

Manuale dell'utente di IBM SPSS
Modeler 14.2



Nota: Prima di utilizzare queste informazioni e il relativo prodotto, leggere le informazioni generali disponibili in Note a pag. 265.

Il presente documento contiene informazioni di proprietà di SPSS Inc, una società IBM. Viene fornito su contratto di licenza ed è protetto dalle leggi sul copyright. Le informazioni contenute nella presente pubblicazione non includono alcuna garanzia di prodotto e qualsiasi dichiarazione inclusa nel presente manuale non deve essere interpretata come tale.

Al momento dell'invio di informazioni a IBM o SPSS, l'utente concede a IBM e SPSS un diritto non esclusivo a utilizzare o distribuire tali informazioni nel modo che ritiene appropriato senza incorrere in alcun obbligo nei confronti dell'utente.

© **Copyright IBM Corporation 1994, 2011..**

Prefazione

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

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

Informazioni su IBM Business Analytics

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

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

Supporto tecnico

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

Contenuto

1	Informazioni su IBM SPSS Modeler	1
	IBM SPSS Modeler Server	1
	Opzioni di IBM SPSS Modeler.	2
	IBM SPSS Text Analytics	2
	Documentazione di IBM SPSS Modeler	2
	Esempi di applicazioni.	4
	Cartella Demos.	5
2	Nuove funzionalità	6
	Funzionalità nuove e modificate di IBM SPSS Modeler 14.2.	6
	Nuovi nodi in IBM SPSS Modeler 14.2.	7
3	Panoramica su IBM SPSS Modeler	8
	Guida Introduttiva.	8
	Avvio di IBM SPSS Modeler	8
	Avvio dalla riga di comando	9
	Connessione a IBM SPSS Modeler Server.	9
	Modifica della directory Temp	13
	Avvio di più sessioni di IBM SPSS Modeler	14
	Nozioni di base sull'interfaccia di IBM SPSS Modeler.	14
	Area di disegno dello stream di IBM SPSS Modeler.	15
	Palette nodi.	15
	Manager di IBM SPSS Modeler.	16
	Progetti di IBM SPSS Modeler.	18
	IBM SPSS Modeler Barra degli strumenti	19
	Personalizzazione della barra degli strumenti.	20
	Personalizzazione della finestra di IBM SPSS Modeler	21
	Utilizzo del mouse in IBM SPSS Modeler	22
	Utilizzo dei tasti di scelta rapida.	22
	Stampa.	23
	Automatizzazione di IBM SPSS Modeler	24

4 *Nozioni sul data mining* **25**

Panoramica del data mining	25
Valutazione dei dati	27
Strategia per il data mining	28
Modello di elaborazione CRISP-DM	29
Tipi di modelli	30
Esempi di Oracle Data Mining	37

5 *Creazione di stream* **38**

Panoramica sulla creazione degli stream	38
Creazione degli stream di dati	38
Utilizzo dei nodi	39
Utilizzo degli stream	50
Descrizioni stream	64
Esecuzione degli stream	67
Utilizzo dei modelli	68
Aggiunta di commenti e annotazioni a nodi e stream	69
Salvataggio degli stream di dati	79
Caricamento di file	82
Mappatura degli stream di dati	83
Suggerimenti e tasti di scelta rapida	88

6 *Gestione dei valori mancanti* **91**

Panoramica sui valori mancanti	91
Gestione dei valori mancanti	92
Gestione dei record con valori mancanti	93
Gestione dei campi con valori mancanti	93
Assegnazione o riempimento dei valori mancanti	94
Funzioni CLEM per i valori mancanti	95

7 *Creazione di espressioni CLEM* **97**

Informazioni su CLEM	97
Esempi CLEM	100

Valori e tipi di dati	102
Espressioni e condizioni	103
Parametri stream, sessione e Supernodo	104
Utilizzo di stringhe.	105
Gestione di vuoti e valori mancanti	106
Utilizzo di numeri	107
Utilizzo di date e ore	107
Riepilogo di più campi.	108
Utilizzo di dati a risposta multipla	109
Generatore di espressioni.	110
Accesso al Generatore di espressioni	111
Creazione di espressioni	111
Selezione di funzioni	112
Selezione di campi, parametri e variabili globali.	113
Visualizzazione o selezione di valori.	113
Controllo di espressioni CLEM	115
Trova e sostituisci.	115

8 Riferimenti al linguaggio CLEM 119

Panoramica di riferimento su CLEM	119
Tipi di dati CLEM.	119
Interi	120
Numeri reali.	120
Caratteri	121
Stringhe.	121
Elenchi	121
Campi	121
Date	122
Time	123
Operatori CLEM	123
Riferimenti alle funzioni.	125
Convenzioni utilizzate nelle descrizioni delle funzioni	126
Funzioni di informazione.	127
Funzioni di conversione	128
Funzioni di confronto	129
Funzioni logiche.	131
Funzioni numeriche	132
Funzioni trigonometriche	133
Funzioni di probabilità	134
Operazioni bitwise sui numeri interi	134

Funzioni casuali	136
Funzioni stringa	136
Funzioni SoundEx	141
Funzioni di data e ora	142
Funzioni sequenza	146
Funzioni globali	152
Funzioni per la gestione di valori vuoti e nulli	153
Campi speciali	154

9 Utilizzo di IBM SPSS Modeler con un repository 156

Informazioni su IBM SPSS Collaboration and Deployment Services Repository	156
Archiviazione e deployment di oggetti di IBM SPSS Collaboration and Deployment Services Repository	158
Connessione a IBM SPSS Collaboration and Deployment Services Repository	159
Immissione di credenziali per il IBM SPSS Collaboration and Deployment Services Repository	160
Esplorazione dei contenuti di IBM SPSS Collaboration and Deployment Services Repository	161
Archiviazione di oggetti in IBM SPSS Collaboration and Deployment Services Repository	163
Impostazione delle proprietà degli oggetti	163
Archiviazione di stream	169
Archiviazione di progetti	169
Archiviazione di nodi	170
Archiviazione di oggetti di output	170
Archiviazione di modelli e palette di modelli	171
Recupero di oggetti da IBM SPSS Collaboration and Deployment Services Repository	171
Scelta di un oggetto da recuperare	173
Selezione della versione di un oggetto	174
Ricerca di oggetti in IBM SPSS Collaboration and Deployment Services Repository	174
Modifica degli oggetti di IBM SPSS Collaboration and Deployment Services Repository	177
Creazione, ridenominazione ed eliminazione delle cartelle	177
Blocco e sblocco degli oggetti di IBM SPSS Collaboration and Deployment Services Repository	178
Eliminazione degli oggetti da IBM SPSS Collaboration and Deployment Services Repository	178
Gestione delle proprietà degli oggetti di IBM SPSS Collaboration and Deployment Services Repository	179
Visualizzazione delle proprietà delle cartelle	179
Visualizzazione e modifica delle proprietà degli oggetti	181
Gestione delle etichette di versione degli oggetti	184
Deployment di stream	184
Opzioni di deployment degli stream	186
Il ramo di calcolo del punteggio	189

10 Esportazione in applicazioni esterne **197**

Informazioni sull'esportazione in applicazioni esterne	197
Apertura di uno stream in IBM SPSS Modeler Advantage	197
Creazione guidata applicazioni predittive 4.x	198
Operazioni preliminari all'utilizzo di Creazione guidata applicazioni predittive	199
Esportazione di previsioni binarie come punteggi di propensione	200
Passo 1: Panoramica	200
Passo 2: Selezione di un nodo terminale	201
Passo 3: Selezione di un nodo UCV	202
Passo 4: Specifica di un pacchetto	204
Passo 5: Creazione del pacchetto	204
Passo 6: Riepilogo	206
Importazione ed esportazione di modelli come PMML	206
Tipi di modello che supportano il formato PMML	209

11 Progetti e report **211**

Introduzione ai progetti	211
Visualizzazione CRISP-DM	212
Visualizzazione Classi	213
Creazione di un progetto	213
Creazione di un nuovo progetto	214
Aggiunta di oggetti a un progetto	214
Trasferimento di progetti in IBM SPSS Collaboration and Deployment Services Repository	216
Impostazione delle proprietà del progetto	216
Annotazione di un progetto	218
Proprietà oggetto	220
Chiusura di un progetto	220
Generazione di report	220
Salvataggio ed esportazione dei report generati	224

12 Personalizzazione di IBM SPSS Modeler **226**

Personalizzazione delle opzioni di IBM SPSS Modeler	226
Impostazione delle opzioni di IBM SPSS Modeler	226
Opzioni di sistema	226
Impostazione delle directory di default	228

Impostazione delle opzioni utente	228
Impostazione delle informazioni sull'utente	237
Personalizzazione della palette dei nodi	237
Personalizzazione del Manager palette	237
Modifica della visualizzazione di una scheda della palette	243
Gestione dei nodi CEMI	243

13 Considerazioni sulle performance degli stream e dei nodi 244

Ordine dei nodi	244
Cache dei nodi	245
Performance: Nodi di elaborazione	247
Performance: Nodi Modelli	248
Performance: espressioni CLEM	249

Appendici

A Accesso facilitato in IBM SPSS Modeler 250

Panoramica delle funzionalità per l'accesso facilitato in IBM SPSS Modeler	250
Tipi di funzionalità per l'accesso facilitato	250
Accesso facilitato per utenti con problemi di vista	250
Accesso facilitato per utenti non vedenti	251
Accesso facilitato mediante la tastiera	252
Uso di un'utilità di lettura dello schermo	260
Suggerimenti per l'utilizzo	261
Interferenze con altri programmi	262
JAWS e Java	263
Utilizzo dei grafici in IBM SPSS Modeler	263

B Supporto Unicode 264

Supporto Unicode in IBM SPSS Modeler	264
--	-----

C Note

265

Indice

268

Informazioni su IBM SPSS Modeler

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

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

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

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. Oltre a quelli elencati di seguito, potrebbero essere disponibili altri prodotti e aggiornamenti. Per ulteriori informazioni, vedere <http://www.ibm.com/software/analytics/spss/products/modeler/>.

SPSS Modeler. SPSS Modeler è una versione del prodotto completa nelle funzioni, che viene installata ed eseguita sul desktop dell'utente. È possibile eseguirla 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.

SPSS Modeler Server. SPSS Modeler Server viene eseguito continuamente in modalità di analisi distribuita assieme a una o più installazioni di IBM® SPSS® Modeler. Una configurazione di questo tipo consente di ottenere performance superiori 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 necessità di scaricare i dati sul computer client. SPSS Modeler Server supporta inoltre funzionalità di ottimizzazione SQL e modellazione in-database, con ulteriori vantaggi in termini di prestazioni e grado di automazione. Per eseguire un'analisi deve essere disponibile almeno un'installazione SPSS Modeler.

Opzioni di IBM SPSS Modeler

I seguenti componenti e funzionalità possono essere acquistati separatamente e concessi in licenza per l'utilizzo con SPSS Modeler. Tenere presente che potrebbero rendersi disponibili anche prodotti e aggiornamenti aggiuntivi. Per ulteriori informazioni, vedere <http://www.ibm.com/software/analytics/spss/products/modeler/>.

- Accesso a SPSS Modeler Server per una maggiore scalabilità e prestazioni migliori su insiemi di dati di grandi dimensioni, nonché per il supporto di funzionalità di ottimizzazione SQL e modellazione in-database.
- SPSS Modeler Solution Publisher, per il calcolo automatico o in tempo reale del punteggio al di fuori dall'ambiente di SPSS Modeler. Per ulteriori informazioni, vedere l'argomento [IBM SPSS Modeler Solution Publisher in il capitolo 2 in IBM SPSS Modeler 14.2 Solution Publisher](#).
- Adattatori per il deployment in IBM SPSS Collaboration and Deployment Services o nell'applicazione thin-client IBM SPSS Modeler Advantage. Per ulteriori informazioni, vedere l'argomento [Archiviazione e deployment di oggetti di IBM SPSS Collaboration and Deployment Services Repository in il capitolo 9 a pag. 158](#).

IBM SPSS Text Analytics

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

- Il nodo Text Mining offre funzioni di modellazione di concetti e categorie nonché un workbench interattivo in cui è possibile eseguire esplorazioni avanzate di collegamenti e cluster di testo, creare categorie personalizzate e rifinire i modelli di risorse linguistiche.
- Sono supportati numerosi formati di importazione, inclusi blog e altre sorgenti basate sul Web.
- Sono inclusi inoltre modelli, librerie e dizionari personalizzati per domini specifici, quali CRM e genomica.

Nota: per l'accesso a questo componente, è necessaria una licenza distinta. Per ulteriori informazioni, vedere <http://www.ibm.com/software/analytics/spss/products/modeler/>.

Documentazione di IBM SPSS Modeler

La documentazione completa 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 è disponibile nella cartella *\Documentation* del DVD di ciascun prodotto.

- **Manuale dell'utente di IBM SPSS Modeler.** Introduzione generale all'utilizzo di SPSS Modeler che illustra come creare stream di dati, gestire valori mancanti, generare espressioni CLEM, utilizzare progetti e report e assemblare stream per il deployment tramite IBM SPSS Collaboration and Deployment Services, le applicazioni predittive o IBM SPSS Modeler Advantage.
- **IBM SPSS Modeler Source, Process, and Output Nodes.** Descrizioni di tutti i nodi utilizzati per leggere, elaborare e generare dati di output in vari formati, ovvero di nodi ad eccezione dei nodi Modelli.
- **IBM SPSS Modeler Nodi Modelli.** Descrizioni di tutti i nodi utilizzati per creare modelli di data mining. IBM® SPSS® Modeler offre numerosi metodi di modellazione ricavati dall'apprendimento automatico, dall'intelligenza artificiale e dalla statistica. [Per ulteriori informazioni, vedere l'argomento Panoramica sui nodi Modelli in il capitolo 3 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)
- **IBM SPSS Modeler Algorithms Guide.** Descrizione dei fondamenti di matematica per i metodi di modellazione utilizzati in SPSS Modeler.
- **IBM SPSS Modeler Guida alle applicazioni.** Gli esempi inclusi in questa guida forniscono indicazioni mirate e sintetiche su specifici metodi e tecniche di modellazione. Una versione in linea di questa guida è inoltre disponibile dal menu Aiuto. [Per ulteriori informazioni, vedere l'argomento Esempi di applicazioni a pag. 4.](#)
- **IBM SPSS Modeler Script e automazione.** Informazioni sulle modalità di automazione del sistema tramite script, incluse le proprietà che è possibile utilizzare per manipolare nodi e stream.
- **IBM SPSS Modeler Deployment Guide.** Informazioni sull'esecuzione di stream e scenari SPSS Modeler come fasi dell'elaborazione di lavori in IBM® SPSS® Collaboration and Deployment Services Deployment Manager.
- **IBM SPSS Modeler Guida per lo sviluppatore CLEF.** CLEF consente di integrare programmi di terze parti (quali routine di elaborazione di dati o algoritmi di modellazione) come nodi in SPSS Modeler.
- **IBM SPSS Modeler Guida alla modellazione in-database.** Informazioni sulle modalità per utilizzare al meglio la potenza del database in uso al fine di ottenere prestazioni migliori ed estendere la gamma di funzionalità analitiche tramite algoritmi di terze parti.
- **IBM SPSS Modeler Server e Guida della performance.** Informazioni su come configurare e amministrare IBM® SPSS® Modeler Server.
- **Manuale dell'utente di IBM SPSS Modeler Administration Console.** Informazioni sull'installazione e l'utilizzo dell'interfaccia utente della console per il monitoraggio e la configurazione di SPSS Modeler Server. La console viene implementata come plug-in dell'applicazione Deployment Manager.
- **IBM SPSS Modeler Solution Publisher Guide.** SPSS Modeler Solution Publisher è un componente aggiuntivo che consente di pubblicare gli stream al di fuori dell'ambiente SPSS Modeler standard.
- **Guida CRISP-DM di IBM SPSS Modeler.** Guida passo a passo al data mining tramite la metodologia CRISP-DM con SPSS Modeler.

