y se si vuole che il file *message.log* (nella cartella *log* dell'installazione di SPSS Modeler Server) per includere i singoli tempi di esecuzione di ogni nodo del flusso.

use\_bigint\_for\_count. Quando il numero di record da conteggiare è superiore ad un intero normale ( $2^{32}-1$ ) è possibile mantenere impostata questa opzione su Y. Quando questa opzione è impostata su Y e un flusso è connesso a DB2 o SQL Server, COUNT\_BIG() verrà utilizzata laddove è necessario un conteggio record (ad esempio, il campo **Record\_Count** generato da un nodo di aggregazione).

## Monitoraggio di SPSS Modeler Server

Il riquadro di monitoraggio di Modeler Administration Console visualizza una snapshot 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 una snapshot 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 in questo elenco solo i processi di SPSS Modeler Server, fare clic sul pulsante **Escludi i processi non SPSS Modeler**.

## 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 questo file di configurazione, ma con alcune eccezioni. Per ulteriori informazioni, consultare l'argomento "Opzioni visibili in options.cfg" a pagina 28.

Con l'utilizzo di IBM SPSS 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, consultare l'argomento "Utilizzo di IBM SPSS Modeler Administration Console" a pagina 23.

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:

1. Aprire il file *options.cfg* con un editor di testo.
2. Individuare le opzioni da modificare. Per un elenco completo delle opzioni, consultare "Configurazione di SPSS Modeler Server" a pagina 24.
3. 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.

4. Salvare il file.
5. Arrestare e riavviare IBM SPSS Modeler Server per rendere effettive le modifiche. Per ulteriori informazioni, consultare l'argomento "Avvio e arresto di IBM SPSS Modeler Server" a pagina 21.

## Chiusura delle connessioni di database inutilizzate

Di default, IBM SPSS Modeler memorizza 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 IBM SPSS Modeler non debba riconnettersi al database ogni volta che viene eseguito un flusso. Tuttavia, in alcuni ambienti, è importante che le applicazioni liberino le risorse del database non appena possibile. Se troppe sessioni di IBM 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 IBM SPSS Modeler nel file di configurazione di un database personalizzato. Questa operazione può anche rendere IBM SPSS Modeler meno soggetto 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 IBM SPSS Modeler.

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

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

---

## 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.

1. Quando un client si connette a un server, il client autentica il server con il certificato di chiave pubblica.
2. Il client genera quindi un numero random, lo crittografa con la chiave pubblica del server e invia il messaggio crittografato al server.
3. Il server decrittografa il numero random tramite la chiave privata.
4. Dal numero random, 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:

1. Installare OpenSSL sul server.
2. Ottenere e installare il certificato e le chiavi SSL.
3. Abilitare e configurare SSL nell'applicazione di amministrazione del server (IBM SPSS Collaboration and Deployment Services Deployment Manager).
4. Se lo si desidera, installare la crittografia "Unlimited strength" sui computer client.
5. Aggiungere il certificato al keystore client.
6. Informare gli utenti finali che è necessario abilitare SSL quando ci si connette 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, IBM SPSS Statistics Server è il *client*.

### Installazione di OpenSSL

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

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

Tabella 1. Versioni OpenSSL.

Prodotto server	Versione OpenSSL compatibile
IBM SPSS Statistics 20-22	1.0.0
IBM SPSS Statistics 17-19 (no Linux® su System z®) IBM SPSS Modeler 16 (no Linux® su System z®)	0.9.8 e le relative sottoversioni (0.9.8a, 0.9.8b e così via)
IBM SPSS Statistics 19 (Linux® su System z®) IBM SPSS Modeler 16 (Linux® su System z®)	1.0.0

Nota: IBM SPSS Modeler Server per UNIX, quando è configurato per SSL, tenta di collegarsi a due librerie con nomi come (.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 IBM SPSS Modeler Server.

2. 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.

3. 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_LIBRARY_PATH_64
```

## Ottenere e installare il certificato e le chiavi SSL

1. 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 certificazione autofirmata interna. OpenSSL fornisce un apposito strumento di gestione dei certificati, oppure è possibile cercare su Internet le istruzioni per creare un certificato SSL autofirmato.
2. Copiare il file o i file della chiave e del certificato in una directory locale o nelle directory sul server. Le chiavi pubbliche e private possono essere archiviate in directory separate. Possono anche essere memorizzate in un singolo file. Verificare che le autorizzazioni di accesso vietino l'accesso casuale alla directory che contiene la chiave privata.

## Abilitazione e configurazione SSL in IBM SPSS Collaboration and Deployment Services Deployment Manager

1. Avviare l'applicazione di amministrazione del server ( IBM SPSS Collaboration and Deployment Services Deployment Manager ) e connettersi al server.
2. Nella pagina di configurazione, impostare **Secure Sockets Layer** su Sì.
3. In **File di chiave pubblica SSL**, specificare il percorso completo al file della chiave pubblica.
4. In **File di chiave privata SSL**, specificare il percorso completo al file della chiave privata.  
*Nota:* se la chiave 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**.
5. Dai menu, scegliere:  
**File > Salva**
6. 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**.
- 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**.
- 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 l'aggiornamento alla crittografia "Unlimited strength".

### IBM J9

1. Scaricare i file JCE (Java Cryptography Extension) Unlimited Strength Jurisdiction Policy relativi alla propria versione SDK dal sito web IBM.com.
2. Estrarre i file di Unlimited Jurisdiction Policy che sono impacchettati nel file compresso. Il file compresso contiene un file *US\_export\_policy.jar* ed 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*.
3. Sostituire i propri file *US\_export\_policy.jar* e *local\_policy.jar* con i due file scaricati.
4. 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**.
5. Scollegarsi da Administration Console.
6. Arrestare il server.
7. Riavviare il server.

### Sun Java

1. Scaricare i file JCE (Java Cryptography Extension) Unlimited Strength Jurisdiction Policy relativi alla propria versione SDK dal sito web Sun Java.
2. Decomprimere il file scaricato.
3. Copiare i due file *.jar local\_policy.jar* e *US\_export\_policy.jar* in *<cartella installazione>/jre/lib/security*, dove *<cartella installazione>* è la cartella in cui è stato installato il prodotto.

## Istruzione agli utenti per l'attivazione di SSL

Quando gli utenti si connettono al server tramite un prodotto client, devono attivare SSL nella finestra di dialogo per connettersi al server. Assicurarsi di comunicare agli utenti la casella di controllo appropriata da selezionare. Ad esempio, in IBM SPSS Modeler, gli utenti devono selezionare la casella di controllo **Assicura connessione protetta** nella finestra di dialogo Repository: server, accessibile selezionando l'opzione **Repository** del menu **Strumenti**.



---

## Capitolo 5. 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 un flusso. 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 Capitolo 2, "Architettura e suggerimenti relativi all'hardware", a pagina 7.
- 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 i flussi per ottenere il massimo delle prestazioni.

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 IBM SPSS Modeler Server. Per ulteriori informazioni, consultare l'argomento "Utilizzo di IBM SPSS Modeler Administration Console" a pagina 23.

Le impostazioni sono riunite sotto l'intestazione **Performance e ottimizzazione** nella finestra di configurazione di IBM SPSS 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 dei flussi. Per ulteriori informazioni, consultare l'argomento "Performance/ottimizzazione" a pagina 25.

---

### 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 flusso > Opzioni > Ottimizzazione**

Per ottimizzare le prestazioni dei flussi 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 nel database e ottimizzazione SQL richiedono che la connettività IBM SPSS Modeler Server venga abilitata sul computer IBM SPSS Modeler. Con questa impostazione attivata, è possibile accedere agli algoritmi di database, restituire codice SQL direttamente da IBM SPSS Modeler, e accedere a IBM SPSS Modeler Server. Per verificare lo stato attuale della licenza, scegliere le seguenti opzioni dal menu IBM SPSS Modeler.

#### **Guida > Informazioni su > Dettagli aggiuntivi**

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

Per ulteriori informazioni, consultare l'argomento "Connessione a IBM SPSS Modeler Server" a pagina 15.

*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 16, vedere il sito aziendale dedicato al supporto tecnico all'indirizzo <http://www.ibm.com/support>.

**Consenti riscrittura flusso.** Selezionare questa opzione per consentire la riscrittura dei flussi in IBM SPSS Modeler. Sono disponibili quattro tipi di riscrittura ed è possibile selezionarne uno o entrambi. La riscrittura dei flussi riordina in background i nodi in un flusso per migliorarne il funzionamento, senza alterare la semantica dei flussi.

- **Ottimizza generazione SQL.** Questa opzione consente di riordinare i nodi nel flusso 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 interpretato in codice SQL, l'ottimizzatore controlla più avanti l'eventuale presenza di nodi downstream che è possibile interpretare in SQL e spostare in modo sicuro davanti al nodo del problema senza compromettere la semantica del flusso. Non solo il database può eseguire le operazioni in modo più efficiente di IBM SPSS Modeler, ma questi pushback o "rinvii" hanno anche la possibilità di ridurre la dimensione dell'insieme di dati restituito a IBM 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.
- **Ottimizza espressione CLEM.** Questa opzione consente all'ottimizzatore di ricercare le espressioni CLEM che è possibile pre-elaborare prima di eseguire il flusso in modo da incrementare la velocità di elaborazione. Come semplice esempio, se si ha un'espressione come ad esempio  $\log(\text{salario})$ , l'ottimizzatore dovrebbe calcolare il valore del salario effettivo e passarlo per l'elaborazione. Ciò può essere utilizzato per migliorare il push back SQL e le prestazioni IBM SPSS Modeler Server.
- **Ottimizza esecuzione sintassi.** Questo metodo di riscrittura dei flussi 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 altra esecuzione.** Questo metodo di riscrittura dei flussi aumenta l'efficienza delle operazioni (quelle che non possono essere delegate al database). L'ottimizzazione viene raggiunta riducendo la quantità dei dati nel flusso non appena possibile. Pur mantenendo l'integrità dei dati, il flusso 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.

**Abilitare 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'utilizzo di più nodi o l'utilizzo dei singoli nodi seguenti possono trarre vantaggio dall'elaborazione parallela: C5.0, Unione (per chiave), Ordina, Raccogli (metodi di rango e tile) e Aggregazione (utilizzando uno o più campi chiave).

**Genera SQL.** Selezionare questa opzione per attivare la generazione SQL, consentendo di rinviare operazioni del flusso 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 del flusso.

- **Memorizzazione in cache del database.** Per i flussi che generano SQL da eseguire nel database, i dati possono essere memorizzati 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 un flusso che unisce più tabelle per creare una visualizzazione data mining può essere memorizzato nella cache e riutilizzato in un secondo tempo in base alle esigenze. Con la memorizzazione nella cache del database attivata, è sufficiente fare clic con il pulsante destro del mouse su qualsiasi nodo non terminale per memorizzare i dati nella cache in quel punto e la

cache viene creata automaticamente direttamente nel database alla successiva esecuzione del flusso. Ciò consente la generazione di SQL per i nodi downstream, 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 la memorizzazione nella cache del database o l'ottimizzazione SQL non sono attivate, la cache verrà scritta nel file system.

- **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, i flussi eseguiti in un database potrebbero restituire risultati leggermente diversi da quelli che restituiscono quando vengono eseguiti in IBM SPSS Modeler. Per ragioni analoghe, tali differenze possono variare anche in base al fornitore di database.

---

## Ottimizzazione e utilizzo del database

**Server di 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 IBM SPSS Modeler possono assorbire molte attività sulla macchina IBM SPSS Modeler Server ed eseguono contemporaneamente istruzioni SQL nello stesso database.

**Mining nel database.** Molti fornitori di database offrono 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 di database o su un server dedicato separato. Le funzionalità di mining nel database di IBM SPSS Modeler completano ed estendono le funzioni di generazione SQL, offrendo un sistema per controllare le estensioni database specifiche del fornitore. In alcuni casi, questo approccio consente di evitare overhead di trasferimento dei dati potenzialmente dispendiosi fra IBM 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 del flusso che non possono essere compilate in SQL devono essere eseguite in IBM SPSS Modeler Server. Per ulteriori informazioni, consultare Capitolo 6, "Ottimizzazione SQL", a pagina 39.

### 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 insiemi di dati intermedi derivanti dalla preparazione dei dati e i risultati della distribuzione.

IBM SPSS Modeler può interfacciarsi con i programmi di caricamento 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 IBM 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 programma di caricamento esterno è il programma di utilità *sqlldr* di Oracle.

Tabella 2. Vantaggi del caricamento di massa in termini di performance

<b>Opzione di esportazione</b>	<b>Tempo (in secondi)</b>
Default (ODBC)	409
Caricamento di massa tramite ODBC	52
Caricamento di massa tramite programma di caricamento esterno	33

---

## Capitolo 6. 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 dell'insieme di risultati che devono essere trasformati da DBMS a IBM 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 Selezione, Campione e Aggregazione), che in genere riducono le dimensioni dell'insieme di dati da trasferire. I dati possono anche essere memorizzati nella cache in una tabella temporanea nel database, in punti critici del flusso (per esempio dopo un nodo Unione o Selezione) 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, IBM 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 IBM 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, consultare l'argomento "Suggerimenti per l'ottimizzazione della generazione SQL" a pagina 43.

*Nota:* a causa di piccole differenze di implementazione del codice SQL, i flussi eseguiti in un database potrebbero restituire risultati leggermente diversi quando vengono eseguiti in IBM SPSS Modeler. Per ragioni analoghe, tali differenze possono variare anche in base al fornitore di database. Per esempio, in base alla configurazione del database per la distinzione tra caratteri maiuscoli/minuscoli nel confronto e nella collazione di stringhe, i flussi di IBM 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 IBM SPSS Modeler, i confronti tra stringhe del database devono fare distinzione tra caratteri maiuscoli/minuscoli.

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

### Guida > Informazioni su > Dettagli aggiuntivi

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

Per ulteriori informazioni, consultare l'argomento "Connessione a IBM SPSS Modeler Server" a pagina 15.

*Nota:* quando i flussi vengono eseguiti in un database Netezza, i dettagli di data e ora sono prelevati dal database. Questi valori potrebbero non corrispondere alla data e ora locali o di IBM 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 IBM 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.

---

## Funzionamento della generazione SQL

I frammenti iniziali di un flusso dai nodi origine 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 dei flussi e prima dell'esecuzione, il processo di generazione SQL si verifica come descritto di seguito:

- Il server riordina i flussi per spostare i nodi downstream nella “zona SQL” dove l'operazione può essere eseguita in sicurezza. Questa funzione può essere disattivata nel server.
- Procedendo dai nodi origine 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 origine, 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 del flusso. In base ai risultati, è possibile riorganizzare ulteriormente il flusso, dove appropriato, per beneficiare al massimo dell'esecuzione del database. Per ulteriori informazioni, consultare l'argomento “Suggerimenti per l'ottimizzazione della generazione SQL” a pagina 43.

## Punti di miglioramento

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

- **Unioni (unione per chiave).** Le operazioni di unione 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 del campo.** In un database, i nuovi campi vengono generati in modo più efficiente.
- **Proiezione dei campi.** IBM 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 flat: anche se il server deve leggere i campi superflui, non deve allocare alcuno spazio per archivarli.
- **Calcolo dei punteggi.** La generazione di un SQL può derivare da strutture ad albero delle decisioni, 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.

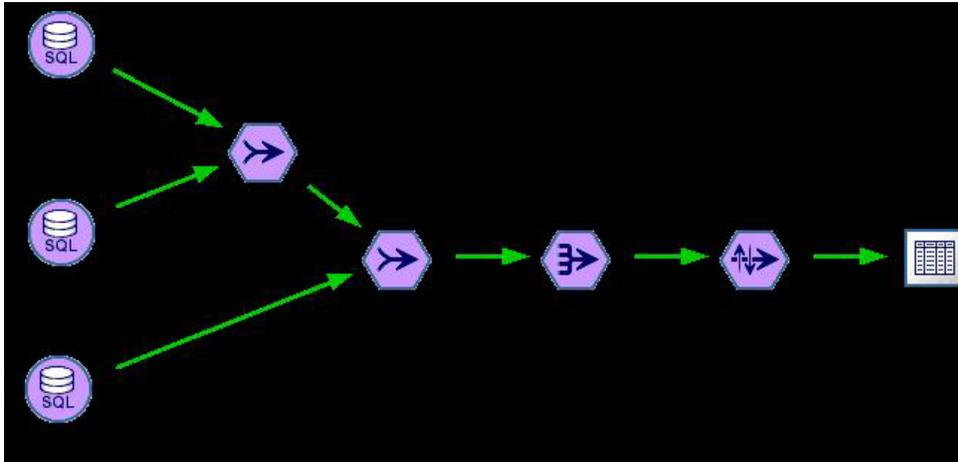


Figura 3. Flusso ottimizzato con nodi porpora che indicano push back SQL (operazioni eseguite nel database)

### 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 del flusso

Quando il flusso termina con un nodo di esportazione del database, è possibile che l'intero stream venga eseguito nel database.

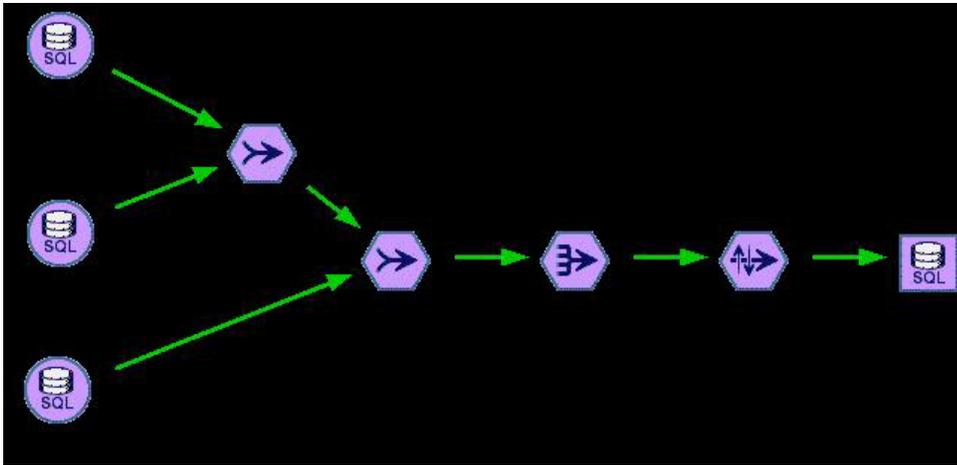


Figura 4. Flusso intero eseguito nel database

## Configurazione dell'ottimizzazione SQL

1. Installare un driver ODBC e configurare una sorgente dati per il database che si desidera utilizzare. Per ulteriori informazioni, consultare l'argomento "Accesso ai dati" a pagina 11.
2. Creare un flusso che utilizza un nodo origine per estrarre i dati dal database.
3. Assicurarsi che la generazione SQL sia attivata sul client e sul server, se applicabile. Per default, è attivata su entrambi.

### Per attivare l'ottimizzazione SQL sul client

1. Dal menu Strumenti, scegliere **Proprietà flusso > Opzioni**.
2. 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, consultare l'argomento "Impostazioni relative alle performance e all'ottimizzazione del client" a pagina 35.

### Per abilitare l'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 su come modificare le impostazioni IBM SPSS Modeler Server, consultare la sezione "Performance/ottimizzazione" a pagina 25

### Per abilitare l'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 un flusso.

---

## 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 il flusso per migliorare le performance.

1. Assicurarsi che le opzioni **Visualizza SQL nel registro messaggi durante l'esecuzione del flusso** e **Visualizza dettagli generazione SQL nel registro messaggi durante la preparazione del flusso** siano selezionate nella finestra di dialogo Opzioni utente. Per ulteriori informazioni, consultare l'argomento "Impostazioni relative alle performance e all'ottimizzazione del client" a pagina 35.
2. Nell'area del flusso, selezionare il nodo o il flusso che si desidera visualizzare in anteprima.
3. 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 il flusso sarà eseguito) sono colorati in porpora nell'area del flusso.
4. Per vedere l'anteprima dell'SQL generato, dai menu scegliere:  
**Strumenti > Proprietà flusso > Messaggi**

---

## Visualizzazione di SQL per i nugget del modello

Per alcuni modelli, è possibile generare SQL per il nugget del modello, 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 ai flussi contenenti questi nugget. Per ulteriori informazioni, consultare l'argomento "Nodo con supporto per generazione SQL" a pagina 44.

Per visualizzare SQL per un nugget del modello che supporta la generazione SQL:

1. Selezionare la scheda Impostazioni nel nugget del modello.
2. Scegliere una delle opzioni **Genera con (senza) supporto per valori mancanti** o **Genera SQL per questo modello** in base alle esigenze.
3. Nel menu del nugget del modello, scegliere:  
**File > Esporta SQL**
4. Salvare il file.
5. 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 flusso.** La generazione SQL può essere interrotta quando 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 supportate da SQL standard. Quando si verifica tale situazione, la generazione SQL viene eliminata anche per tutti i nodi downstream. 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, consultare l'argomento "Nodo con supporto per generazione SQL" a pagina 44.

**CLEM espressioni.** Se un flusso 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 **Ricava**, **Seleziona** e simili possono essere interpretati in codice SQL, purché tutti gli operatori delle espressioni CLEM abbiano equivalenti SQL. La maggior

parte degli operatori può essere interpretata, 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, consultare l'argomento “Espressioni e operatori CLEM con supporto di generazione SQL” a pagina 48.

**Più nodi origine.** Qualora un flusso abbia più nodi origine 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'unione del database e la generazione può proseguire a valle.

**Algoritmo 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.

**Modelli di calcolo del punteggio.** Per alcuni modelli, per il calcolo del punteggio in-database il modello generato viene interpretato 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, consultare l'argomento “Nodo con supporto per generazione SQL”.

Quando si eseguono test sui nodi Modelli per verificare se la generazione SQL per i modelli funziona correttamente, si consiglia di salvare prima tutti i flussi 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 IBM 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 del flusso (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 memorizzati nella cache in una tabella temporanea nel database anziché nel file system. Per ulteriori informazioni, consultare l'argomento “Configurazione dell'ottimizzazione SQL” a pagina 42.

**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 grado 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, consultare l'argomento “Anteprima di SQL generato” a pagina 43.

Tabella 3. 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, consultare l'argomento "Scrittura di query SQL" a pagina 51.

Tabella 4. 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 null, 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 nel variare i gradi in rapporto al database. Consultare Tabella 5.
Aggrega	Il supporto della generazione SQL dell'aggregazione dipende dal tipo di archiviazione dei dati. Consultare Tabella 6 a pagina 46.
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>date_time_date(YEAR,MONTH,DAY)</code> viene rinviata.
Ordina	
Unisci	SQL non generato per l'unione per ordine.  L'unione per chiave con unione esterna completa 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 origine.  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 5. Supporto della generazione SQL nel nodo Campione per il campionamento semplice.

Modalità	Campione	Dimensione massima	Schema	DB2 per z/OS	DB2 per OS/400	DB2 per Windows/UNIX	Netezza (vedi Nota)	Oracle	SQL Server	Teradata	
Include	Primo	n/d		Y	Y	Y	Y	Y	Y	Y	
	1-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		Y
on					Y		Y	Y			
Discard	Primo	off				Y	Y				

Tabella 5. Supporto della generazione SQL nel nodo Campione per il campionamento semplice (Continua).

Modalità	Campione	Dimensione massima	Insieme	DB2 per z/OS	DB2 per OS/400	DB2 per Windows/UNIX	Netezza (vedi Nota)	Oracle	SQL Server	Teradata
		max					Y	Y		
	1-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		

Nota: per i database Netezza, la limitazione riguarda l'uso di parole chiave in interrogazioni secondarie, che vengono utilizzate per eseguire la generazione SQL. Una soluzione alternativa consiste nel consentire la memorizzazione nella cache sul nodo Campione.

Tabella 6. 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		

Tabella 7. Operazioni su campi

Nodo con supporto per generazione SQL	Note
Tipo	Supporta la generazione SQL se il nodo Tipo viene istanziato e nessuna verifica di tipo TERMINA o AVVISA viene specificata.
Filtro	
Ricava	Supporta la generazione SQL se consente l'SQL generato per l'espressione nuovo campo (vedere espressioni di seguito).
Insieme	Supporta la generazione SQL per i target continui. Per gli altri obiettivi, supporta la generazione solo se viene utilizzato il metodo dell'insieme "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 i target continui e la generazione SQL parziale per gli obiettivi nominali e flag.
Ricodifica	
Discretizzazione	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 downstream non la supportano.

Tabella 7. Operazioni su campi (Continua)

Nodo con supporto per generazione SQL	Note
Partition	Supporta l'assegnazione di record alle partizioni da parte della generazione SQL.
Crea flag	
Ristruttura	

Tabella 8. 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	

Tabella 9. Modellazione

Nodo con supporto per generazione SQL	Note
Classificatore automatico	Supporta la generazione SQL per i target continui. Per gli altri obiettivi, supporta la generazione solo se viene utilizzato il metodo dell'insieme "Confidenza più elevata vince".
Numerico automatico	

Per alcuni modelli, è possibile generare SQL per il nugget del modello, 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 ai flussi contenenti questi nugget. Per ulteriori informazioni, consultare l'argomento "Visualizzazione di SQL per i nugget del modello" a pagina 43.

Tabella 10. Nugget del modello

Nugget del modello con supporto per generazione SQL	Note
C&R Tree	Supporta la generazione SQL per l'opzione struttura ad albero singola, ma non per le opzioni di boosting, bagging o insiemi di dati di grandi dimensioni.
QUEST	
CHAID	
C5.0	
Elenco di decisioni	
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.

Tabella 10. Nugget del modello (Continua)

Nugget del modello con supporto per generazione SQL	Note
Insiemi di regole generati	

Tabella 11. 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".
Analisi	Supporta la generazione, in base alle opzioni selezionate.
Trasforma	
Statistiche	Supporta la generazione se non viene utilizzata l'opzione Correla.
Report	
Calcola globali	

Tabella 12. 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.

Tabella 13. Operatori

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

Tabella 14. Operatori relazionali

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

Tabella 14. Operatori relazionali (Continua)

Operazione con supporto di generazione SQL	Note
<	
<=	

Tabella 15. Funzioni

Operazione con supporto di generazione SQL	Note
abs	
allbutfirst	
allbutlast	
and	
arccos	
arcsin	
arctan	
arctanh	
cos	
div	
exp	
fracof	
hasstartstring	
hassubstring	
integer	
intof	
isaplhacode	
islowercode	
isnumbercode	
isstartstring	
issubstring	
isuppercode	
last	
length	
locchar	
log	
log10	
lowertoupper	
max	
member	
min	
negate	
not	
number	

Tabella 15. Funzioni (Continua)

Operazione con supporto di generazione SQL	Note
or	
pi	
real	
rem	
round	
sign	
sin	
sqrt	
string	
strmember	
subscrs	
substring	
substring_between	
uppertolower	
to_string	

Tabella 16. 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 Calcola globali.
@GLOBAL_SUM	
@GLOBAL_MAX	
@GLOBAL_MEAN	
@GLOBAL_MIN	
@GLOBALSDEV	

Tabella 17. Funzioni di aggregazione

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

## Utilizzo di funzioni SQL in espressioni CLEM

È possibile utilizzare la funzione @SQLFN 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 funzione non è contemplato nell'accordo di assistenza standard di IBM SPSS Modeler, poiché l'esecuzione dipende da componenti di database esterni al di fuori del controllo di IBM Corp., ma in casi particolari può essere distribuita, di solito nell'ambito di un accordo 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 downstream.

IBM SPSS Modeler utilizza istruzioni SELECT nidificate per rinviare l'SQL nei flussi che utilizzano una query SQL nel nodo origine del database. In altre parole, il flusso nidifica la query specificata nel nodo origine del database all'interno di una o più istruzioni SELECT generate durante l'ottimizzazione dei nodi downstream. 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 l'unione 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 origine:

```
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;
```



---

## Appendice A. 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.

#### Specifica delle opzioni internazionali

Se IBM 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 IBM SPSS Modeler Server viene eseguito con opzioni internazionali diverse o se queste ultime non sono state definite completamente nella shell in cui IBM SPSS Modeler Server è stato avviato, non è possibile ottimizzare completamente l'SQL generato in Oracle. I motivi sono i seguenti:

- IBM 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 IBM 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 IBM SPSS Modeler Server viene restituito un codice non tradotto, anziché la stringa *Oracle*. Ciò significa che IBM SPSS Modeler non è in grado di ottimizzare stream per Oracle.

Per controllare e reimpostare le specifiche delle opzioni internazionali:

1. In una shell UNIX, eseguire:

```
#locale
```

Vengono restituite le informazioni relative alle opzioni internazionali per la shell. Ad 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
```

2. Passare alla directory Connect ODBC/locale. Qui si trova una directory singola, *en\_US*.
3. 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.ISO8859-15
```

Per opzioni internazionali non inglesi, quali *fr\_FR.ISO8859-1*, è possibile creare il collegamento di tipo soft come segue:

```
#ln -s en_US fr_FR.ISO8859-1
```

4. Una volta creato il collegamento, riavviare IBM SPSS Modeler Server dalla stessa shell. (IBM SPSS Modeler Server riceve le informazioni sulle impostazioni internazionali dalla shell dalla quale è stato 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 nel database, è necessario definire sia `LANG` sia `LC_ALL` nella shell usata per avviare IBM SPSS Modeler Server. È possibile non definire `LANG` nella shell come per altre variabili di ambiente prima di riavviare IBM SPSS Modeler Server. Per esempio, vedere la definizione seguente:  