Esempi di applicazioni

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

È possibile accedere agli esempi facendo clic su Esempi di applicazioni nel menu Aiuto di SPSS Modeler. I file di dati e gli stream di esempio sono installati nella cartella *Demos* nella directory di installazione del prodotto. [Per ulteriori informazioni, vedere l'argomento Cartella Demos a pag. 5.](#)

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

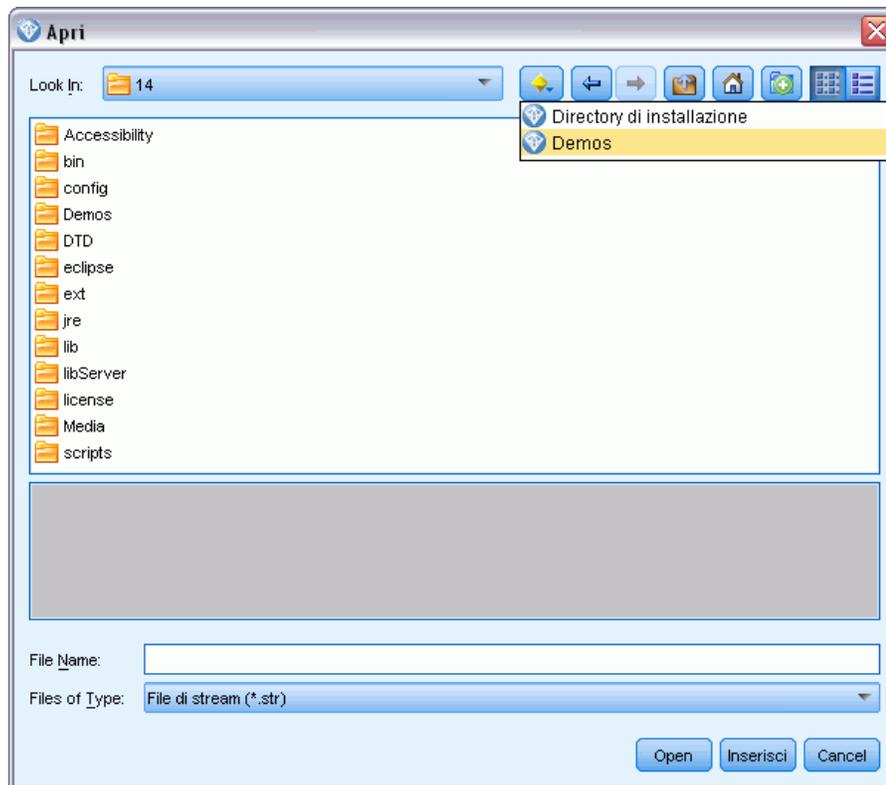
Esempi di script. Vedere gli esempi nella *IBM SPSS Modeler Guida per script e automazione*.

Cartella Demos

I file di dati e gli stream di esempio utilizzati negli esempi di applicazioni sono installati nella cartella *Demos* nella directory di installazione del prodotto. A questa cartella è possibile accedere anche dal gruppo di programmi IBM SPSS Modeler 14.2 nel menu Start di Windows oppure facendo clic su *Demos* nell'elenco delle directory recenti nella finestra di dialogo Apri file.

Figura 1-1

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



Nuove funzionalità

Funzionalità nuove e modificate di IBM SPSS Modeler 14.2

Nella versione IBM® SPSS® Modeler 14.2 sono state aggiunte le seguenti nuove funzionalità.

Supporto mining in-database per IBM Netezza Analytics. Questa versione supporta il mining in-database per la tecnologia IBM Netezza Analytics. [Per ulteriori informazioni, vedere l'argomento IBM SPSS Modeler e IBM Netezza Analytics in il capitolo 6 in *IBM SPSS Modeler 14.2 Guida alla modellazione in-database*.](#)

Il supporto viene fornito mediante due nuovi nodi di modellazione database: gli alberi decisionali Netezza e K-Means Netezza. [Per ulteriori informazioni, vedere l'argomento Nuovi nodi in IBM SPSS Modeler 14.2 a pag. 7.](#)

Supporto di livello 1 per i database IBM Netezza. I database IBM Netezza ora sono supportati al livello 1 e questo significa che ora è disponibile tutto il push back SQL possibile, con ottimizzazione SQL specifica del database. [Per ulteriori informazioni, vedere l'argomento Accesso ai dati in il capitolo 2 in *IBM SPSS Modeler Server 14.2 Guida della performance e amministrazione*.](#)

Supporto avanzato per la generazione SQL nel nodo Campione. Il supporto per la generazione SQL nel nodo Campione quando si utilizza il campionamento semplice è stato ottimizzato per i database DB2 eseguiti in Windows e UNIX e per i database IBM Netezza e Teradata. [Per ulteriori informazioni, vedere l'argomento Nodo con supporto per generazione SQL in il capitolo 6 in *IBM SPSS Modeler Server 14.2 Guida della performance e amministrazione*.](#)

Supporto per Cognos 10. I nodi di input e di esportazione di IBM Cognos BI ora supportano Cognos BI versione 10.1 oltre alla versione 8.4. Per passare dalla versione 8.4 (default) alla versione 10.1, eseguire l'utilità appropriata per Cognos 10.1 (file batch di Windows o script di shell UNIX) disponibile sul sito del supporto tecnico aziendale all'indirizzo <http://www.ibm.com/support>. Se in un secondo tempo si desidera tornare alla versione 8.4, sarà necessario eseguire l'utilità corrispondente per Cognos 8.4, disponibile allo stesso indirizzo.

Supporto lato server per IBM System z. SPSS Modeler dispone ora del supporto sul lato server per i sistemi IBM System z con Red Hat Enterprise Linux, Red Hat Enterprise Linux Advanced Platform o SuSE Linux Enterprise Server. Per ulteriori informazioni, vedere le *Istruzioni di installazione di IBM SPSS Modeler Server 14.2 per UNIX* contenute nel DVD di IBM® SPSS® Modeler Server.

Nuovi nodi in IBM SPSS Modeler 14.2



Il nodo Alberi decisionali di Netezza consente di creare un albero decisionale da utilizzare per la previsione o la classificazione di osservazioni future. Il metodo utilizza l'esecuzione ricorsiva di partizioni per suddividere i record di addestramento in sottogruppi riducendo l'impurità ad ogni passaggio. Un nodo dell'albero è considerato "puro" quando il 100% dei casi nel nodo rientra in una categoria specifica, detta "etichetta di classe", del campo obiettivo. [Per ulteriori informazioni, vedere l'argomento Alberi decisionali di Netezza in il capitolo 6 in IBM SPSS Modeler 14.2 Guida alla modellazione in-database.](#)



Il nodo K-Means di Netezza esegue l'analisi dei cluster e consente così di suddividere i membri di un insieme di dati in gruppi di record simili. L'algoritmo esegue varie iterazioni dello stesso processo di base, in cui ogni istanza di addestramento viene assegnata al cluster più vicino. Tutti i centri dei cluster vengono in seguito ricalcolati come vettori del valore medio degli attributi delle istanze assegnate a determinati cluster. [Per ulteriori informazioni, vedere l'argomento Netezza K-Means in il capitolo 6 in IBM SPSS Modeler 14.2 Guida alla modellazione in-database.](#)

Panoramica su IBM SPSS Modeler

Guida Introduttiva

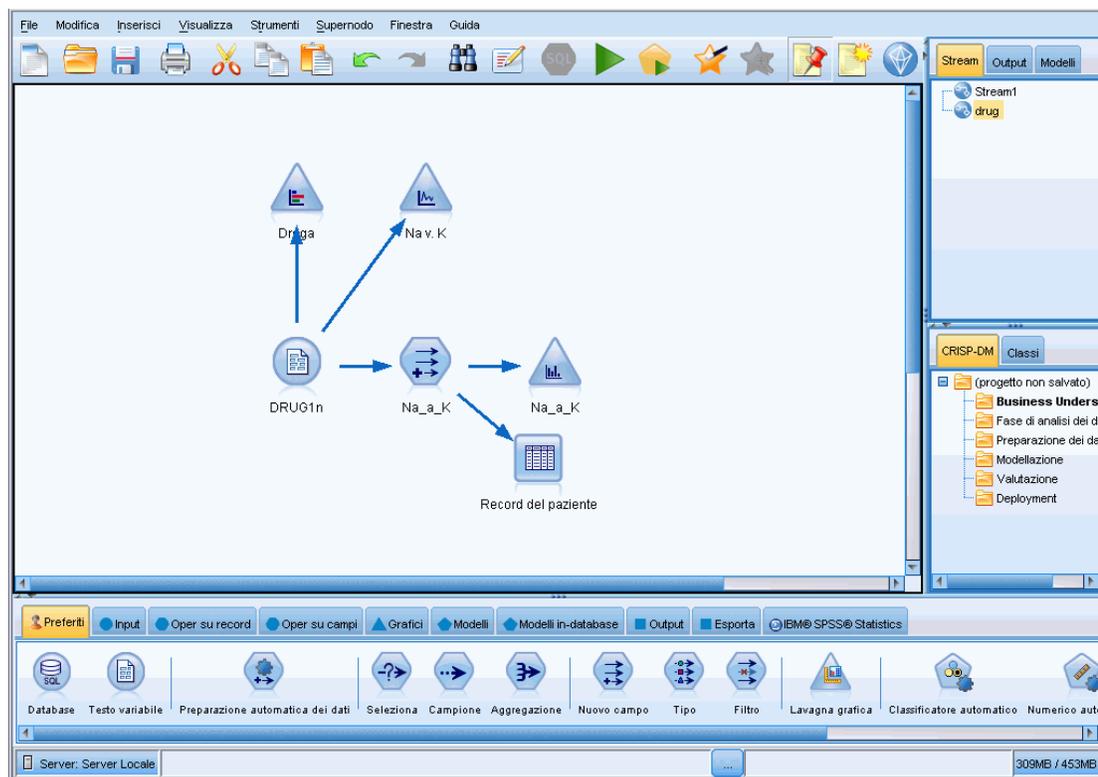
IBM® SPSS® Modeler è un'applicazione di data mining che offre un approccio strategico al rilevamento di relazioni utili all'interno di insiemi di dati di grandi dimensioni. Rispetto ai metodi statistici più tradizionali, non è necessario conoscere fin dall'inizio che cosa cercare esattamente. È possibile esplorare i dati, adattarvi modelli diversi e analizzare le diverse relazioni, fino a individuare informazioni utili.

Avvio di IBM SPSS Modeler

Per avviare l'applicazione, fare clic su
Start > [Tutti] i programmi > IBM SPSS Modeler14.2 > IBM SPSS Modeler14.2

Dopo qualche secondo viene visualizzata la finestra principale.

Figura 3-1
Finestra principale di IBM SPSS Modeler



Avvio dalla riga di comando

È possibile utilizzare la riga di comando del sistema operativo per avviare IBM® SPSS® Modeler:

- ▶ Sul computer in cui è installato IBM® SPSS® Modeler, aprire una finestra DOS (prompt dei comandi).
- ▶ Per avviare l'interfaccia di SPSS Modeler in modalità interattiva, digitare il comando `modelerclient` seguito dagli argomenti desiderati; per esempio:

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

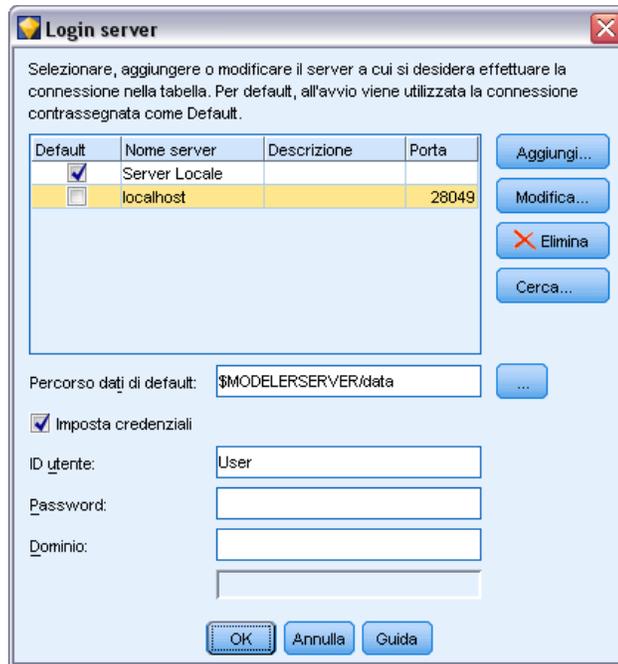
Gli argomenti disponibili (flag) consentono di connettersi a un server, caricare stream, eseguire script o specificare altri parametri.

Connessione a IBM SPSS Modeler Server

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

Per connettersi a un server, è possibile immettere manualmente il nome del server al quale connettersi oppure selezionare un nome definito in precedenza. Tuttavia, se si dispone di IBM SPSS Collaboration and Deployment Services, è possibile cercare nell'elenco di server o cluster di server disponibili nella finestra di dialogo Login server. Il plug-in Coordinator of Processes consente di spostarsi tra i servizi Statistics in esecuzione in una rete. [Per ulteriori informazioni, vedere l'argomento Bilanciamento del carico con cluster di server in l'appendice D in IBM SPSS Modeler Server 14.2 Guida della performance e amministrazione.](#)

Figura 3-2
Finestra di dialogo Login server



Per connettersi a un server

- ▶ Nel menu Strumenti, fare clic su Login server. Viene visualizzata la finestra di dialogo Login server. In alternativa, fare doppio clic sull'area dello stato di connessione nella finestra di SPSS Modeler.
- ▶ Nella finestra di dialogo specificare le opzioni per la connessione al server locale oppure selezionare una connessione dalla tabella.
 - Fare clic su Aggiungi o Modifica per aggiungere o modificare una connessione. [Per ulteriori informazioni, vedere l'argomento Aggiunta e modifica della connessione di IBM SPSS Modeler Server a pag. 11.](#)
 - Fare clic su Cerca per accedere al server o al cluster di server in Coordinator of Processes. [Per ulteriori informazioni, vedere l'argomento Ricerca di server in IBM SPSS Collaboration and Deployment Services a pag. 12.](#)

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

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

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

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

Password. Immettere la password associata al nome utente specificato.

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

- ▶ Fare clic su OK per completare la connessione.

Per disconnettersi da un server

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

Aggiunta e modifica della connessione di IBM SPSS Modeler Server

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

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

Figura 3-3

Finestra di dialogo Login server: Aggiungi/Modifica server



Per aggiungere connessioni server

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

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

Per modificare connessioni server

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

Ricerca di server in IBM SPSS Collaboration and Deployment Services

Anziché immettere una connessione server manualmente, è possibile selezionare un server o un cluster di server della rete tramite **Coordinator of Processes**, disponibile in IBM® SPSS® **Collaboration and Deployment Services**. Un cluster di server è un gruppo di server tra i quali **Coordinator of Processes** stabilisce quello più adatto a rispondere a una richiesta di elaborazione. [Per ulteriori informazioni, vedere l'argomento **Bilanciamento del carico con cluster di server** in l'appendice D in *IBM SPSS Modeler Server 14.2 Guida della performance e amministrazione*.](#)

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

Nota: se non si dispone della funzionalità **Coordinator of Processes**, è comunque possibile immettere manualmente il nome del server a cui connettersi oppure selezionare un nome definito in precedenza. [Per ulteriori informazioni, vedere l'argomento **Aggiunta e modifica della connessione di IBM SPSS Modeler Server** a pag. 11.](#)

Figura 3-4
Finestra di dialogo Cerca server



Per cercare server e cluster

- ▶ Nel menu Strumenti, fare clic su Login server. Viene visualizzata la finestra di dialogo Login server.
- ▶ Nella finestra di dialogo, fare clic su Cerca per aprire la finestra di dialogo Cerca server. Se non si è ancora connessi a IBM SPSS Collaboration and Deployment Services quando si tenta di sfogliare Coordinator of Processes, viene richiesto di connettersi. [Per ulteriori informazioni, vedere l'argomento Connessione a IBM SPSS Collaboration and Deployment Services Repository in il capitolo 9 a pag. 159.](#)
- ▶ Selezionare un server o un cluster di server dall'elenco.
- ▶ Fare clic su OK per chiudere la finestra di dialogo e aggiungere la connessione alla tabella nella finestra di dialogo Login server.