```
#LANG=en_US.IS08859-15; export LANG
```
- A ogni avvio di IBM 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`.

---

## Appendice B. 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 IBM SPSS Modeler Server per avviarsi automaticamente al riavvio del computer server.
- Arrestare e riavviare manualmente IBM SPSS Modeler Server.
- Modificare le autorizzazioni relative ai file creati da IBM SPSS Modeler Server.
- Configurare IBM SPSS Modeler Server per funzionare con i driver ODBC Connect forniti con il programma IBM SPSS Modeler Server. Per ulteriori informazioni, consultare l'argomento "IBM SPSS Modeler Server e Data Access Pack" a pagina 56.
- Configurare IBM 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, consultare l'argomento "Configurazione di un driver per il nodo Enterprise View" a pagina 61.

---

### Script

IBM SPSS Modeler Server utilizza numerosi script, fra cui:

- **modelersrv.sh.** Lo script di avvio manuale di IBM SPSS Modeler Server è contenuto nella directory di installazione di IBM 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 IBM 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 18. 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 IBM SPSS Modeler Server.

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

1. Accedere come utente root.
2. Passare alla directory di installazione di IBM SPSS Modeler Server.

3. 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 IBM 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

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

```
./modelersrv.sh start
```
3. 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

1. Arrestare IBM SPSS Modeler Server. (Per ulteriori informazioni, consultare l'argomento "Avvio e arresto manuale di IBM SPSS Modeler Server".)
2. Individuare la posizione dello script desiderato. (Per ulteriori informazioni, consultare l'argomento "Script" a pagina 55.)
3. Aprire lo script in un editor di testo, apportare le modifiche e salvare il file.
4. Avviare IBM 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, consultare "Modifica degli script".) Si consiglia di utilizzare 077 come impostazione UMASK più restrittiva. Impostazioni più restrittive potrebbero causare problemi di autorizzazioni per IBM 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 IBM 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, consultare l'argomento "Modifica degli script".)

Per ulteriori informazioni, vedere il sito Web del supporto tecnico all'indirizzo <http://www.ibm.com/support>. Per domande su come creare o impostare le autorizzazioni per le origini dati ODBC, contattare l'amministratore di database.



*Nota:* è importante che qui si rispetti la sintassi. Assicurarsi di lasciare uno spazio tra il primo punto e il percorso del file.

11. Salvare *modelersrv.sh*.

### Per verificare la connessione

1. Riavviare IBM SPSS Modeler Server.
2. Connettersi a IBM SPSS Modeler Server da un client.
3. Sul client, aggiungere un nodo origine del database nell'area.
4. 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, consultare "Risoluzione dei problemi di una configurazione ODBC".

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

Quando la connessione al database può essere stabilita da IBM 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 IBM SPSS Modeler Solution Publisher Runtime.

1. Modificare lo script *modelerrun* in IBM 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 che qui si rispetti la sintassi. Assicurarsi di lasciare uno spazio tra il primo punto e il percorso del file.

2. 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 IBM SPSS Modeler Server viene infatti effettuata da IBM SPSS Modeler Batch per eseguire i flussi. Verificare che la configurazione ODBC di IBM SPSS Modeler Server sia stata eseguita e funzioni correttamente, come descritto in questa sezione.

### Per aggiungere o modificare il nome di una sorgente dati

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

Se la connessione con IBM SPSS Modeler Server funziona correttamente, anche la sorgente dati nuova o modificata dovrebbe funzionare correttamente con IBM SPSS Modeler Solution Publisher Runtime e IBM 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 origine del database e l'elenco delle sorgenti dati disponibili è vuoto o contiene immissioni inattese, il problema può essere dovuto allo script di avvio.

1. 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*.
2. Nel secondo caso, assicurarsi che ODBCINI punti al percorso completo del file *odbc.ini* utilizzato per la definizione delle sorgenti dati ODBC.
3. Se la specifica del percorso in ODBCINI è corretta, verificare il valore di \$ODBCINI utilizzato nell'ambiente di IBM SPSS Modeler Server eseguendo l'echo della 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
```
4. Salvare ed eseguire *modelersrv.sh*. Il valore di \$ODBCINI impostato nell'ambiente di IBM SPSS Modeler Server viene scritto in *stdout* per la verifica.
5. 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 viene eseguito l'echo del valore corretto in *stdout* nel file *modelersrv.sh* in esecuzione, dovrebbe essere possibile visualizzare i nomi delle sorgenti dati nel nodo origine del database quando si riavvia IBM 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 IBM SPSS Modeler Server, eseguire l'echo del 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. Verrà eseguito l'echo del valore del percorso della libreria sul terminale alla successiva esecuzione dello script.

Se viene fatto riferimento a *odbc.sh* da *modelersrv.sh* per impostare l'ambiente ODBC IBM SPSS Modeler Server, eseguire l'echo del percorso della libreria dalla riga dopo quella in cui viene fatto riferimento allo script *odbc.sh*. Per eseguire l'echo del 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 IBM 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,

eseguire l'echo di 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 origine dati non trovato e nessun driver predefinito 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 DNS (data source name), come definito all'interno della sezione [ODBC Data Sources] nella parte superiore 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*:

```
[Origini dati ODBC]
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 stato 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:

```
[Origini dati ODBC]
Oracle=Oracle Wire Protocol

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

1. Verificare che l'oggetto condiviso specificato dal parametro Driver esista.
2. Se errato, correggere il percorso dell'oggetto condiviso.
3. 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, consultare l'argomento "IBM SPSS Modeler Server e Data Access Pack" a pagina 56. Dopo aver eseguito questo script, la stringa "ODBCHOME" dovrebbe essere stata sostituita con il percorso della propria 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 confermare che il driver non può essere caricato. 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:

1. Utilizzare `ivtestlib` per confermare che il driver non può essere caricato. Per esempio, al prompt dei comandi di UNIX, digitare:

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

dove ODBC DIR viene sostituito con il percorso della propria directory di installazione di ODBC.

2. Leggere il messaggio per individuare l'eventuale errore. Per esempio, il messaggio:

**Caricamento di MFor815.so non riuscito: ld.so.1: bin/ivtestlib: errore irreversibile: libclntsh.so: apertura non riuscita: Nessun file o 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).