Modifica della directory Temp

Alcune operazioni eseguite da IBM® SPSS® Modeler Server potrebbero richiedere la creazione di file temporanei. Per default, per la creazione di file temporanei in IBM® SPSS® Modeler viene utilizzata la directory temporanea di sistema. È possibile modificare il percorso della directory temporanea, attenendosi alla procedura seguente.

- ▶ Creare una nuova directory con nome *spss* e una sottodirectory con nome *servertemp*.
- ▶ Modificare il file *options.cfg*, disponibile nella directory */config* dell'installazione di SPSS Modeler in uso. Modificare il parametro *temp_directory* in tale file come segue: *temp_directory*, "C:/spss/servertemp".
- ▶ Dopo avere eseguito tale operazione, è necessario riavviare il servizio SPSS Modeler Server, facendo clic sulla scheda Servizi nel Pannello di controllo di Windows. Arrestare il servizio, quindi avviarlo per attivare le modifiche apportate. È possibile riavviare il servizio anche riavviando il computer.

Tutti i file temporanei verranno quindi scritti nella nuova directory.

Nota: l'errore più comune in questo tipo di operazione è l'utilizzo del tipo di barra non corretto. In conformità con UNIX, in SPSS Modeler vengono utilizzate le barre (/).

Avvio di più sessioni di IBM SPSS Modeler

Per avviare più sessioni di IBM® SPSS® Modeler contemporaneamente occorre apportare alcune modifiche alle impostazioni di IBM® SPSS® Modeler e di Windows. Questo può essere necessario, per esempio, se si dispone di due licenze server separate e si desidera eseguire due stream su due server diversi dallo stesso computer client.

Per attivare sessioni multiple di SPSS Modeler:

- ▶ Fare clic su
Start > [Tutti] i programmi > IBM SPSS Modeler14.2
- ▶ Con il pulsante destro del mouse, fare clic sul collegamento IBM SPSS Modeler14.2 (quello con l'icona) e selezionare Proprietà.
- ▶ Nella casella di testo Destinazione, aggiungere -noshare alla fine della stringa.
- ▶ In Esplora risorse, selezionare:
Strumenti > Opzioni cartella...
- ▶ Nella scheda Tipi di file, selezionare l'opzione Stream di SPSS Modeler e fare clic su Avanzate.
- ▶ Nella finestra di dialogo Modifica tipo file, selezionare Open with SPSS Modeler e fare clic su Modifica.
- ▶ Nella casella di testo Applicazione utilizzata per eseguire l'operazione, aggiungere -noshare prima dell'argomento -stream.

Nozioni di base sull'interfaccia di IBM SPSS Modeler

In ogni fase del processo di data mining, l'interfaccia visiva di IBM® SPSS® Modeler sollecita le capacità commerciali specifiche dell'utente. Gli algoritmi di creazione di modelli, quali quelli di previsione, classificazione, segmentazione e individuazione delle associazioni, assicurano modelli potenti e precisi. I modelli possono essere facilmente implementati e letti nei database, nell'IBM® SPSS® Statistics e in molte altre applicazioni.

Il processo di elaborazione dei dati in SPSS Modeler è articolato in tre passaggi.

- In primo luogo, si leggono i dati in SPSS Modeler.
- Quindi li si sottopone a una serie di manipolazioni.
- Infine, li si invia a una destinazione.

Questa sequenza di operazioni è nota come **stream di dati** perché i dati fluiscono, record dopo record, dalla sorgente attraverso ogni manipolazione e infine arrivano alla destinazione, che può essere un modello o un tipo di output di dati.

Figura 3-5
Uno stream semplice



Area di disegno dello stream di IBM SPSS Modeler

L'area di disegno dello stream è l'area di maggiori dimensioni della finestra di IBM® SPSS® Modeler. In quest'area vengono costruiti e manipolati gli stream di dati.

Gli stream vengono creati disegnando diagrammi delle operazioni sui dati pertinenti per la propria azienda nell'area di disegno principale dell'interfaccia. Ogni operazione è rappresentata da un'icona o **nodo** e i nodi sono collegati insieme in uno **stream** che rappresenta il flusso dei dati attraverso ciascuna operazione.

Con SPSS Modeler è possibile utilizzare più stream contemporaneamente, nella stessa area di disegno o aprendo una nuova area di disegno dello stream. Nel corso di una sessione, gli stream vengono archiviati nel relativo Manager, nella parte superiore destra della finestra di SPSS Modeler.

Palette nodi

La maggior parte dei dati e degli strumenti di modellazione di IBM® SPSS® Modeler si trovano nella **palette dei nodi**, visualizzata nella sezione sottostante l'area di disegno dello stream.

Per esempio, la scheda Oper su record della palette contiene i nodi utilizzabili per eseguire operazioni sui **record** di dati quali la selezione, l'unione e l'accodamento.

Per aggiungere nodi all'area di disegno, fare doppio clic sulle icone dalla palette dei nodi oppure trascinare e rilasciare le icone nell'area di disegno. Connettere quindi le icone per creare uno **stream** che rappresenti il flusso di dati.

Figura 3-6
Scheda Oper su record della palette dei nodi



Ogni scheda della palette contiene un insieme di nodi correlati che viene utilizzato per le diverse fasi delle operazioni degli stream, per esempio:

- **Input.** Nodi che consentono di inserire i dati in SPSS Modeler.
- **Oper su record.** Nodi utilizzati per le operazioni sui **record** di dati, per esempio la selezione, l'unione e l'accodamento.
- **Oper su campi.** Nodi utilizzati per le operazioni sui **campi** di dati, per esempio l'applicazione di filtri, la derivazione di nuovi campi e la definizione del livello di misurazione per campi specifici.
- **Grafici.** Nodi utilizzati per rappresentare graficamente i dati prima e dopo la fase di modellazione. I grafici includono i plot, gli istogrammi, i nodi Web e i grafici di valutazione.
- **Creazione di modelli.** Nodi che utilizzano gli algoritmi di creazione di modelli disponibili in SPSS Modeler, per esempio le reti neurali, gli alberi decisionali, gli algoritmi cluster e la sequenzializzazione dei dati.
- **Modelli in-database.** Nodi che utilizzano gli algoritmi di modellazione disponibili nei database di Microsoft SQL Server, IBM DB2 Oracle.
- **Output.** Nodi che generano output di svariati tipi per dati, grafici e risultati di modelli, visualizzabili in SPSS Modeler.
- **Esporta.** Nodi che generano vari tipi di output, visualizzabili in applicazioni esterne quali IBM® SPSS® Data Collection o Excel.
- **SPSS Statistics.** Nodi utilizzati per importare o esportare dati da IBM® SPSS® Statistics e per eseguire le procedure di SPSS Statistics.

Una volta acquisita una certa dimestichezza con le funzioni di SPSS Modeler, sarà possibile personalizzare il contenuto della palette a piacere. [Per ulteriori informazioni, vedere l'argomento Personalizzazione della palette dei nodi in il capitolo 12 a pag. 237.](#)

Il riquadro Report, posto sotto la palette dei nodi, consente di controllare l'andamento delle varie operazioni, per esempio in fase di lettura dei dati nello stream di dati. Anche il riquadro dello stato si trova sotto la palette dei nodi. Esso fornisce informazioni sulle operazioni in corso nell'applicazione e offre indicazioni qualora sia richiesto il feedback dell'utente.

Manager di IBM SPSS Modeler

Il riquadro dei manager si trova nella parte superiore destra della finestra ed è formato da tre schede, utilizzate per gestire gli stream, l'output e i modelli.

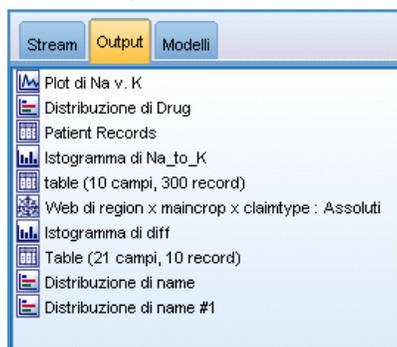
La scheda Stream può essere utilizzata per aprire, rinominare, salvare ed eliminare gli stream creati in una sessione.

Figura 3-7
Scheda Stream



La scheda Output contiene tutti i file creati (quali grafici e tabelle) tramite le operazioni su stream in IBM® SPSS® Modeler. I grafici, le tabelle e i report elencati in questa scheda possono essere visualizzati, salvati, rinominati e chiusi.

Figura 3-8
Scheda Output



La scheda Modelli è la scheda più importante della finestra dei manager. Essa contiene tutti gli **insiemi di modelli**, che contengono i modelli generati in SPSS Modeler per la sessione corrente. I modelli possono essere visualizzati direttamente dalla scheda Modelli oppure possono essere aggiunti allo stream nell'area di disegno.

Figura 3-9
Scheda Modelli contenente degli insiemi di modelli



Progetti di IBM SPSS Modeler

Nella parte inferiore destra della finestra viene visualizzato il riquadro dei progetti, utilizzato per creare e gestire **progetti** di data mining (gruppi di file relativi a un'attività di data mining). Sono disponibili due modalità di visualizzazione per i progetti creati in IBM® SPSS® Modeler—: le visualizzazioni Classi e CRISP-DM.

La scheda CRISP-DM consente di organizzare i progetti in base allo standard CRISP-DM (Cross-Industry Standard Process for Data Mining), una metodologia del settore consolidata e non proprietaria. L'utilizzo dello strumento CRISP-DM consente ai data miner esperti e principianti di organizzare e comunicare al meglio le operazioni eseguite.

Figura 3-10
visualizzazione CRISP-DM



La scheda Classi consente di organizzare il proprio lavoro in SPSS Modeler per categorie, in base ai tipi di oggetti creati. Si tratta di una visualizzazione utile per l'inventario di dati, stream e modelli.

Figura 3-11
visualizzazione Classi



IBM SPSS Modeler Barra degli strumenti

Nella parte superiore della finestra di IBM® SPSS® Modeler vi è una barra degli strumenti a icone che offre diverse funzioni utili. Di seguito vengono illustrati i pulsanti della barra degli strumenti e le relative funzioni.



Crea un nuovo stream



Apri stream



Salva stream



Stampa stream corrente



Taglia e sposta negli Appunti



Copia negli Appunti



Incollare la selezione



Annulla l'ultima operazione



Ripeti



Cerca nodi



Modifica proprietà dello stream



Anteprima generazione SQL

	Esegui stream corrente		Esegui selezione di stream
	Interrompi stream (attivato solo durante l'esecuzione dello stream)		Aggiunge un Supernodo
	Zoom avanti (solo Supernodi)		Zoom indietro (solo Supernodi)
	No markup in stream		Inserisci commento
	Nascondi markup stream (se presente)		Mostra markup dello stream nascosto
	Apri stream in IBM® SPSS® Modeler Advantage		

Il markup dello stream comprende i commenti agli stream, i collegamenti dei modelli e l'indicazione dei rami di calcolo del punteggio.

Per maggiori informazioni sui commenti agli stream, vedere [Aggiunta di commenti e annotazioni a nodi e stream a pag. 69](#).

Per maggiori informazioni sull'indicazione dei rami di calcolo del punteggio, vedere [Il ramo di calcolo del punteggio a pag. 189](#).

I collegamenti dei modelli sono descritti nel manuale *Nodi Modelli in IBM SPSS*.

Personalizzazione della barra degli strumenti

È possibile modificare vari aspetti della barra degli strumenti, per esempio:

- Se è visualizzata o meno
- Se viene visualizzato il testo descrittivo delle icone
- Se utilizzare le icone grandi o piccole

Per attivare o disattivare la visualizzazione della barra degli strumenti:

- ▶ Nel menu principale, fare clic su:
Visualizza > Barra degli strumenti > Visualizza

Per modificare l'impostazione del testo descrittivo o delle dimensioni delle icone:

- ▶ Nel menu principale, fare clic su:
Visualizza > Barra degli strumenti > Personalizza

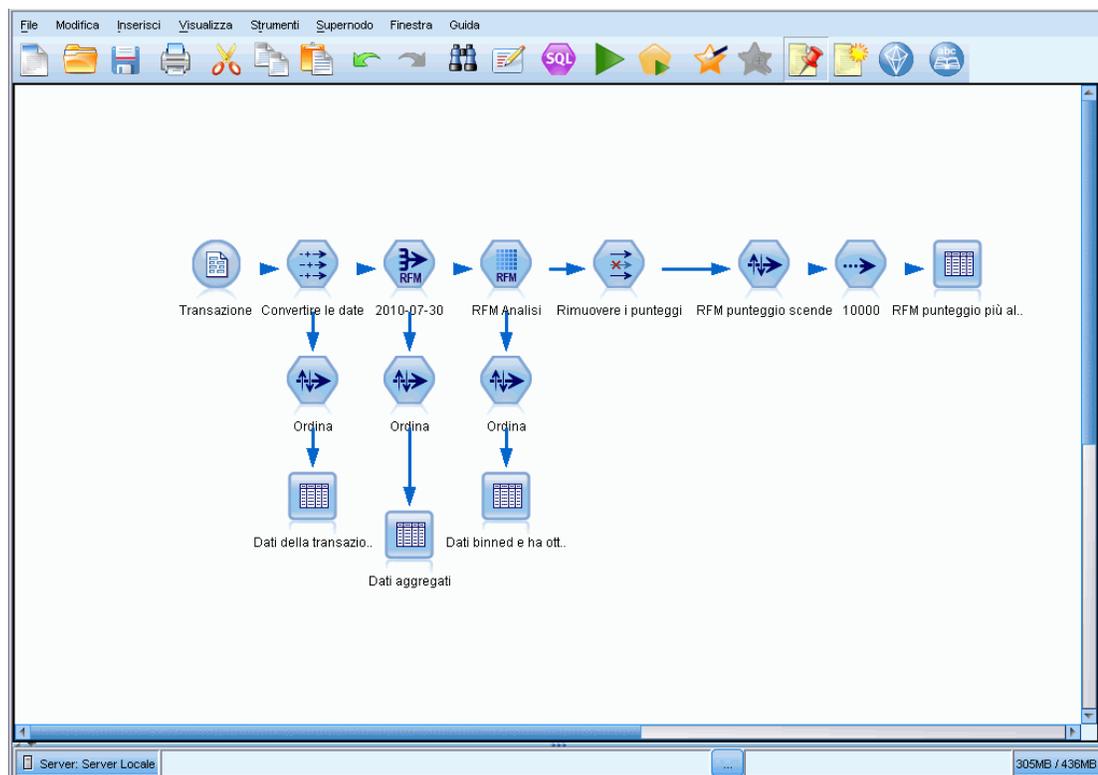
Fare clic su Mostra descrizioni o Pulsanti grandi a seconda dei casi.

Personalizzazione della finestra di IBM SPSS Modeler

I divisori tra le varie parti dell'interfaccia di IBM® SPSS® Modeler consentono di ridimensionare o chiudere a piacere gli strumenti. Per esempio, per uno stream di grandi dimensioni è possibile utilizzare le piccole frecce poste su ciascun divisore per chiudere la palette dei nodi, il riquadro dei manager e il riquadro dei progetti. In questo modo l'area di disegno viene ingrandita e consente di utilizzare più stream o stream di grandi dimensioni.

In alternativa, nel menu Visualizza fare clic su Palette nodi, Manager o Progetto per attivare o disattivare la visualizzazione di questi elementi.

Figura 3-12
Area di disegno dello stream ingrandita



In alternativa alla chiusura della palette dei nodi e dei riquadri dei manager e dei progetti è possibile utilizzare l'area di disegno come pagina scorrevole spostandosi in verticale e in orizzontale tramite le barre di scorrimento poste di fianco e in fondo alla finestra di SPSS Modeler.

È possibile anche controllare la visualizzazione del markup, costituito dai commenti agli stream, dai collegamenti dei modelli e dall'indicazione dei rami di calcolo del punteggio. Per attivare o disattivare la visualizzazione, fare clic su:

Visualizza > Markup stream

Utilizzo del mouse in IBM SPSS Modeler

Di seguito sono riportati gli utilizzi più comuni del mouse in IBM® SPSS® Modeler:

- **Clic singolo.** Utilizzare il pulsante destro o sinistro del mouse per scegliere opzioni dai menu, aprire menu di scelta rapida e accedere a numerosi altri controlli e opzioni standard. Fare clic e tenere premuto il pulsante per spostare e trascinare i nodi.
- **Doppio clic.** Fare doppio clic con il pulsante sinistro del mouse per posizionare i nodi nell'area di disegno dello stream e modificare i nodi esistenti.
- **Clic con il pulsante centrale.** Fare clic con il pulsante centrale del mouse e trascinare il cursore per connettere i nodi nell'area di disegno dello stream. Fare doppio clic con il pulsante centrale del mouse per disconnettere un nodo. Se non si dispone di un mouse con tre pulsanti, è possibile simulare questa funzionalità tenendo premuto Alt mentre si seleziona e si trascina l'elemento desiderato.

Utilizzo dei tasti di scelta rapida

In IBM® SPSS® Modeler, numerose operazioni di programmazione visuale sono associate a tasti di scelta rapida. Per esempio, per eliminare un nodo è possibile fare clic su di esso e premere il tasto Canc. Analogamente, per salvare rapidamente uno stream tenere premuto il tasto Ctrl e premere contemporaneamente il tasto S. I comandi di controllo di questo tipo sono indicati da una combinazione di Ctrl più un altro tasto, per esempio Ctrl+S.

Nelle operazioni standard di Windows sono utilizzati diversi tasti di scelta rapida, per esempio Ctrl+X per tagliare. Questi tasti di scelta rapida sono supportati in SPSS Modeler insieme ai tasti di scelta rapida seguenti specifici per l'applicazione.

Nota: in alcuni casi, i precedenti tasti di scelta rapida utilizzati in SPSS Modeler sono in conflitto con quelli standard di Windows. I precedenti tasti di scelta rapida sono supportati con l'aggiunta del tasto Alt. Per esempio, Ctrl+Alt+C può essere utilizzato per attivare/disattivare la cache.

Tabella 3-1
Tasti di scelta rapida supportati

Tasto di scelta rapida	Funzione
Ctrl+A	Seleziona tutto
Ctrl+X	Taglia
Ctrl+N	Nuovo stream

Tasto di scelta rapida	Funzione
Ctrl+O	Apri stream
Ctrl+P	Stampa
Ctrl+C	Copia
Ctrl+V	Incolla
Ctrl+Z	Annulla
Ctrl+Q	Seleziona tutti i nodi a valle del nodo selezionato
Ctrl+W	Deseleziona tutti i nodi a valle del nodo selezionato (si alterna con Ctrl+Q)
Ctrl+E	Esegui a partire dal nodo selezionato
Ctrl+S	Salva stream corrente
Alt+tasti freccia	Sposta i nodi selezionati nell'area di disegno dello stream in direzione della freccia utilizzata
Maiusc+F10	Apri il menu di scelta rapida per il nodo selezionato

Tabella 3-2

Scelte rapide supportate per i precedenti tasti di scelta rapida

Tasto di scelta rapida	Funzione
Ctrl+Alt+D	Duplica il nodo
Ctrl+Alt+L	Carica nodo
Ctrl+Alt+R	Rinomina il nodo
Ctrl+Alt+U	Crea il nodo input utente
Ctrl+Alt+C	Attiva/disattiva la cache
Ctrl+Alt+F	Svuota cache
Ctrl+Alt+X	Espandi Supernodo
Ctrl+Alt+Z	Zoom avanti/indietro
Canc	Elimina il nodo o la connessione

Stampa

In IBM® SPSS® Modeler è possibile stampare i seguenti oggetti:

- Diagrammi di stream
- Grafici
- Tabelle
- Report (dal nodo Report e dai report del progetto)
- Script (dalle finestre di dialogo Proprietà stream, Script locale o Script Supernodo)
- Modelli (browser Modello, schede di finestre di dialogo con stato attivo corrente, visualizzatori albero)
- Annotazioni (utilizzando la scheda Annotazioni per l'output)

Per stampare un oggetto:

- Per stampare senza anteprima, fare clic sul pulsante Stampa sulla barra degli strumenti.
- Per impostare la pagina prima della stampa, scegliere Imposta pagina dal menu File.
- Per richiamare l'anteprima prima della stampa, scegliere Anteprima di stampa dal menu File.
- Per visualizzare la finestra di dialogo di stampa standard con le opzioni di selezione delle stampanti e specificare le opzioni relative all'aspetto, scegliere Stampa dal menu File.

Automatizzazione di IBM SPSS Modeler

Il data mining può essere un processo complesso e talvolta lungo. Per questo motivo IBM® SPSS® Modeler comprende diversi tipi di supporto per la codifica e l'automazione.

- **Control Language for Expression Manipulation (CLEM)** è un linguaggio per l'analisi e la manipolazione dei flussi di dati negli stream di SPSS Modeler. I data miner fanno ampio ricorso al linguaggio CLEM nelle operazioni stream per eseguire attività sia semplici, come la derivazione dei profitti dai dati relativi ai costi e alle entrate, sia complesse, come la trasformazione dei dati dei registri Web in un insieme di campi e record contenenti informazioni utilizzabili. [Per ulteriori informazioni, vedere l'argomento Informazioni su CLEM in il capitolo 7 a pag. 97.](#)
- **Gli script** sono un potente strumento per automatizzare i processi nell'interfaccia utente. Possono eseguire lo stesso tipo di azioni compiute dagli utenti con un mouse o una tastiera. È possibile impostare le opzioni per i nodi ed eseguire creazioni di campi utilizzando un sottoinsieme di CLEM. Inoltre è possibile specificare l'output e manipolare i modelli generati. [Per ulteriori informazioni, vedere l'argomento Panoramica sugli script in il capitolo 2 in IBM SPSS Modeler 14.2 Guida per script e automazione.](#)

Nozioni sul data mining

Panoramica del data mining

Grazie a una vasta gamma di tecniche, il **data mining** consente di identificare insiemi di informazioni utili all'interno di corpi di dati. Con il data mining, le informazioni vengono estratte per essere utilizzate in settori quali il supporto decisionale e l'elaborazione di previsioni e stime. Ingenti quantità di dati si rivelano in genere di scarsa utilità quando si presentano sotto forma di dati non elaborati. Ad avere valore sono infatti le informazioni nascoste all'interno dei dati.

Nel data mining, il successo è garantito da una corretta combinazione della conoscenza dei dati di cui dispongono gli utenti o gli esperti aziendali con tecniche avanzate di analisi attiva, che consentono di individuare le relazioni e le caratteristiche sottostanti i dati. Mediante il processo di data mining vengono generati modelli basati su dati cronologici, che vengono successivamente utilizzati per l'elaborazione di previsioni, il rilevamento di schemi e così via. Le tecniche per la creazione di questi modelli vengono definite di **apprendimento automatico** o **modellazione**.

Tecniche di modellazione

In IBM® SPSS® Modeler sono disponibili numerose tecniche di apprendimento automatico e di modellazione, raggruppabili in base alla tipologia dei problemi che sono in grado di risolvere.

- I metodi di modellazione predittiva includono gli alberi decisionali, le reti neurali e i modelli statistici.
- I modelli di raggruppamento tramite cluster sono incentrati sull'identificazione di gruppi di record simili e sull'assegnazione di etichette ai record in base al gruppo di appartenenza. Tali metodi includono i metodi Kohonen, K-Means e TwoStep.
- Le regole di associazione consentono di associare una conclusione specifica, per esempio l'acquisto di un particolare prodotto, a un insieme di condizioni, come l'acquisto di numerosi altri prodotti.
- I modelli di screening possono essere utilizzati per sottoporre a screening i dati, in modo da individuare i campi e i record di maggior interesse ai fini della modellazione e identificare i valori anomali che potrebbero non rientrare in schemi già noti. I metodi disponibili includono la selezione delle funzioni e il rilevamento di anomalie.

Manipolazione e individuazione dei dati

SPSS Modeler include inoltre numerose funzionalità per l'applicazione delle proprie conoscenze ai dati:

- **Manipolazione dei dati.** Consente di creare nuovi elementi di dati derivanti da quelli esistenti e di suddividere i dati in sottoinsiemi significativi. È possibile unire e filtrare dati da numerose origini.

- **Esplorazione e visualizzazione.** È possibile visualizzare aspetti diversi dei dati utilizzando il nodo Esplora, che consente di analizzare i dati e visualizzare grafici e statistiche. Fra le opzioni avanzate vi è la possibilità di creare grafici interattivi, esportabili nei report dei progetti.
- **Statistiche.** Consentono di trovare conferma delle presunte relazioni tra le variabili all'interno dei dati. In SPSS Modeler possono essere utilizzate anche statistiche fornite da IBM® SPSS® Statistics.
- **Verifica delle ipotesi.** Consentono di creare modelli del comportamento dei dati e di verificare tali modelli.

Queste funzionalità vengono in genere utilizzate per identificare un insieme interessante di attributi all'interno dei dati. Questi attributi possono quindi essere utilizzati nell'ambito delle tecniche di creazione di modelli per identificare regole e relazioni sottostanti.

Applicazioni comuni

Tra le applicazioni più comuni delle tecniche di data mining sono incluse le seguenti:

Direct mailing. È possibile stabilire quali gruppi demografici presentano il tasso di risposta più elevato e utilizzare queste informazioni per massimizzare la risposta ai messaggi futuri.

Calcolo del merito creditizio. È possibile basare una decisione relativa a un credito sulla storia personale dei crediti di un singolo.

Risorse umane. È possibile analizzare le procedure di assunzione precedenti e creare regole decisionali che razionalizzano il processo di assunzione.

Ricerca medica. È possibile creare regole decisionali in grado di suggerire le procedure appropriate in base alla documentazione medica.

Analisi di mercato. È possibile stabilire quali sono le variabili associate alle vendite, per esempio gli aspetti geografici, il prezzo e le caratteristiche dei clienti.

Controllo della qualità. È possibile analizzare i dati a partire dalla produzione e identificare le variabili che determinano i difetti dei prodotti.

Studio di linee guida. È possibile utilizzare i dati dei sondaggi per formulare linee guida applicando regole decisionali per selezionare le variabili più importanti.

Servizi sanitari. È possibile combinare le indagini e i dati clinici per individuare le variabili che contribuiscono allo stato di buona salute.

Terminologia

I termini **attributo**, **campo** e **variabile** fanno riferimento a un singolo elemento di dati comune a tutti i casi considerati. Un insieme di valori di attributo che fa riferimento a un caso specifico viene definito **record**, **esempio** o **caso**.

Valutazione dei dati

Il data mining si rivela utile solo se i dati che si desidera utilizzare soddisfano criteri specifici. Nelle sezioni riportate di seguito vengono illustrati alcuni aspetti dei dati e della loro applicazione che è opportuno considerare.

Assicurarsi che i dati siano disponibili

Sebbene questo possa apparire ovvio, è importante tenere in considerazione che i dati eventualmente disponibili potrebbero non essere in un formato facilmente utilizzabile. IBM® SPSS® Modeler supporta l'importazione di dati da database (tramite ODBC) o da file. È tuttavia possibile che i dati siano memorizzati in un altro formato su un computer non accessibile direttamente. In questo caso, prima dell'utilizzo è necessario eseguirne il download o il dump in un formato appropriato. Se sono distribuiti in origini e database diversi, i dati dovranno inoltre essere raccolti. È infine possibile che non siano disponibili in linea. Se esistono solo in formato cartaceo, prima di avviare il data mining è necessario immettere i dati.

Controllare se i dati coprono gli attributi pertinenti

Poiché l'obiettivo del data mining consiste nell'identificazione di attributi pertinenti, includere questo controllo può apparire singolare. È tuttavia estremamente utile esaminare i dati disponibili e tentare di identificare i fattori probabilmente pertinenti non registrati. Se si tenta di prevedere le vendite di gelati, per esempio, è possibile che si disponga di ingenti quantità di informazioni sui punti vendita o sulla cronologia delle vendite, ma non di informazioni meteorologiche o sulle temperature, che avranno probabilmente un ruolo significativo. Sebbene non impediscano di ottenere risultati utili con il data mining, gli attributi mancanti possono limitare la precisione delle risultanti previsioni.

Un modo rapido per valutare la situazione consiste nell'eseguire un controllo completo dei dati. Prima di proseguire, è opportuno valutare l'opportunità di associare un nodo Esplora all'origine dati e generare un report completo. [Per ulteriori informazioni, vedere l'argomento Nodo Esplora in il capitolo 6 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)

Attenzione ai dati rumorosi

I dati contengono spesso errori o possono comprendere giudizi soggettivi e pertanto variabili. Questi fenomeni vengono definiti collettivamente **rumore**. La presenza di rumore nei dati è talvolta normale. Le regole sottostanti non possono infatti controllare il 100% dei casi.

In genere, la quantità di rumore presente nei dati è direttamente proporzionale alla difficoltà di ottenere risultati precisi. I metodi di apprendimento automatico di SPSS Modeler, tuttavia, sono in grado di gestire dati di questo tipo e hanno fornito esito positivo su insiemi di dati con quasi il 50% di rumore.

Assicurarsi che i dati a disposizione siano sufficienti

Nell'ambito del data mining, le dimensioni di un insieme di dati non rivestono necessariamente un ruolo importante e risulta nettamente più significativa la *rappresentatività* dell'insieme di dati, così come la copertura dei risultati e delle combinazioni di variabili possibili.

In genere, per fornire una copertura rappresentativa sono necessari tanti più record quanti più attributi vengono presi in considerazione.

In presenza di dati rappresentativi e di regole sottostanti generali, è possibile che un campione di dati di alcune migliaia o addirittura centinaia di record fornisca risultati altrettanto validi di quelli garantiti da un milione di record, restituendo i risultati con maggiore rapidità.

Confrontarsi con gli esperti sui dati

In numerosi casi si utilizzano dati propri e si dispone quindi di un'elevata familiarità con il relativo contenuto e significato. Se tuttavia si gestiscono dati per un altro reparto aziendale o per un cliente, è preferibile potersi avvalere di esperti con una conoscenza approfondita dei dati, che possano semplificare l'identificazione degli attributi significativi e l'interpretazione dei risultati del data mining, distinguendo gli effettivi nuclei di informazioni da dati falsati a causa di anomalie negli insiemi di dati.

Strategia per il data mining

Come la maggior parte delle attività aziendali, il data mining si rivela più efficiente se eseguito in modo sistematico e pianificato. Anche con strumenti di data mining all'avanguardia come IBM® SPSS® Modeler, per la maggior parte delle operazioni di data mining è richiesto un analista aziendale competente, in grado di gestire correttamente il processo. Per completare la pianificazione è necessario individuare quanto segue:

- Il problema di base che si desidera risolvere
- Le origini dati disponibili e le parti dei dati pertinenti per il problema corrente
- Le operazioni di pre-elaborazione e di pulitura dei dati da eseguire prima dell'avvio del data mining
- Le tecniche di data mining da utilizzare
- La modalità di valutazione dei risultati dell'analisi di data mining
- Le possibilità di utilizzo delle informazioni ottenute dal data mining

Un tipico processo di data mining può spesso risultare complicato. È infatti necessario tenere traccia di complessi problemi aziendali, di più origini dati, della diversa qualità dei dati nelle diverse origini, di numerose tecniche di data mining, di metodi differenti di valutazione dell'esito del data mining e così via.

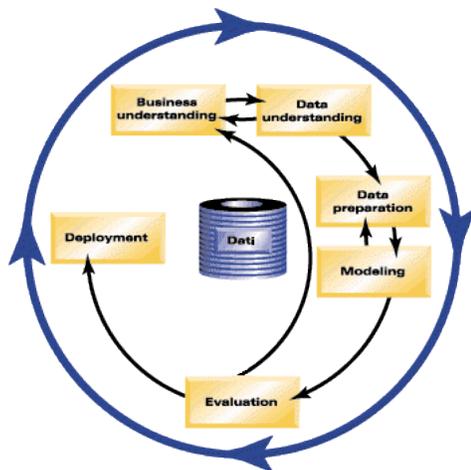
La gestione di tutti questi aspetti può essere semplificata da un modello di elaborazione esplicitamente definito per il data mining. Il modello di elaborazione consente di rispondere alle domande elencate in precedenza in questa sezione e di assicurarsi che siano affrontati i punti importanti. Inoltre, tale modello guida l'utente nel processo di data mining in modo da superare la complessità dei dati.