3. 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 IBM 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
```

4. 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.

5. Correggere il percorso della libreria nello script di avvio di IBM SPSS Modeler.

6. Riavviare IBM 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 19. 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 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. Dopo l'installazione, aggiungere un riferimento a *pev.sh* allo script di avvio in *modelersrv.sh*, *rc.modeler* o in entrambi, con il seguente formato:

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

---

## Appendice C. 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 IBM 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.

**Esecuzione come non-root.** L'esecuzione di IBM 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, IBM 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, consultare l'argomento "Configurazione non root senza un database privato di password".
- **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, consultare l'argomento "Configurazione non root con un database di password privato".

---

### 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 SPSS Modeler Server *options.cfg* per la modifica.
2. Impostare l'opzione `start_process_as_login_user` su Y.
3. Salvare e chiudere il file *options.cfg*.

Considerare che l'esecuzione come non root probabilmente richiede aggiornamenti della configurazione. Per ulteriori informazioni, consultare l'argomento "Risoluzione dei problemi di mancata autenticazione degli utenti" a pagina 65.

*Attenzione:* Non abilitare l'impostazione `start_process_as_login_user` e avviare IBM 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 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 IBM SPSS Modeler Server. L'utilizzo di questo account è riservato al processo daemon di IBM SPSS Modeler Server: In questo esempio, all'account viene assegnato il nome *modelerserv*.

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

- Il gruppo principale deve essere il gruppo *<modelerusers>* creato in precedenza.
  - La home directory può essere la directory di installazione di IBM 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 IBM SPSS Modeler Server utilizzando l'account appena creato. Individuare lo script di avvio appropriato e aprirlo in un editor di testo. Per ulteriori informazioni, consultare l'argomento "Script" a pagina 55.
    - a. 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. Impostare l'opzione *start\_process\_as\_login\_user* su *Y*.
6. 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 IBM SPSS Modeler. *Nota:* queste sono riservate a IBM 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 IBM 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", 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 singolo *>*, il contenuto di *passwords.cfg* sarà sovrascritto ogni volta, con la sostituzione degli utenti impostati in precedenza. Ricordare che, indipendentemente da questo, tutti gli utenti condividono lo stesso account utente UNIX.

*Nota:* se si aggiungono nuovi utenti al database delle password riservate mentre è in esecuzione SPSS Modeler Server, sarà necessario riavviare SPSS Modeler Server in modo che sia possibile riconoscere gli utenti appena definiti. Finché non si fa questa operazione, i login non avranno esito positivo per i nuovi utenti aggiunti via *pwutil* dall'ultimo riavvio di SPSS Modeler Server.

7. Modificare in modo ricorsivo i proprietari della directory di installazione di IBM 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. Ad esempio:  

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

8. Nella directory dei dati può essere utile creare delle sottodirectory per gli utenti di IBM 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 è in esecuzione con la configurazione del file di opzione `start_process_as_login_user` abilitata, è possibile avviare, arrestare e verificare lo stato di SPSS Modeler Server. Per ulteriori informazioni, consultare l'argomento "Per avviare, arrestare e verificare lo stato di Clementine Server in UNIX" a pagina 21.

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 venga visualizzato nella console l'errore `lsattr not found`. Questo perché `lsattr` si trova nella directory `/usr/sbin`, ma quando IBM 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 IBM 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 sicurezza descritto in precedenza.

### Esempio di un ACL (Access Control List)

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:---
```

---

## Appendice D. Configurazione ed esecuzione di SPSS Modeler Server utilizzando un file di password privato su Windows

---

### Introduzione

Queste istruzioni forniscono le informazioni relative all'esecuzione IBM SPSS Modeler Server utilizzando un file di password privato sui sistemi Windows. Con questo metodo IBM SPSS Modeler autentica gli utenti in base ad un database di password privato distinto dall'autenticazione di sistema su Windows.

---

### Configurazione utilizzando un database di password privato

Se si sceglie di autenticare gli utenti mediante un database di password privato, tutte le sessioni utente vengono avviate sullo stesso account utente.

Per configurare IBM SPSS Modeler Server in questo modo, seguire questi passi:

1. Creare l'account utente su cui eseguire IBM SPSS Modeler Server. L'utilizzo di questo account è riservato al processo daemon di IBM SPSS Modeler Server: In questo esempio, all'account viene assegnato il nome *modelerserv*.
2. Modificare il file di opzioni del server, *config/options.cfg*, per impostare l'opzione `start_process_as_login_user` su `Ye` e specificare l'autenticazione rispetto al database di password privato aggiungendo la riga:  
`authentication_methods, "pasw_modeler"`
3. 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 IBM SPSS Modeler. *Nota:* Queste sono private per IBM SPSS Modeler e non hanno nessun rapporto con i nomi utente e password utilizzati per accedere a Windows. 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 IBM 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", 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 singolo `>`, il contenuto di *passwords.cfg* sarà sovrascritto ogni volta, con la sostituzione degli utenti impostati in precedenza. Ricordare che, indipendentemente da questo, tutti gli utenti condividono lo stesso account utente UNIX.

*Nota:* se si aggiungono nuovi utenti al database delle password riservate mentre è in esecuzione SPSS Modeler Server, sarà necessario riavviare SPSS Modeler Server in modo che sia possibile riconoscere gli utenti appena definiti. Finché non si fa questa operazione, i login non avranno esito positivo per i nuovi utenti aggiunti via *pwutil* dall'ultimo riavvio di SPSS Modeler Server.

4. Consentire all'utente creato nella fase 1 il pieno controllo sul file delle opzioni del server: directory *config/options.cfg* and the `%ALLUSERSPROFILE%\IBM\SPSS`.

5. Nei servizi del sistema, arrestare il servizio *IBM SPSS Modeler Server* e modificare l'*accesso* da Account sistema locale nell'account utente creato nella fase 1 e riavviare il servizio.