Il processo di data mining consigliato per l'uso con SPSS Modeler è rappresentato dal modello CRISP-DM (Cross-Industry Standard Process for Data Mining), progettato come modello generale applicabile a una vasta gamma di settori e problemi aziendali.

Modello di elaborazione CRISP-DM

Il modello di processo generale CRISP-DM comprende sei fasi relative alle principali problematiche del data mining. Le sei fasi costituiscono complessivamente un processo ciclico, studiato per integrare il data mining nell'attività aziendale globale.

Figura 4-1
modello di elaborazione CRISP-DM



Di seguito vengono descritte queste sei fasi.

- **Business Understanding.** Costituisce probabilmente la fase più importante del data mining. Comprende l'individuazione degli obiettivi aziendali, la valutazione della situazione, la definizione degli scopi del data mining e la creazione di un piano di progetto.
- **Data Understanding.** I dati costituiscono la "materia prima" del data mining. Questa fase, in cui viene soddisfatta l'esigenza di analizzare le risorse di dati disponibili e le caratteristiche di tali risorse, comprende la raccolta dei dati iniziali, la descrizione e l'esplorazione dei dati e la verifica della relativa qualità. Il nodo Esplora, disponibile nella palette di nodi Output, è uno strumento indispensabile per l'analisi dei dati.
- **Data Preparation.** Dopo la catalogazione delle risorse di dati è necessario preparare i dati per il data mining. Le operazioni di preparazione comprendono la selezione, la pulitura, la creazione, l'integrazione e la formattazione.
- **Creazione di modelli.** Costituisce la parte più accattivante del data mining, in cui vengono utilizzati avanzati metodi di analisi per estrarre le informazioni dai dati. Questa fase comprende la selezione di tecniche di creazione di modelli, la generazione di strutture di test e la creazione e la valutazione di modelli.
- **Valutazione.** Una volta selezionati i modelli, è possibile valutare i vantaggi che può offrire il data mining ai fini del raggiungimento degli obiettivi aziendali. Gli elementi costitutivi di questa fase comprendono la valutazione dei risultati, la revisione del processo di data mining e l'individuazione dei passaggi successivi.
- **Deployment.** A questo punto è possibile usufruire dei vantaggi. Questa fase è incentrata sull'integrazione delle nuove conoscenze nei comuni processi aziendali allo scopo di risolvere il problema originale. Comprende la distribuzione, il monitoraggio e la gestione del piano, la creazione di un report finale e la revisione del progetto.

Questo modello di elaborazione è caratterizzato da alcuni punti principali. Innanzitutto, sebbene generalmente le fasi del processo seguano l'ordine illustrato nei paragrafi precedenti, esistono vari punti in cui si influenzano reciprocamente in modo non lineare. La fase di preparazione dei dati, per esempio, precede in genere la fase di modellazione. Tuttavia, le decisioni definite e le informazioni raccolte durante la fase di modellazione possono spesso indurre a riconsiderare alcune parti della fase di preparazione dei dati, che possono a loro volta presentare nuove problematiche a livello di modellazione. Le due fasi continuano a influenzarsi reciprocamente finché entrambe non vengono risolte in modo appropriato. Analogamente, la fase Evaluation può determinare una rivalutazione della fase Business Understanding iniziale, e questo potrebbe portare alla scoperta che si sta tentando di ottenere una risposta alla domanda sbagliata. A questo punto, si può rivedere la fase Business Understanding e proseguire nuovamente con il resto del processo con un obiettivo più appropriato.

Il secondo punto principale è costituito dal carattere iterativo del data mining. Raramente il processo si esaurisce con la pianificazione di un progetto di data mining, il relativo completamento e l'ottenimento dei dati. L'utilizzo del data mining per soddisfare le richieste dei clienti costituisce un'attività costante. Le conoscenze acquisite con un ciclo di data mining generano solitamente nuove domande, nuove problematiche e nuove opportunità di individuare e soddisfare le esigenze dei clienti. Queste nuove domande, problematiche e opportunità possono in genere essere gestite con un nuovo processo di data mining. Questo processo di estrazione e identificazione delle nuove opportunità deve diventare parte dell'approccio allo svolgimento delle attività aziendali e un elemento di base della strategia aziendale complessiva.

Nella presente introduzione viene fornita solo una breve panoramica del modello di elaborazione CRISP-DM. Per informazioni dettagliate sull'utilizzo del modello, consultare le seguenti risorse:

- La *CRISP-DM Guide*, accessibile insieme ad altra documentazione dalla cartella *Documentation* nel disco di installazione.
- La Guida in linea di CRISP-DM, accessibile dal menu Start o facendo clic su Guida in linea di CRISP-DM nel menu Aiuto di IBM® SPSS® Modeler.

Tipi di modelli

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.

Il documento *SPSS Modeler Guida alle applicazioni* fornisce numerosi esempi di questi metodi, oltre a un'introduzione generale sul processo di modellazione. Tale guida è disponibile come esercitazione online e in formato PDF. [Per ulteriori informazioni, vedere l'argomento Esempi di applicazioni in il capitolo 1 a pag. 4.](#)

I metodi di modellazione sono suddivisi in tre categorie:

- Classificazione
- Associazione
- Segmentazione

Modelli Classificazione

I *Modelli Classificazione* utilizzano i valori di uno o più campi di **input** per prevedere il valore di uno o più campi di output o **obiettivo**. Alcuni esempi di queste tecniche sono: alberi decisionali (C&R Tree, QUEST, CHAID e algoritmi C5.0), regressione (lineare, logistica, lineare generalizzata e algoritmo di regressione di Cox), reti neurali, SVM (support vector machines) e reti bayesiane.

I modelli Classificazione consentono di prevedere un risultato noto, per esempio se un cliente effettuerà l'acquisto o vi rinuncerà oppure se una transazione rientra in un modello noto di comportamento fraudolento. Le tecniche di modellazione includono l'apprendimento automatico, l'induzione di regole, l'identificazione dei sottogruppi, i metodi statistici e la creazione di più modelli.

Nodi di classificazione



Il nodo Classificatore automatico crea e confronta svariati tipi di modelli per risultati binari (sì o no, abbandono oppure no e così via), consentendo di scegliere l'approccio migliore per una determinata analisi. Sono supportati numerosi algoritmi di modellazione ed è possibile selezionare i metodi da utilizzare, le opzioni specifiche per ognuno di essi e i criteri per confrontare i risultati. Il nodo genera un insieme di modelli basati sulle opzioni specificate e classifica i candidati migliori in base ai criteri indicati. [Per ulteriori informazioni, vedere l'argomento Nodo Classificatore automatico in il capitolo 5 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il nodo Numerico automatico stima e confronta i modelli per i risultati di intervalli numerici continui utilizzando svariati metodi. Il nodo funziona in modo analogo al nodo Classificatore automatico e consente di scegliere gli algoritmi da utilizzare e di sperimentare più combinazioni di opzioni in un singolo passaggio di modellazione. Gli algoritmi supportati includono reti neurali, C&R Tree, CHAID, regressione lineare, regressione lineare generalizzata e SVM (Support Vector Machine). I modelli si possono confrontare in base a correlazione, errore relativo o numero di variabili utilizzato. [Per ulteriori informazioni, vedere l'argomento Nodo Numerico automatico in il capitolo 5 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il nodo Albero di classificazione e regressione (C&R) genera un albero decisionale che consente di prevedere o classificare osservazioni future. Il metodo utilizza l'esecuzione ricorsiva di partizioni per suddividere i record di addestramento in segmenti riducendo l'impurità ad ogni passaggio. Un nodo dell'albero è considerato "puro" quando il 100% dei casi nel nodo fa parte di una categoria specifica del campo obiettivo. I campi obiettivo e di input possono essere intervalli numerici o categoriali (nominali, ordinali o flag); tutte le suddivisioni sono binarie (solo due sottogruppi). [Per ulteriori informazioni, vedere l'argomento Nodo C&R Tree in il capitolo 6 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il nodo QUEST offre un metodo di classificazione binario per la creazione di alberi decisionali, progettato per ridurre i tempi di elaborazione necessari per le analisi C&R Tree più complesse, riducendo inoltre la tendenza dei metodi per gli alberi di classificazione a favorire gli input che consentono un numero maggiore di suddivisioni. I campi di input possono essere intervalli numerici (continui), ma il campo obiettivo deve essere categoriale. Tutte le suddivisioni sono binarie. [Per ulteriori informazioni, vedere l'argomento Nodo QUEST in il capitolo 6 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il nodo CHAID genera un albero decisionale utilizzando statistiche chi-quadrato per identificare suddivisioni ottimali. A differenza dei nodi C&R Tree e QUEST, CHAID può generare alberi non binari e pertanto alcune suddivisioni possono avere più di due rami. I campi obiettivo e di input possono essere intervallo numerico (continui) o categoriali. Un CHAID completo è una modificazione di CHAID che esegue operazioni avanzate per l'analisi di tutte le suddivisioni possibili, ma richiede tempi di elaborazione maggiori. [Per ulteriori informazioni, vedere l'argomento Nodo CHAID in il capitolo 6 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il nodo C5.0 crea un albero decisionale o un insieme di regole. Il modello suddivide il campione in base al campo che fornisce il massimo guadagno di informazioni a ogni livello. Il campo obiettivo deve essere categoriale. Sono consentite suddivisioni multiple in più di due sottogruppi. [Per ulteriori informazioni, vedere l'argomento nodo C5.0 in il capitolo 6 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il nodo Elenco decisionale identifica i sottogruppi o i segmenti che mostrano una probabilità maggiore o minore che si verifichi un determinato risultato binario rispetto alla popolazione globale. Per esempio, è possibile che si cerchino i clienti non a rischio di abbandono o quelli che più probabilmente rispondano in modo favorevole a una campagna. È possibile incorporare le proprie conoscenze aziendali nel modello aggiungendo propri segmenti personalizzati e visualizzando in anteprima modelli alternativi uno accanto all'altro per confrontarne i risultati. I modelli Elenco decisionale consistono in un elenco di regole in cui ogni regola ha una condizione e un risultato. Le regole vengono applicate in ordine e la prima regola corrispondente determina il risultato. [Per ulteriori informazioni, vedere l'argomento Elenco decisionale in il capitolo 9 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



I modelli di regressione lineare prevedono un obiettivo continuo basato sulle relazioni lineari tra l'obiettivo e uno o più predittori. [Per ulteriori informazioni, vedere l'argomento Modelli lineari in il capitolo 10 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



La regressione lineare è una tecnica statistica molto comune per riassumere i dati ed eseguire previsioni individuando un'area o una linea retta in grado di ridurre le discrepanze tra i valori di output previsti e quelli osservati.



Il nodo Fattoriale offre potenti tecniche di riduzione dei dati che consentono di diminuirne la complessità. L'analisi dei componenti principali (PCA, Principal Components Analysis) trova le combinazioni lineari dei campi di input che catturano meglio la varianza nell'intero insieme di campi, dove i componenti sono ortogonali (perpendicolari) l'uno rispetto all'altro. L'analisi fattoriale tenta di identificare i concetti sottostanti, o fattori, che spiegano lo schema delle correlazioni all'interno dell'insieme di campi osservati. Entrambi gli approcci mirano a trovare un numero ridotto di campi derivati che riassumono in modo efficace le informazioni presenti nell'insieme originale di campi. [Per ulteriori informazioni, vedere l'argomento Nodo Fattoriale in il capitolo 10 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il nodo Selezione funzioni effettua lo screening dei campi di input, rimuovendoli in base a un insieme di criteri quali la percentuale di valori mancanti. Classifica quindi gli input restanti in ordine di importanza rispetto a un determinato obiettivo. Per esempio, dato un insieme di dati con centinaia di input potenziali, quali sono quelli con la maggiore probabilità di essere utili nella modellazione di risultati clinici? [Per ulteriori informazioni, vedere l'argomento Nodo Selezione funzioni in il capitolo 4 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



L'analisi discriminante prevede presupposti più rigidi rispetto alla regressione logistica, ma può essere una valida alternativa o un complemento dell'analisi di regressione logistica quando vengono soddisfatti tali presupposti. [Per ulteriori informazioni, vedere l'argomento Nodo Discriminante in il capitolo 10 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



La regressione logistica, una tecnica statistica che consente di classificare i record in base ai valori dei campi di input, è analoga alla regressione lineare ma, al posto di un intervallo numerico, prende un campo obiettivo categoriale. [Per ulteriori informazioni, vedere l'argomento Nodo Logistica in il capitolo 10 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il modello Lineare generalizzato amplia il modello lineare generale in modo che la variabile dipendente venga linearmente correlata ai fattori e alle covariate tramite una funzione di collegamento specifica. Inoltre, il modello consente alla variabile dipendente di avere una distribuzione non normale. Copre la funzionalità di un grande numero di modelli statistici, inclusi modelli di regressione lineare, modelli di regressione logistica, modelli loglineari per dati dei conteggi e modelli di sopravvivenza censurati per intervallo. [Per ulteriori informazioni, vedere l'argomento Nodo GenLin in il capitolo 10 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il nodo Regressione di Cox consente di generare un modello di sopravvivenza per i dati della relazione tempo-evento in presenza di record censurati. Il modello produce una funzione di sopravvivenza che prevede la probabilità che l'evento di interesse si sia verificato a una determinata ora (t) per i valori dati delle variabili di input. [Per ulteriori informazioni, vedere l'argomento Nodo Cox in il capitolo 10 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il nodo SVM (Support Vector Machine) consente di classificare i dati in uno di due gruppi senza sovradattamento. Il nodo SVM è particolarmente indicato per l'utilizzo con insiemi di dati di grandi dimensioni, cioè quelli con un elevato numero di campi di input. [Per ulteriori informazioni, vedere l'argomento Nodo SVM in il capitolo 15 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il nodo Rete bayesiana consente di generare un modello di probabilità combinando elementi osservati e registrati con conoscenze del mondo reale per stabilire la probabilità di occorrenze. Il nodo si concentra sulle reti TAN (Tree Augmented Naïve Bayes) e Markov Blanket, che sono prevalentemente utilizzate a scopo di classificazione. [Per ulteriori informazioni, vedere l'argomento Nodo Rete bayesiana in il capitolo 7 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il nodo Modello risposta autoapprendimento consente di creare un modello in cui è possibile utilizzare un unico nuovo caso oppure un numero limitato di nuovi casi per eseguire una nuova stima del modello senza doverlo riaddestrare con tutti i dati. [Per ulteriori informazioni, vedere l'argomento Nodo SLRM in il capitolo 14 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il nodo Serie storica stima i modelli di livellamento esponenziale, i modelli ARIMA (Autoregressive Integrated Moving Average, autoregressivi integrati a media mobile) univariati e ARIMA (o a funzione di trasferimento) multivariati per i dati di serie storica e genera previsioni di prestazioni future. Il nodo Serie storica deve sempre essere preceduto da un nodo Intervalli di tempo. [Per ulteriori informazioni, vedere l'argomento Nodo Modelli Serie storica in il capitolo 13 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il nodo Vicini più vicini K (KNN) associa un nuovo caso alla categoria o valore degli oggetti K più vicini ad esso nello spazio di predittori, dove K è un numero intero. I casi simili sono vicini gli uni agli altri, mentre i casi dissimili sono distanti gli uni dagli altri. [Per ulteriori informazioni, vedere l'argomento Nodo KNN in il capitolo 16 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)

Modelli Associazione

I *Modelli Associazione* trovano nei dati gli schemi nei quali una o più entità (eventi, acquisti o attributi) vengono associate a un'altra o a diverse altre entità. I modelli creano insiemi di regole che definiscono queste relazioni. I campi all'interno dei dati possono essere sia di input che obiettivo. Queste associazioni possono essere trovate manualmente, ma gli algoritmi delle regole di associazione funzionano più velocemente e possono esplorare schemi più complessi. I modelli Apriori e Carma sono esempi dell'utilizzo di questi algoritmi. Un altro tipo di modello di associazione è un modello di rilevamento di sequenze, che trova gli schemi sequenziali nei dati a struttura temporale.

I modelli Associazione sono utili soprattutto nella previsione di risultati multipli; per esempio, i clienti che hanno acquistato il prodotto X hanno acquistato anche Y e Z. I modelli Associazione associano una particolare conclusione (per esempio la decisione di acquistare qualcosa) a un insieme di condizioni. Rispetto agli algoritmi dell'albero decisionale più tradizionali (C5.0 e C&R tree), quelli delle regole di associazione presentano il vantaggio di poter definire associazioni tra qualsiasi tipo di attributo. Mentre un algoritmo di albero decisionale genera regole con un'unica conclusione, gli algoritmi di associazione tentano di individuare più regole, ciascuna delle quali può fornire una diversa conclusione.

Nodi di associazione



Il nodo Apriori estrae un insieme di regole dai dati, estrapolando le regole con il più alto contenuto di informazioni. Apriori offre cinque diversi metodi per la selezione delle regole e utilizza uno schema di indicizzazione sofisticato per elaborare in modo efficiente insiemi di dati di grandi dimensioni. In caso di problemi complessi, l'addestramento di Apriori è in genere più rapido. Apriori non ha un limite arbitrario per quanto riguarda il numero di regole che possono essere mantenute e può gestire regole con un massimo di 32 precondizioni. Apriori richiede che tutti i campi di input e output siano categoriali ma garantisce prestazioni migliori perché è ottimizzato per questo tipo di dati. [Per ulteriori informazioni, vedere l'argomento Nodo Apriori in il capitolo 12 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il modello CARMA estrae un insieme di regole dai dati senza che venga richiesto all'utente di specificare i campi di input o obiettivo. A differenza di Apriori il nodo CARMA fornisce le impostazioni di creazione per il supporto delle regole (sia per l'antecedente che per il conseguente) anziché solo per il supporto antecedente. Pertanto, le regole generate possono essere utilizzate per una gamma più vasta di applicazioni, per esempio per trovare un elenco di prodotti o di servizi (antecedenti) il cui conseguente è rappresentato dall'articolo che si desidera promuovere per le festività correnti. [Per ulteriori informazioni, vedere l'argomento Nodo CARMA in il capitolo 12 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il nodo Sequenza consente di scoprire le regole di associazione nei dati sequenziali o basati su valori temporali. Per sequenza si intende un elenco di insiemi di elementi che tendono a ricorrere secondo un ordine prevedibile. Ad esempio, un cliente che acquista un rasoio e la lozione dopobarba potrebbe in seguito acquistare la schiuma da barba. Il nodo Sequenza si basa sull'algorithmo delle regole di associazione CARMA, che utilizza un metodo efficiente in due passaggi per trovare le sequenze. [Per ulteriori informazioni, vedere l'argomento Nodo Sequenza in il capitolo 12 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)

Modelli Segmentazione

I *Modelli Segmentazione* suddividono i dati in segmenti, o cluster, di record con schemi simili di campi di input. I modelli di segmentazione si concentrano soltanto sui campi di input e quindi non tengono in considerazione i campi di output o obiettivo. Tra gli esempi di modelli di segmentazione citiamo le reti di Kohonen, il raggruppamento in cluster K-Means, il raggruppamento in cluster TwoStep e il rilevamento delle anomalie.

I modelli Segmentazione (denominati anche “modelli di cluster”) sono utili nei casi in cui il risultato specifico non è conosciuto, per esempio nell'identificazione di nuovi modelli di comportamento fraudolento oppure nell'identificazione di gruppi di interesse nella base clienti. I modelli di raggruppamento tramite cluster sono incentrati sull'identificazione di gruppi di record simili e sull'assegnazione di etichette ai record in base al gruppo di appartenenza. Ciò avviene senza che siano necessarie preve conoscenze relative ai gruppi e alle loro caratteristiche e distingue i modelli di raggruppamento tramite cluster da altre tecniche di modellazione - non esiste un campo di destinazione o un output predefinito per il modello da prevedere. Per questi modelli non esistono risposte corrette o errate. Il valore è determinato dalla capacità di acquisire gruppi significativi all'interno dei dati e di fornire descrizioni utili di tali raggruppamenti. Spesso i modelli di raggruppamento tramite cluster vengono utilizzati per creare cluster o segmenti poi

utilizzati come input nelle analisi successive, per esempio segmentando i potenziali clienti in sottogruppi omogenei.

Nodi di segmentazione



Il nodo Cluster automatico stima e confronta i modelli di cluster che identificano gruppi di record con caratteristiche simili. Il nodo funziona in modo analogo ad altri nodi Modelli automatici e consente di sperimentare varie combinazioni di opzioni in un singolo passaggio di modellazione. I modelli si possono confrontare utilizzando misure di base con cui tentare di filtrare e classificare l'utilità dei modelli di cluster e fornire una misura in base all'importanza di determinati campi. [Per ulteriori informazioni, vedere l'argomento Nodo Cluster automatico in il capitolo 5 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il nodo K-Means raggruppa l'insieme di dati in gruppi distinti (o cluster). Il metodo definisce un numero fisso di cluster, esegue un'assegnazione iterativa dei record ai cluster e modifica i centri dei cluster finché un'ulteriore ridefinizione non consente più un miglioramento del modello. Invece di tentare di prevedere un risultato, il nodo K-Means utilizza un processo denominato apprendimento non supervisionato per scoprire gli schemi nell'insieme di campi di input. [Per ulteriori informazioni, vedere l'argomento Nodo K-Means in il capitolo 11 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il nodo Kohonen genera un tipo di rete neurale che può essere utilizzato per raggruppare l'insieme di dati in gruppi distinti. Al termine dell'apprendimento della rete, i record analoghi dovranno essere vicini nella mappa di output, mentre i record diversi saranno a notevole distanza. Per identificare le unità forti, è possibile controllare il numero di osservazioni catturate da ciascuna unità nell'insieme di modelli. In questo modo è possibile avere un'idea del numero appropriato di cluster. [Per ulteriori informazioni, vedere l'argomento Nodo Kohonen in il capitolo 11 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il nodo TwoStep è un metodo di raggruppamento tramite cluster in due fasi. La prima fase esegue un singolo passaggio nei dati per comprimere i dati di input non elaborati in un insieme gestibile di sottocluster. Nella seconda fase viene utilizzato un metodo di raggruppamento tramite cluster gerarchico per unire progressivamente i sottocluster in cluster sempre più grandi. Il nodo TwoStep offre il vantaggio di stimare automaticamente il numero ottimale di cluster per i dati di addestramento. Può gestire in modo efficiente tipi di campo misti e insiemi di dati di grandi dimensioni. [Per ulteriori informazioni, vedere l'argomento Nodo Cluster TwoStep in il capitolo 11 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)



Il nodo Rilevamento anomalie identifica casi insoliti, o valori anomali, non conformi a schemi di dati "normali". Con questo nodo è possibile identificare valori anomali anche se questi non rientrano in schemi precedentemente conosciuti e anche se l'utente non sa esattamente ciò che sta cercando. [Per ulteriori informazioni, vedere l'argomento Nodo Rilevamento anomalie in il capitolo 4 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)

Modelli Mining in-database

SPSS Modeler supporta l'integrazione con gli strumenti di data mining e di modellazione forniti da sviluppatori di database, quali Oracle Data Miner, IBM DB2 InfoSphere Warehouse e Microsoft Analysis Services. Operando all'interno dell'applicazione SPSS Modeler è infatti possibile sia creare modelli che calcolarne il punteggio e archivarli nel database. Per ulteriori

informazioni, consultare la *SPSS Modeler Guida alla modellazione in-database* disponibile nel DVD del prodotto.

IBM SPSS Statistics Models

Se si dispone di una copia di IBM® SPSS® Statistics installata con licenza sul proprio computer, è possibile accedere a determinate routine SPSS Statistics ed eseguirle dall'interno di SPSS Modeler per creare e determinare il punteggio dei modelli. [Per ulteriori informazioni, vedere l'argomento Nodi IBM SPSS Statistics - Panoramica in il capitolo 8 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)

Ulteriori informazioni

È inoltre disponibile una documentazione esaustiva sugli algoritmi di modellazione. Per ulteriori informazioni, vedere la *SPSS Modeler Algorithms Guide* disponibile sul DVD del prodotto.

Esempi di Oracle Data Mining

Il metodo migliore per apprendere i concetti pratici del data mining consiste nel partire da un esempio. Numerosi esempi di applicazioni sono disponibili in *IBM® SPSS® Modeler Guida alle applicazioni*, che fornisce brevi indicazioni mirate su specifici metodi e tecniche di modellazione. [Per ulteriori informazioni, vedere l'argomento Esempi di applicazioni in il capitolo 1 a pag. 4.](#)

Creazione di stream

Panoramica sulla creazione degli stream

Il data mining di IBM® SPSS® Modeler è incentrato sul processo di elaborazione dei dati attraverso una serie di nodi, denominato **stream**. La serie di nodi rappresenta le operazioni che verranno eseguite sui dati, mentre i collegamenti tra i nodi indicano la direzione del flusso di dati. Con uno stream di dati, è in genere possibile leggere i dati di SPSS Modeler, eseguire una serie di manipolazioni e quindi inviarli a una destinazione, per esempio a una tabella o a un visualizzatore.

Si supponga, per esempio, che sia necessario aprire una sorgente dati, aggiungere un nuovo campo, selezionare i record in base ai valori del nuovo campo e quindi visualizzare i risultati in una tabella. In questo caso, lo stream di dati includerà quattro nodi:



Un nodo Testo File, che è possibile impostare per leggere i dati dalla sorgente dati.



Un nodo Nuovo campo, mediante il quale è possibile aggiungere il nuovo campo calcolato all'insieme di dati.



Un nodo Seleziona, che consente di impostare i criteri di selezione per l'esclusione di record dallo stream di dati.



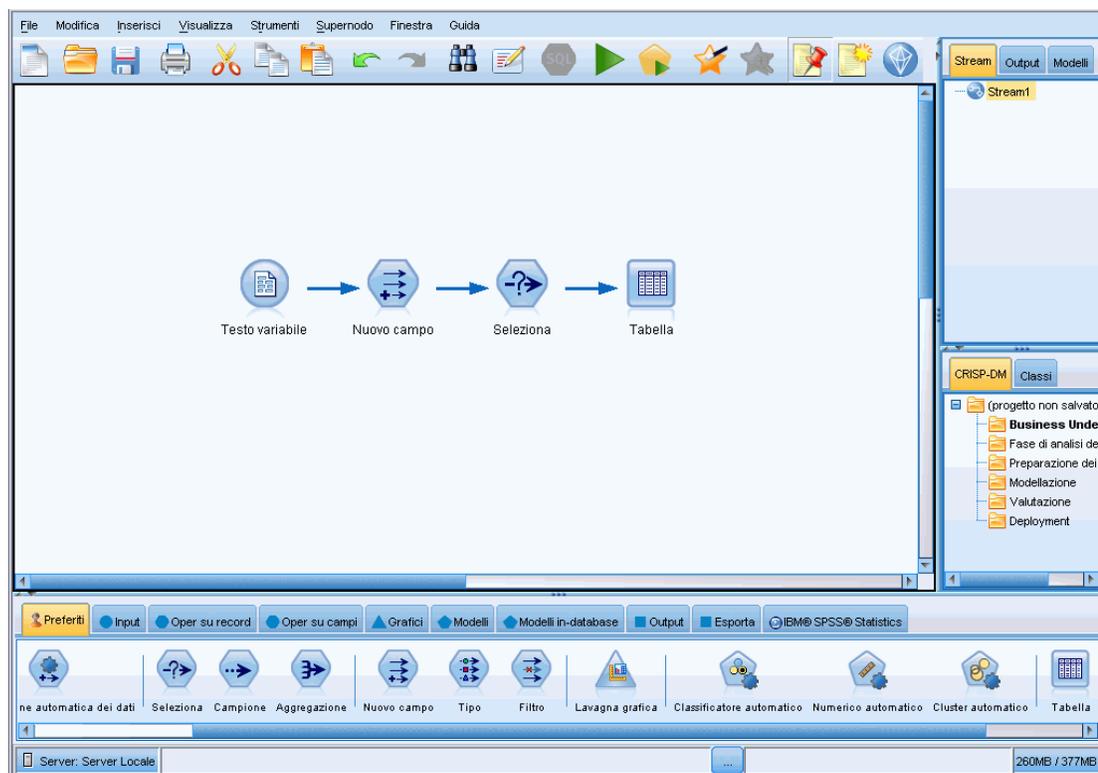
Un nodo Tabella, che consente di visualizzare sullo schermo i risultati delle manipolazioni.

Creazione degli stream di dati

L'interfaccia di IBM® SPSS® Modeler consente di eseguire il data mining visuale grazie all'utilizzo dei diagrammi degli stream di dati. La creazione degli stream di dati prevede i seguenti passaggi fondamentali:

- Aggiunta di nodi nell'area di disegno dello stream.
- Connessione dei nodi per creare uno stream.
- Definizione delle opzioni dei nodi o dello stream.
- Esecuzione dello stream.

Figura 5-1
Stream completato nell'area di disegno dello stream



In questa sezione vengono fornite informazioni dettagliate sull'utilizzo dei nodi per la creazione di stream di dati più complessi e vengono inoltre illustrate le opzioni e le impostazioni relative ai nodi e agli stream. Per esempi di procedure dettagliate relative alla creazione di stream mediante i dati disponibili in SPSS Modeler (cartella *Demos* del programma di installazione), vedere [Esempi di applicazioni a pag. 4](#).

Utilizzo dei nodi

I nodi vengono utilizzati in IBM® SPSS® Modeler per semplificare l'analisi dei dati. I vari nodi presenti nell'area di lavoro rappresentano oggetti e operazioni diversi. La palette disponibile nella parte inferiore della finestra di SPSS Modeler contiene tutti i possibili nodi utilizzati per la creazione degli stream.

Sono disponibili vari tipi di nodi. **I nodi di input** importano i dati nello stream e si trovano nella scheda Input della palette dei nodi. **I nodi di elaborazione** eseguono operazioni su singoli campi e record di dati e sono disponibili nelle schede Oper su record e Oper su campi della palette. **I nodi di output** generano vari tipi di output per dati, grafici e risultati di modelli e sono inclusi nelle schede Grafici, Output ed Esporta della palette dei nodi. **I nodi Modelli** utilizzano algoritmi statistici per creare insiemi di modelli e sono contenuti nella scheda Modelli e (se attivata) nella scheda Modelli database della palette dei nodi. [Per ulteriori informazioni, vedere l'argomento Palette nodi in il capitolo 3 a pag. 15.](#)

È possibile connettere i nodi per creare gli stream che, una volta eseguiti, consentono di visualizzare le relazioni e di trarre le conclusioni appropriate. Gli stream sono simili agli script, è infatti possibile salvarli e riutilizzarli con file di dati diversi.

Un nodo eseguibile che elabora i dati di uno stream è detto **nodo terminale**. Un nodo Modelli o di output è un nodo terminale se compare alla fine di uno stream o di un ramo dello stesso. Ai nodi terminali non è possibile collegare altri nodi.

Nota: è possibile personalizzare la palette Nodi. [Per ulteriori informazioni, vedere l'argomento Personalizzazione della palette dei nodi in il capitolo 12 a pag. 237.](#)

Aggiunta di nodi a uno stream

È possibile aggiungere nodi a uno stream dalla palette dei nodi in vari modi:

- Fare doppio clic su un nodo nella palette. *Nota:* facendovi doppio clic sopra, il nodo verrà connesso automaticamente allo stream corrente. [Per ulteriori informazioni, vedere l'argomento Connessione di nodi in uno stream a pag. 40.](#)
- Selezionare un nodo e trascinarlo dalla palette nell'area di disegno dello stream.
- Fare clic su un nodo nella palette e quindi sull'area di disegno dello stream.
- Selezionare l'opzione desiderata dal menu Inserisci di IBM® SPSS® Modeler.