## Appendice E. 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.

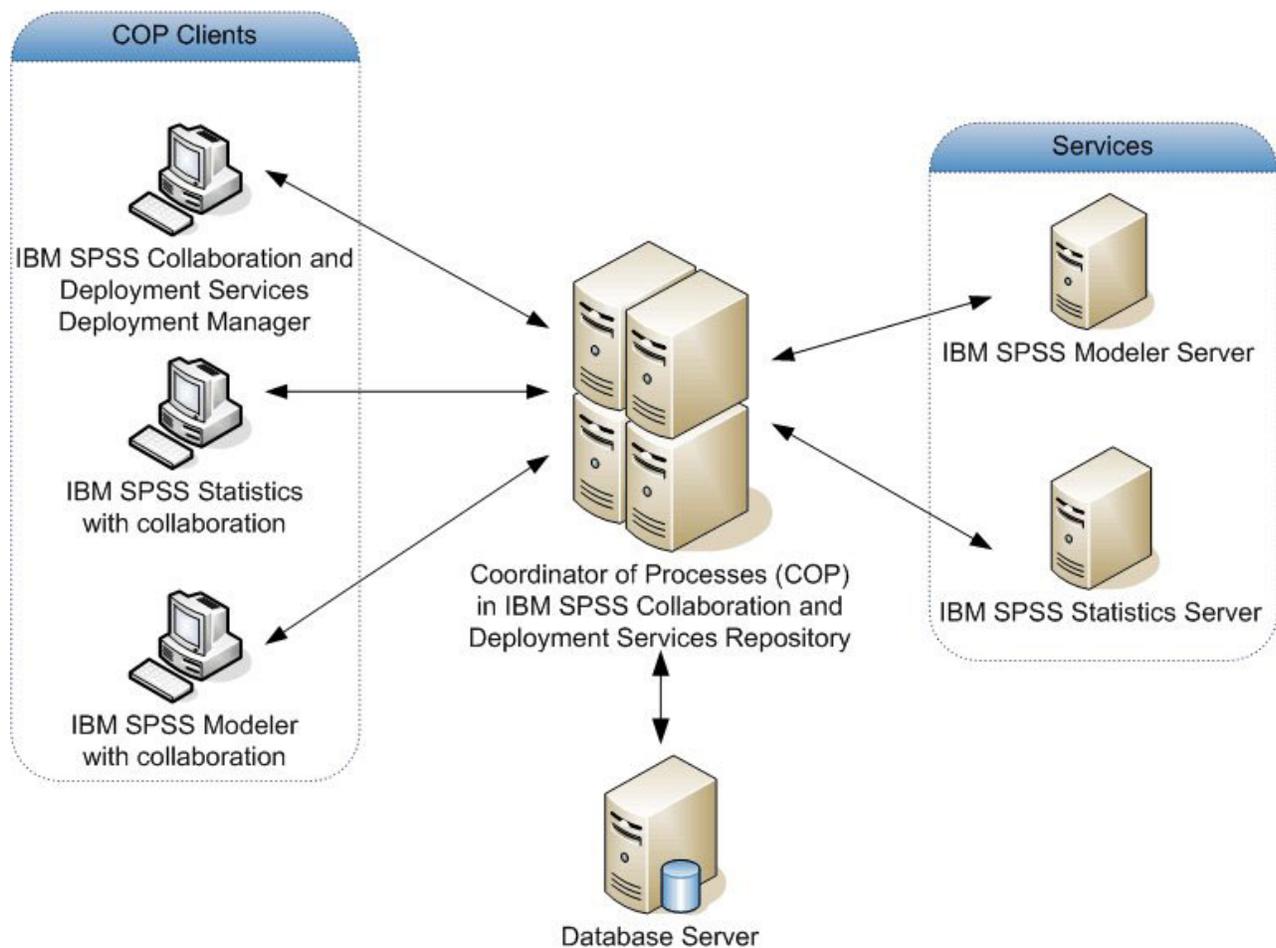


Figura 5. Architettura di Coordinator of Processes

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, consultare *Coordinator of Processes Service Developer's Guide*, disponibile su richiesta da IBM Corp.

Ogni volta che ci si connette a un server o a un cluster di server in IBM SPSS Modeler, è possibile immettere un server manualmente o cercare un server o un cluster con Coordinator of Processes. Per ulteriori informazioni, consultare l'argomento "Connessione a IBM SPSS Modeler Server" a pagina 15.

---

## Note

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

È possibile che IBM non offra i prodotti, servizi o funzioni illustrati in questa documentazione. Consultare il rappresentante locale IBM per le informazioni sui prodotti e servizi attualmente disponibili nella propria zona. Qualsiasi riferimento a un prodotto, programma o servizio IBM non implica o intende dichiarare che può essere utilizzato solo quel prodotto, programma o servizio IBM. In sostituzione a quelli forniti da IBM è possibile utilizzare qualsiasi prodotto, programma o servizio funzionalmente equivalente che non comporti la violazione dei diritti di proprietà intellettuale IBM o altri diritti. Tuttavia, è responsabilità dell'utente valutare e verificare il funzionamento di qualsiasi prodotto, programma o servizio non IBM.

IBM potrebbe avere brevetti o domande di brevetti in corso relativi ad argomenti discussi nella presente pubblicazione. 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 Europe  
Schoenaicher Str.220  
D-7030 Boeblingen  
Deutschland

Per richieste di licenze relative ad informazioni double-byte (DBCS) contattare il Dipartimento di Proprietà Intellettuale IBM nel proprio paese o inviare richieste per iscritto a:

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

Il seguente paragrafo non è valido nel Regno Unito o per tutti i paesi le cui leggi nazionali siano in contrasto con le disposizioni in esso contenute INTERNATIONAL BUSINESS MACHINES CORPORATION FORNISCE QUESTA PUBBLICAZIONE "NELLO STATO IN CUI ESSA SI TROVA" SENZA ALCUNA GARANZIA ESPLICITA O IMPLICITA IVI INCLUSE EVENTUALI GARANZIE DI COMMERCIALIZZABILITÀ ED IDONEITÀ AD UNO SCOPO PARTICOLARE 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 si riserva il diritto di apportare miglioramenti e/o modifiche al prodotto o programma descritto in questa pubblicazione in qualsiasi momento e 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 disponibili sui siti Web non fanno parte di questo prodotto IBM e l'utilizzo di questi è a discrezione dell'utente.

IBM può utilizzare o distribuire qualsiasi informazione fornita dall'utente nel modo che ritiene più idoneo senza incorrere in alcun obbligo nei confronti dell'utente stesso.

Coloro che detengono la licenza su questo programma e desiderano avere informazioni su di esso allo scopo di consentire: (i) lo scambio di informazioni tra programmi indipendenti ed altri (compreso questo) e (ii) l'uso reciproco di tali informazioni dovrebbero contattare:

IBM Software Group  
ATTN: Licensing  
200 W. Madison St.  
Chicago, IL; 60606  
U.S.A.

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

Il programma su licenza descritto in questa documentazione e tutto il materiale su licenza ad esso relativo vengono forniti da IBM nei termini del Customer Agreement IBM IBM International Program License Agreement o di eventuali accordi equivalenti intercorsi 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 testato quei prodotti e non può garantire 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 di business 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 e [ibm.com](http://ibm.com) sono marchi o marchi registrati di International Business Machines Corp., registrati in molte giurisdizioni nel mondo. Altri nomi di prodotti e servizi possono essere marchi di IBM o altre società. Un elenco aggiornato di marchi IBM è disponibile sul sito Web "Copyright and trademark information" all'indirizzo [www.ibm.com/legal/copytrade.shtml](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 logo basati su Java sono marchi o marchi registrati di Oracle e/o suoi affiliati.

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



---

## Indice analitico

### A

- accesso ai dati 11
- accesso come amministratore
  - per IBM SPSS Modeler Server 23
- account utente
  - IBM SPSS Modeler Server 19
- aggiunta di connessioni IBM SPSS Modeler Server 17, 18
- allow\_modelling\_memory\_override
  - file options.cfg 25
- amministrazione
  - di IBM SPSS Modeler Server 22
- anteprima
  - Generazione SQL 43
- architettura
  - lato server 7
- arresto di IBM SPSS Modeler Server
  - in UNIX 21
  - in Windows 21
- autenticazione 19
- autenticazione degli utenti 19
- auto.sh (UNIX)
  - posizione di 55
- autorizzazioni per i file
  - configurazione in UNIX 56
  - su IBM SPSS Modeler Server 19
- avvio automatico del server
  - configurazione in UNIX 55
- avvio di IBM SPSS Modeler Server
  - in UNIX 21
  - in Windows 21

### C

- cache\_compression
  - file options.cfg 25
- cache per nodo
  - Generazione SQL 43
  - scrittura sul database 43
- caching, in-database 29
- caching del database
  - controllo da options.cfg 29
  - Generazione SQL 43
- caching in-database 29
- caricamento in IBM SPSS Modeler Server 15
- chemsrv.sh (UNIX)
  - posizione di 55
- compressione della cache 25
- configurazione di Coordinator of Processes
  - per IBM SPSS Modeler Server 28
- configurazione di COP
  - per IBM SPSS Modeler Server 28
- connessioni
  - a IBM SPSS Modeler Server 15, 17, 18
  - cluster di server 18
- connessioni database
  - chiusura 30

- Coordinator of Processes 18
  - bilanciamento del carico 69
  - cluster di server 69
- COP 18
  - bilanciamento del carico 69
  - cluster di server 69
- cop\_enabled
  - file options.cfg 28
- cop\_host
  - file options.cfg 28
- cop\_password
  - file options.cfg 28
- cop\_port\_number
  - file options.cfg 28
- cop\_service\_description
  - file options.cfg 28
- cop\_service\_host
  - file options.cfg 28
- cop\_service\_name
  - file options.cfg 28
- cop\_service\_weight
  - file options.cfg 28
- cop\_update\_interval
  - file options.cfg 28
- cop\_user\_name
  - file options.cfg 28
- crittografia
  - FIPS 28
  - SSL 30
- crittografia dei dati con SSL
  - abilitazione di IBM SPSS Modeler Server 27
- Crittografia FIPS 28