Dopo avere aggiunto un nodo nell'area di disegno dello stream, farvi doppio clic sopra per aprire la relativa finestra di dialogo. Le opzioni disponibili nella finestra di dialogo variano in base al tipo di nodo da aggiungere. Per informazioni sui controlli specifici della finestra di dialogo, fare clic sul pulsante Aiuto.

Rimozione di nodi

Per rimuovere un nodo dallo stream di dati, fare clic sul nodo e premere il tasto CANCEL oppure farvi clic sopra con il pulsante destro del mouse e scegliere Elimina dal menu.

Connessione di nodi in uno stream

I nodi aggiunti nell'area di disegno dello stream non formano uno stream di dati fino a quando non vengono connessi. Le connessioni tra i nodi indicano la direzione del flusso di dati da un'operazione alla successiva. Esistono molti modi per connettere i nodi in modo da formare uno stream, ovvero è possibile fare doppio clic su di essi, utilizzare il pulsante centrale del mouse oppure connetterli manualmente.

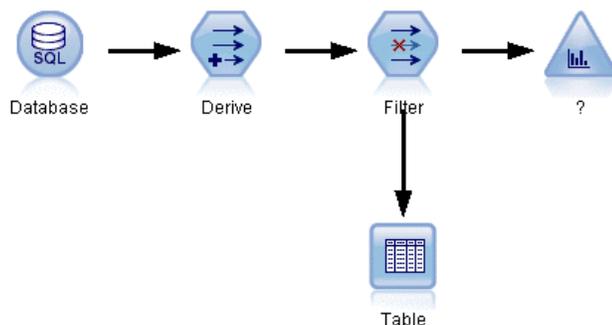
Per aggiungere e connettere i nodi mediante il doppio clic

Il modo più semplice per creare uno stream consiste nel fare doppio clic sui nodi nella palette. Il nuovo nodo verrà connesso automaticamente al nodo selezionato nell'area di disegno dello stream. Se, per esempio, l'area di disegno contiene un nodo Database, è possibile selezionare questo nodo e fare doppio clic sul nodo successivo nella palette, quale un nodo Nuovo campo. Il nodo Nuovo campo verrà connesso automaticamente al nodo Database esistente. È possibile ripetere questa operazione fino a quando si raggiunge un nodo terminale, per esempio un nodo

Istogramma o Tabella. A questo punto, i nuovi nodi verranno connessi all'ultimo nodo a monte diverso da un nodo terminale.

Figura 5-2

Stream creato facendo doppio clic sui nodi nelle palette



Per connettere i nodi utilizzando il pulsante centrale del mouse

Nell'area di disegno dello stream, è possibile fare clic e trascinare da un nodo a un altro utilizzando il pulsante centrale del mouse. (Se il mouse in uso non dispone di un pulsante centrale, è possibile simularlo tenendo premuto il tasto Alt mentre si trascina con il mouse da un nodo a un altro).

Figura 5-3

Utilizzo del pulsante centrale del mouse per connettere i nodi



Per connettere i nodi manualmente

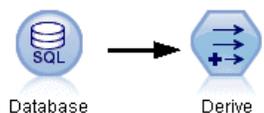
Se non è disponibile un pulsante centrale del mouse e si preferisce connettere manualmente i nodi, utilizzare il menu di scelta rapida di un nodo per connetterlo a un altro nodo già presente nell'area di disegno.

- ▶ Fare clic con il pulsante destro del mouse sul nodo da cui si desidera iniziare la connessione. Viene visualizzato il menu del nodo.
- ▶ Fare clic su Connetti nel menu di scelta rapida.
- ▶ Sul nodo iniziale e in prossimità del cursore viene visualizzata l'icona di una connessione. Fare clic su un secondo nodo per connettere i due nodi.

Figura 5-4
Connessione di nodi utilizzando il comando *Connetti* del menu di scelta rapida



Figura 5-5
Nodi connessi



Per la connessione dei nodi, esistono numerose indicazioni da seguire. Verrà visualizzato un messaggio di errore se si tenta di eseguire i seguenti tipi di connessione:

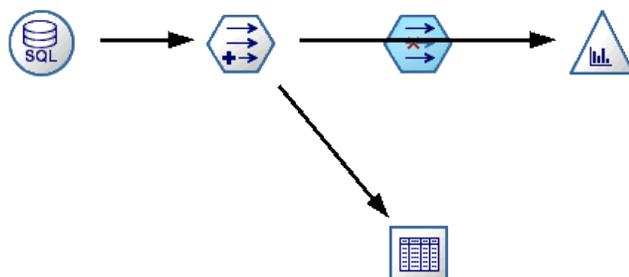
- Una connessione verso un nodo di input
- Una connessione da un nodo finale
- Una connessione relativa a un nodo per il quale è stato superato il numero massimo di connessioni di input consentite
- Una connessione di due nodi già connessi
- Circolarità (i dati tornano a un nodo dal quale sono già passati)

Ignorare i nodi di uno stream

Quando si ignora un nodo dello stream di dati, tutte le relative connessioni di input e output vengono sostituite da connessioni dirette dai nodi di input ai nodi di output. Se il nodo non dispone delle connessioni di input e di output, tutte le relative connessioni vengono eliminate anziché essere reindirizzate.

Si consideri, per esempio, uno stream che deriva un nuovo campo, filtra i campi e quindi analizza i risultati in un istogramma e in una tabella. Se si desidera visualizzare lo stesso grafico e la stessa tabella di dati anche *prima* che i campi vengano filtrati, è possibile aggiungere nuovi nodi Istogramma e Tabella allo stream oppure ignorare il nodo Filtro. Se si ignora il campo Filtro, le connessioni al grafico e alla tabella vengono trasferite direttamente dal nodo Nuovo campo e il nodo Filtro viene disconnesso dallo stream.

Figura 5-6
Ignorare un nodo Filtro già connesso



Per ignorare un nodo

- Nell'area di disegno dello stream, fare doppio clic con il pulsante centrale del mouse sul nodo che si desidera ignorare. In alternativa, è possibile utilizzare Alt+doppio clic.

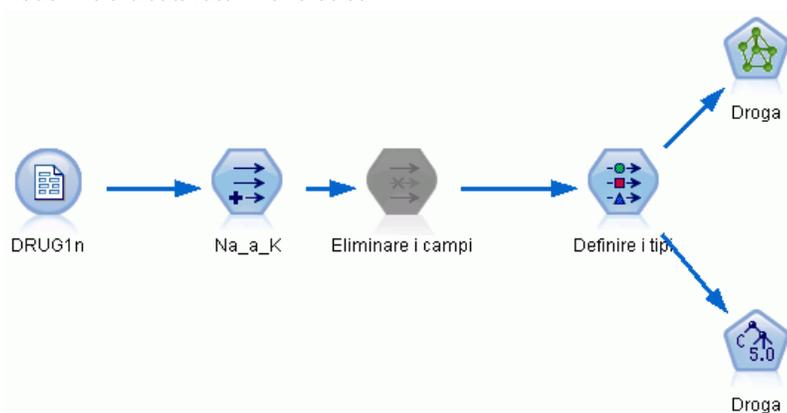
Nota: per annullare l'operazione è possibile fare clic su Annulla nel menu Modifica oppure premere Ctrl+Z.

Disattivazione di nodi in uno stream

I nodi di elaborazione con un unico input all'interno degli stream possono essere disattivati, con il risultato che il nodo viene ignorato durante l'esecuzione. In questo modo non è necessario rimuovere o ignorare il nodo e questo può restare connesso agli altri nodi. Le impostazioni del nodo possono ancora essere aperte e modificate; tuttavia, le modifiche non verranno applicate finché il nodo non viene nuovamente attivato.

Per esempio, si supponga di avere uno stream che filtra diversi campi e in seguito crea i modelli con l'insieme di dati ridotto. Se si desidera creare gli stessi modelli anche *senza* che i campi vengano filtrati, per verificare se i risultati migliorano, è possibile disattivare il nodo Filtro. Quando si disattiva il nodo Filtro, le connessioni ai nodi Modelli vengono trasferite direttamente dal nodo Nuovo campo al nodo Tipo.

Figura 5-7
Nodo Filtro disattivato in uno stream



Per disattivare un nodo

- ▶ Sull'area di disegno dello stream, fare clic con il pulsante destro del mouse sul nodo che si desidera disattivare.
- ▶ Fare clic su Disattiva nodo nel menu di scelta rapida.

In alternativa, fare clic su **Nodo > Disattiva nodo** nel menu Modifica. Quando si desidera includere nuovamente il nodo nello stream, fare clic su **Attiva nodo** allo stesso modo.

Nota: per annullare l'operazione è possibile fare clic su **Annulla** nel menu Modifica oppure premere **Ctrl+Z**.

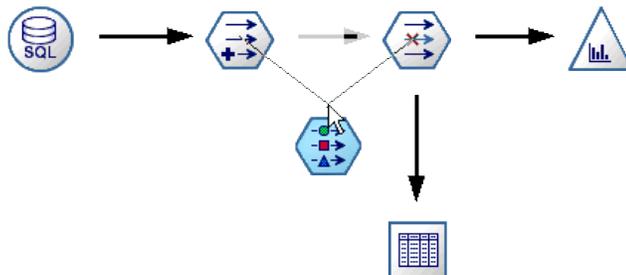
per annullare l'operazione è possibile fare clic su **Annulla** nel menu Modifica oppure premere **Ctrl+Z**.

Aggiunta di nodi a connessioni esistenti

È possibile aggiungere un nuovo nodo tra due nodi già connessi trascinando la freccia che connette i due nodi.

Figura 5-8

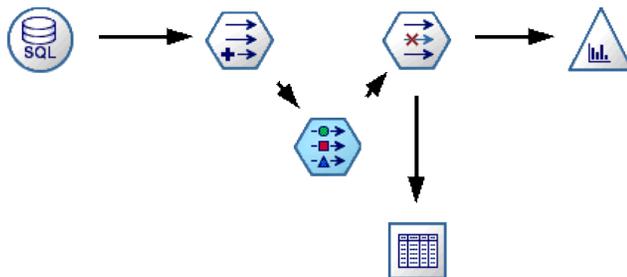
Connessione di un nuovo nodo tra due nodi connessi



- ▶ Fare clic con il pulsante centrale del mouse sulla freccia della connessione in cui si desidera inserire il nuovo nodo. In alternativa, tenere premuto il tasto **Alt** in sostituzione del pulsante centrale del mouse.

Figura 5-9

Nuovo stream



- ▶ Trascinare la connessione sul nodo che si desidera aggiungere e rilasciare il pulsante del mouse.

Nota: per rimuovere le nuove connessioni dal nodo e ripristinare le originali, è possibile **ignorare** il nodo.

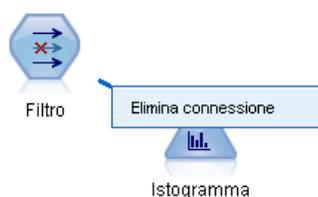
Eliminazione delle connessioni tra nodi

Per eliminare la connessione tra due nodi:

- ▶ Fare clic con il pulsante destro del mouse sulla freccia della connessione.
- ▶ Nel menu, fare clic su Elimina connessione.

Figura 5-10

Eliminazione della connessione tra due nodi di uno stream



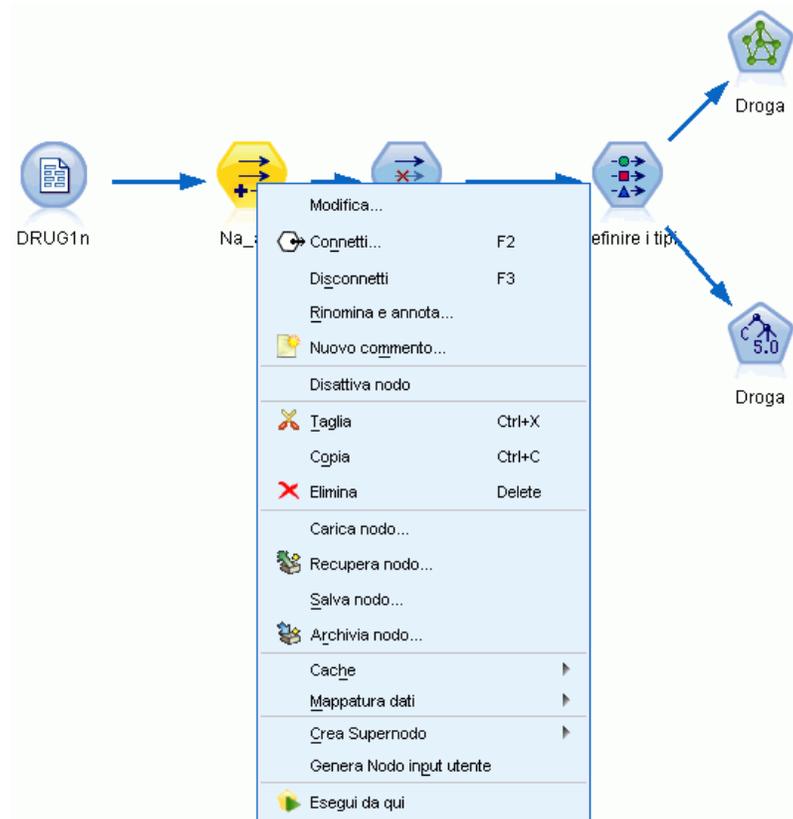
Per eliminare tutte le connessioni di un nodo, procedere in uno dei seguenti modi:

- Selezionare il nodo e premere F3.
- Selezionare un nodo e nel menu principale fare clic su:
Modifica > Nodo > Disconnetti

Impostazione delle opzioni per i nodi

Dopo che i nodi sono stati creati e connessi, è possibile personalizzarli in diversi modi. Fare clic con il pulsante destro del mouse su un nodo e scegliere una delle opzioni del menu.

Figura 5-11
Opzioni del menu di scelta rapida dei nodi



- Fare clic su Modifica per aprire la finestra di dialogo relativa al nodo selezionato.
- Fare clic su Connetti per connettere manualmente un nodo a un altro nodo.
- Fare clic su Disconnetti per eliminare tutte le connessioni a e dal nodo.
- Fare clic su Rinomina e annota per visualizzare la scheda Annotazioni della finestra di dialogo di modifica.
- Fare clic su Nuovo commento per aggiungere un commento relativo al nodo. [Per ulteriori informazioni, vedere l'argomento Aggiunta di commenti e annotazioni a nodi e stream a pag. 69.](#)
- Fare clic su Disattiva nodo per “nascondere” il nodo durante l’elaborazione. Per rendere il nodo nuovamente visibile per l’elaborazione, fare clic su Attiva nodo. [Per ulteriori informazioni, vedere l'argomento Disattivazione di nodi in uno stream a pag. 43.](#)
- Fare clic su Taglia o Elimina per rimuovere il nodo o i nodi selezionati dall’area di disegno dello stream. *Nota:* se si fa clic su Taglia è possibile incollare i nodi, mentre Elimina non lo consente.
- Fare clic su Copia nodo per creare una copia del nodo senza connessioni. La copia potrà essere aggiunta a un nuovo stream o a uno stream esistente.
- Fare clic su Carica nodo per aprire un nodo salvato in precedenza e caricare le relative opzioni nel nodo corrente selezionato. *Nota:* i nodi devono essere dello stesso tipo.