### D

- data\_file\_path
  - file options.cfg 25
- data\_files\_restricted
  - file options.cfg 25
- database
  - accesso 11
- DB2
  - ottimizzazione SQL 39, 40
- directory temp
  - per IBM SPSS Modeler Server 25
- documentazione 3

### E

- elaborazione parallela
  - controllo 25
- errore durante l'esecuzione del flusso 25
- esecuzione di più flussi 25
- esempi
  - cenni generali 5
  - Guida alle applicazioni 3
- esempi di applicazioni 3
- espressioni CLEM
  - Generazione SQL 48

### F

- file di dati
  - IBM SPSS Statistics 12
  - importazione ed esportazione 12
- file di dati IBM SPSS Statistics
  - importazione ed esportazione 12
- file di registro
  - visualizzazione SQL generato 43
- file options.cfg 29
- file temporanei 10
  - autorizzazioni per (IBM SPSS Modeler Server) 19
- funzione @SQLFN 51

### G

- Generazione SQL 39, 40, 44
  - abilitazione di IBM SPSS Modeler Server 27
  - anteprima 43
  - attivazione 42
  - espressioni CLEM 43, 48
  - registrazione 43
  - risrittura dei flussi 43
  - suggerimenti 43
  - visualizzazione per i nugget del modello 43
- gestione della memoria
  - opzioni di amministrazione 25

### I

- IBM SPSS Analytic Server
  - opzioni di configurazione 25
- IBM SPSS Modeler 1
  - documentazione 3
- IBM SPSS Modeler Administration Console 23
  - accesso come amministratore 23
- IBM SPSS Modeler Server 1
  - accesso come amministratore 23
  - account utente 19
  - amministrazione di 23
  - autenticazione degli utenti 19
  - configurazione di Coordinator of Processes 28
  - configurazione di COP 28
  - creazione di file 19
  - directory temp 25
  - ID utente 15
  - informazioni per gli utenti finali 18
  - monitoraggio dell'utilizzo 29
  - nome dominio (Windows) 15
  - nome host 15, 17
  - numero porta 15, 17, 24
  - opzioni di amministrazione 22
  - opzioni di configurazione 24
  - password 15
  - processi bloccati 22
  - processi server 29

- IBM SPSS Modeler Server (*Continua*)
  - risultati diversi rispetto al client 19
  - single sign-on (SSO) 16, 17
- ID utente
  - IBM SPSS Modeler Server 15
- impostazioni del firewall
  - file options.cfg 25
- impostazioni della porta
  - file options.cfg 25
- io\_buffer\_size
  - file options.cfg 25

## L

- limiti del kernel in UNIX 21
- limiti del kernel UNIX 21

## M

- max\_file\_size
  - file options.cfg 25
- max\_login\_attempts
  - file options.cfg 24
- max\_parallelism
  - file options.cfg 25
- max\_sessions
  - file options.cfg 24
- max\_sql\_string\_length
  - file options.cfg 27
- memoria 10
- memory\_usage
  - file options.cfg 25
- messaggi
  - visualizzazione SQL generato 43
- Microsoft SQL Server
  - ottimizzazione SQL 39, 40
- mining nel database 37
- modellazione
  - gestione della memoria 25
- modelling\_memory\_limit\_percentage
  - file options.cfg 25

## N

- nodi
  - con supporto per generazione SQL 44
- nodi porpora
  - ottimizzazione SQL 41
- nome dominio (Windows)
  - IBM SPSS Modeler Server 15
- nome host
  - IBM SPSS Modeler Server 15, 17
- nomi di file
  - UNIX 12
  - Windows 12
- nomi di file UNC 12
- nugget del modello
  - visualizzazione SQL per 43
- numero porta
  - IBM SPSS Modeler Server 15, 17, 24

## O

- ODBC
  - configurazione in UNIX 56
- Operatori
  - Generazione SQL 48
- opzione cache\_connection 30
- opzioni di configurazione
  - accesso a file di dati 25
  - cenni generali 24, 25
  - connessioni e sessioni 24
  - coordinator of processes 28
  - COP 28
  - crittografia dei dati con SSL 27
  - di IBM SPSS Modeler Server 22
  - directory temp 25
  - elaborazione parallela 25
  - generazione SQL automatica 27
  - gestione della memoria 25
  - lunghezza delle stringhe SQL 27
  - numero porta 24
  - performance e ottimizzazione 25
  - riscrittura dei flussi 25
  - tentativi di accesso 24
- Oracle
  - ottimizzazione SQL 39, 40, 53
- origine dati ODBC e UNIX 56
  - ODBC e script UNIX 56
- ottimizzazione
  - Generazione SQL 39, 40, 42

## P

- pacchetto di accesso ai dati
  - configurazione di UNIX per 56
  - e percorsi delle librerie in UNIX 61
  - ODBC, configurazione in UNIX 56
  - risoluzione dei problemi di ODBC in UNIX 58
- password
  - IBM SPSS Modeler Server 15
- percorsi 12
- performance
  - di IBM SPSS Modeler Server 35
- port\_number
  - file options.cfg 24
- processi, bloccati 22
- processi zombie, IBM SPSS Modeler Server 22
- processori 9
  - multipli 25
- program\_file\_path
  - file options.cfg 25
- program\_files\_restricted
  - file options.cfg 25
- push back 39, 40, 42, 44
  - espressioni CLEM 48
- push back SQL. Vedere anche
  - Generazione SQL 39

## R

- RAM 10
- rc.modeler (UNIX)
  - posizione di 55
- ricerca di connessioni in COP 18

- riscrittura dei flussi 43
- risultati
  - arrotondamento di 19
  - differenze tra Client e Server 19
  - ordine dei record 19

## S

- Script UNIX
  - auto.sh 55
  - modelersrv.sh 55
  - modifica 56
  - rc.modeler 55
- Secure Sockets Layer 30
- server
  - accesso a 15
  - aggiunta di connessioni 17
  - ricerca di server in COP 18
  - single sign-on (SSO) 16, 17
  - server, impostazioni della porta
    - file options.cfg 25
  - server di database 37
  - sicurezza
    - configurazione della creazione di file in UNIX 56
    - creazione di file 19
    - SSL 30
  - single sign-on (SSO) 15
  - sistemi operativi
    - a 64-bit 9
  - sistemi operativi a 64-bit 9
  - spazio su disco 10
    - calcolo 10
- SQL
  - anteprima generata 43
  - nomi di colonna duplicati 51
  - ottimizzazione di Oracle 53
  - query 51
  - visualizzazione per i nugget del modello 43
- sql\_generation\_enabled
  - file options.cfg 27
- SQL Server
  - ottimizzazione SQL 39, 40
- SSL 30
  - cenni generali 30
  - protezione delle comunicazioni 31
- ssl\_certificate\_file
  - file options.cfg 27
- ssl\_enabled
  - file options.cfg 27
- ssl\_private\_key\_file
  - file options.cfg 27
- ssl\_private\_key\_password
  - file options.cfg 27
- stato
  - di IBM SPSS Modeler Server su UNIX 21
  - di IBM SPSS Modeler Server su Windows 21
- stream\_rewriting\_enabled
  - file options.cfg 25
- suggerimenti relativi all'hardware per IBM SPSS Modeler Server 9

## T

Tecnologia data access di IBM SPSS  
Statistics 11  
temp\_directory  
file options.cfg 25

## U

unità disco rigido 10  
UNIX  
autenticazione degli utenti 19  
configurazione delle autorizzazioni  
per i file 56  
percorsi delle librerie 61  
single sign-on (SSO) 17







Printed in Ireland