- Fare clic su Recupera nodo per recuperare un nodo da un IBM® SPSS® Collaboration and Deployment Services Repository connesso. [Per ulteriori informazioni, vedere l'argomento Accesso a oggetti in IBM SPSS Collaboration and Deployment Services Repository in il capitolo 5 in IBM SPSS Modeler 14.2 Guida per script e automazione.](#)
- Fare clic su Salva nodo per salvare i dettagli del nodo in un file. È possibile caricare i dettagli del nodo solo in un altro nodo dello stesso tipo.
- Fare clic su Archivia nodo per archiviare il nodo selezionato in un IBM SPSS Collaboration and Deployment Services Repository connesso. [Per ulteriori informazioni, vedere l'argomento Accesso a oggetti in IBM SPSS Collaboration and Deployment Services Repository in il capitolo 5 in IBM SPSS Modeler 14.2 Guida per script e automazione.](#)
- Fare clic su Cache per espandere il menu e visualizzare le opzioni relative all'inserimento nella cache per il nodo selezionato.
- Fare clic su Mappatura dati per espandere il menu e visualizzare le opzioni per la mappatura dei dati a una nuova sorgente o per la specifica dei campi obbligatori.
- Fare clic su Crea Supernodo per espandere il menu e visualizzare le opzioni per la creazione di un Supernodo nello stream corrente. [Per ulteriori informazioni, vedere l'argomento Creazione di Supernodi in il capitolo 9 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)
- Fare clic su Genera Nodo input utente per sostituire il nodo selezionato. Gli esempi generati dal nodo includeranno gli stessi campi del nodo corrente. [Per ulteriori informazioni, vedere l'argomento Nodo input utente in il capitolo 2 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)
- Fare clic su Esegui da qui per eseguire tutti i nodi terminali a valle del nodo selezionato.

Opzioni di copia dei dati nella cache per i nodi

Per ottimizzare l'esecuzione dello stream, è possibile impostare una **cache** in qualsiasi nodo diverso da un nodo terminale. La cache configurata per un nodo viene riempita con i dati che passano attraverso il nodo alla successiva esecuzione dello stream di dati. Da questo momento in poi, i dati vengono letti dalla cache, che è archiviata in una directory temporanea, anziché dalla sorgente dati.

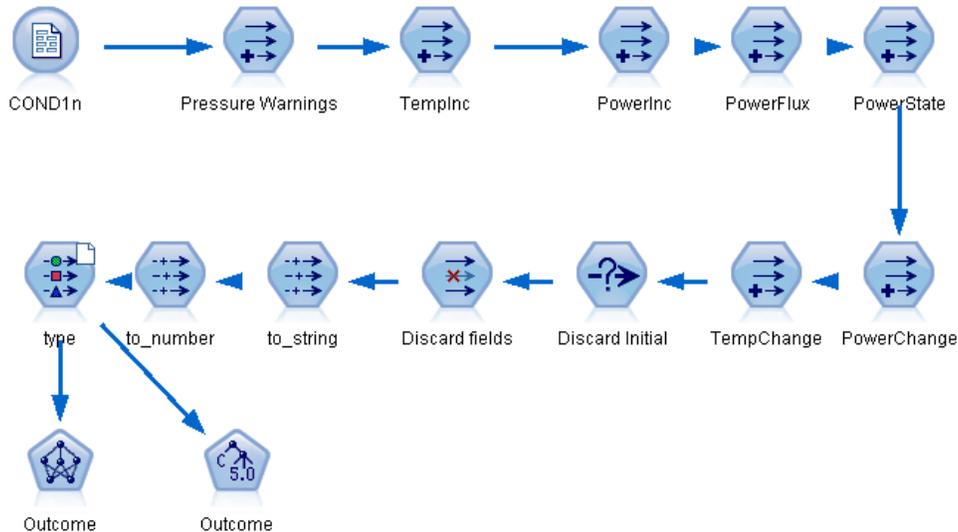
La copia dei dati nella cache è utile soprattutto dopo un'operazione che richiede tempo, per esempio un ordinamento, un'unione o un'aggregazione. Si supponga, per esempio, che esista un nodo di input impostato in modo da leggere i dati relativi alle vendite da un database e un nodo Aggregazione che riepiloga le vendite in base alla posizione. È possibile impostare una cache nel nodo Aggregazione anziché nel nodo di input, in modo tale da memorizzarvi i dati aggregati anziché l'intero insieme di dati.

Nota: l'impostazione di una cache nei nodi di input, che consente semplicemente di archiviare una copia dei dati originali non appena vengono letti in IBM® SPSS® Modeler, nella maggior parte dei casi non migliorerà le prestazioni.

I nodi con la cache attivata sono indicati da una piccola icona di documento nell'angolo in alto a destra. Quando i dati vengono inseriti nella cache del nodo, l'icona del documento diventa di colore verde.

Figura 5-12

Copia dei dati nella cache del nodo Tipo per memorizzare i campi appena derivati



Per attivare una cache:

- ▶ Nell'area di disegno dello stream, fare clic con il pulsante destro del mouse sul nodo e fare clic su Cache nel menu.
- ▶ Fare clic su Attiva nel sottomenu.
- ▶ Per disattivare la cache, fare clic con il pulsante destro del mouse sul nodo e fare clic su Disattiva nel sottomenu.

Inserimento nella cache di nodi nel database

Per gli stream eseguiti nel database, i dati possono essere inseriti nella cache midstream in una tabella temporanea nel database anziché nel file system. Insieme all'ottimizzazione SQL, ciò consente di ottenere significativi vantaggi in termini di prestazioni. Per esempio, l'output di uno stream che unisce più tabelle per creare una visualizzazione data mining può essere inserito nella cache e riutilizzato in un secondo tempo in base alle esigenze. Generando automaticamente codice SQL per tutti i nodi a valle, è possibile migliorare ulteriormente le prestazioni.

Se si utilizza l'inserimento nella cache del database con stringhe di lunghezza superiore a 256 caratteri, assicurarsi che sia presente un nodo Tipo a monte del nodo per cui è attivata la cache e che i valori dei campi vengano letti. In questo modo la colonna corrispondente nella tabella temporanea viene impostata sulla larghezza corretta per adattarsi alle stringhe.

Per sfruttare l'inserimento nella cache del database, è necessario che siano attivate l'ottimizzazione SQL e l'inserimento nella cache del database. Si noti che le impostazioni di ottimizzazione del Server prevalgono su quelle del Client. [Per ulteriori informazioni, vedere l'argomento Impostazione delle opzioni di ottimizzazione in il capitolo 12 a pag. 233.](#)

Con l'inserimento nella cache del database attivato, è sufficiente fare clic con il pulsante destro del mouse su qualsiasi nodo non terminale per inserire i dati nella cache in quel punto e la cache verrà creata automaticamente direttamente nel database alla successiva esecuzione dello stream. Se l'inserimento nella cache del database o l'ottimizzazione SQL non sono attivati, la cache verrà scritta nel file system.

Nota: i seguenti database supportano tabelle temporanee per l'inserimento di dati nella cache: DB2, Netezza, Oracle, SQL Server e Teradata. Altri database utilizzano una tabella normale per l'inserimento di dati nella cache. È possibile personalizzare il codice SQL per database specifici - contattare l'assistenza.

Per svuotare una cache

Un'icona del documento di colore bianco su un nodo indica che la relativa cache è vuota. Quando la cache è piena, l'icona del documento diventa di colore verde. Per sostituire il contenuto della cache, è necessario innanzitutto svuotare la cache e quindi rieseguire lo stream di dati per riempirla di nuovo.

- ▶ Nell'area di disegno dello stream, fare clic con il pulsante destro del mouse sul nodo e fare clic su Cache nel menu.
- ▶ Fare clic su Svuota nel sottomenu.

Per salvare una cache

È possibile salvare il contenuto di una cache come file di dati IBM® SPSS® Statistics (*.sav), e quindi ricaricare il file come una cache oppure impostare un nodo che utilizza il file cache come la propria sorgente dati. È inoltre possibile caricare una cache salvata da un altro progetto.

- ▶ Nell'area di disegno dello stream, fare clic con il pulsante destro del mouse sul nodo e fare clic su Cache nel menu.
- ▶ Fare clic su Salva cache nel sottomenu.
- ▶ Nella finestra di dialogo Salva cache passare alla posizione in cui si desidera salvare il file cache.
- ▶ Digitare un nome nella casella di testo Nome file.
- ▶ Verificare che nell'elenco dei tipi di file sia selezionato *.sav e fare clic su Salva.

Per caricare una cache

I file cache che sono stati salvati prima di essere rimossi dal nodo possono essere ricaricati.

- ▶ Nell'area di disegno dello stream, fare clic con il pulsante destro del mouse sul nodo e fare clic su Cache nel menu.
- ▶ Fare clic su Carica cache nel sottomenu.
- ▶ Nella finestra di dialogo Carica cache passare alla posizione del file cache, selezionarlo e fare clic su Carica.

Anteprima dei dati nei nodi

Per fare in modo che i dati vengano modificati secondo le aspettative durante la creazione di uno stream, è possibile elaborare i dati stessi mediante un nodo Tabella a ogni punto significativo del processo. In alternativa, è possibile generare da ciascun nodo un'anteprima di un campione dei dati che verranno creati, riducendo così il tempo necessario per creare ciascun nodo.

Per i nodi a monte di un insieme di modelli, l'anteprima mostra i campi di input; per un insieme di modelli o per i nodi a valle dell'insieme (ad eccezione dei nodi terminali) l'anteprima mostra i campi di input e generati.

Per impostazione predefinita, vengono visualizzate dieci righe. È tuttavia possibile modificare il valore nelle proprietà dello stream. [Per ulteriori informazioni, vedere l'argomento Impostazione delle opzioni per gli stream a pag. 51.](#)

Figura 5-13
Anteprima dei dati da un insieme di modelli

	year_built	volume_interior	volume_other	lot_size	taxable_value	\$XR-taxable_value	\$XRE-taxable_value
1	379	166	11	100	90500	105523.184	1526.604
2	388	603	73	497	420000	355497.802	14929.477
3	387	303	75	91	152500	147713.631	1733.849
4	326	228	12	55	92000	69313.952	8116.212
5	388	666	145	441	390000	367551.491	16353.591
6	370	563	60	800	435000	395173.708	4759.398
7	302	355	41	225	130000	109865.220	3694.157
8	372	468	78	625	380500	334714.268	6367.557
9	386	315	14	107	158000	145149.106	1263.839
10	375	282	27	141	134000	138747.302	1892.730

Dal menu Genera è possibile creare diversi tipi di nodi. [Per ulteriori informazioni, vedere l'argomento Browser del nodo Tabella in il capitolo 6 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)

Nodi di blocco

Per impedire ad altri utenti di modificare le impostazioni di uno o più nodi di uno stream è possibile incapsulare i nodi in uno speciale tipo di nodo detto Supernodo e quindi bloccare quest'ultimo proteggendolo con una password. [Per ulteriori informazioni, vedere l'argomento Blocco di Supernodi in il capitolo 9 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)

Utilizzo degli stream

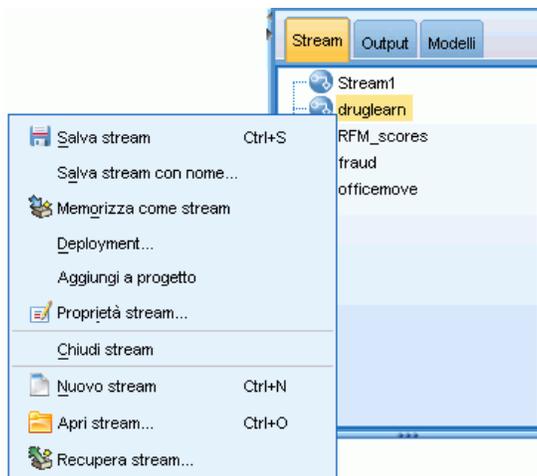
La connessione dei nodi di input, di processo e finali determina la creazione di uno stream. Gli stream non sono altro che un insieme di nodi e possono pertanto essere salvati, annotati e aggiunti ai progetti. È inoltre possibile impostare numerose opzioni per gli stream, per esempio le opzioni

per l'ottimizzazione, le impostazioni di data e ora, i parametri e gli script. Queste proprietà verranno illustrate nelle sezioni seguenti.

In IBM® SPSS® Modeler è possibile utilizzare e modificare contemporaneamente più stream di dati. Nella parte destra della finestra principale è disponibile il riquadro dei manager, che semplifica gli spostamenti negli stream, negli output e nei modelli aperti. Se il riquadro dei manager non è visibile, fare clic su Manager nel menu Visualizza e fare clic sulla scheda Stream.

Figura 5-14

Scheda Stream del riquadro dei manager con le opzioni del menu di scelta rapida



Da questa finestra è possibile:

- Accedere agli stream.
- Salvare gli stream.
- Salvare gli stream nel progetto corrente.
- Chiudere gli stream.
- Aprire nuovi stream.
- Archiviare e recuperare stream da un repository di IBM SPSS Collaboration and Deployment Services (se disponibile sul sito). [Per ulteriori informazioni, vedere l'argomento Informazioni su IBM SPSS Collaboration and Deployment Services Repository in il capitolo 9 a pag. 156.](#)

Fare clic con il pulsante destro del mouse su uno stream nella scheda Stream per accedere a queste opzioni.

Impostazione delle opzioni per gli stream

Per lo stream corrente è possibile specificare numerose opzioni, molte delle quali sono valide per le espressioni CLEM.

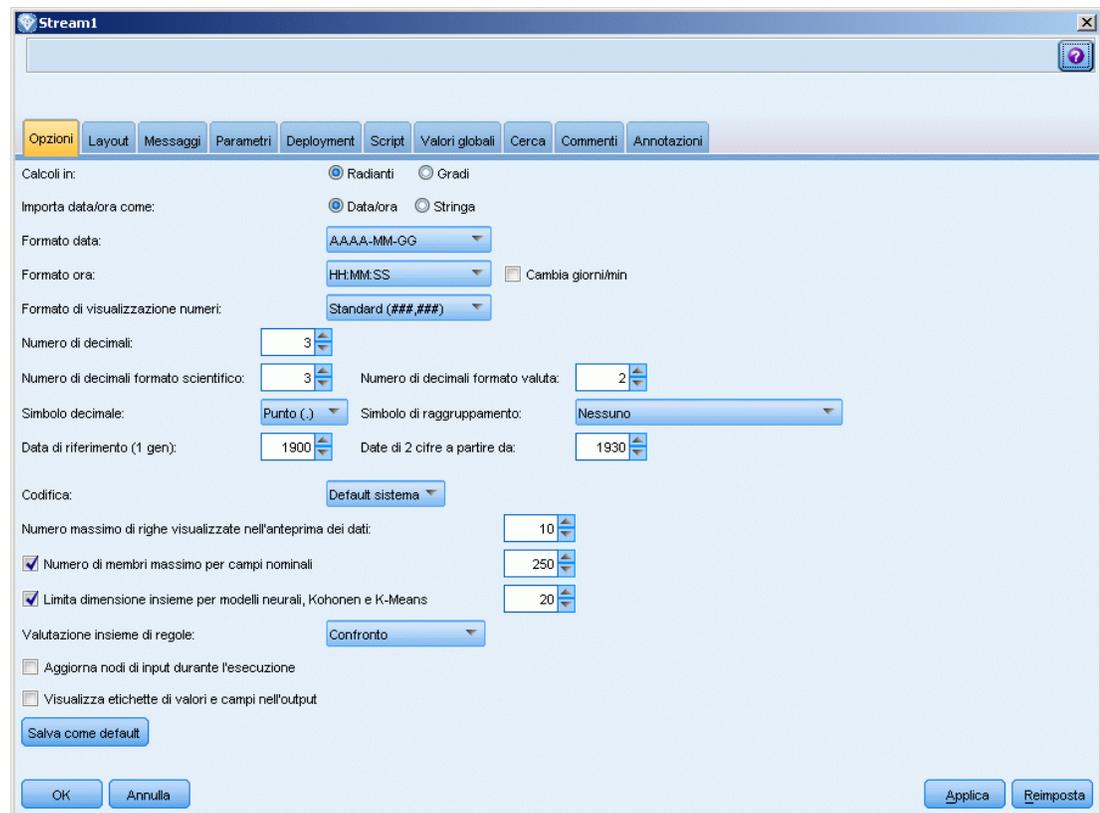
Per impostare le opzioni degli stream

- ▶ Nel menu File fare clic su Proprietà stream. In alternativa, selezionare lo stream dalla scheda Stream del riquadro dei manager, quindi fare clic con il pulsante destro del mouse e scegliere Proprietà stream dal menu di scelta rapida).
- ▶ Fare clic sulla scheda Opzioni.

In alternativa, nel menu Strumenti, fare clic su:
Proprietà stream > Opzioni

Figura 5-15

Scheda Opzioni della finestra di dialogo delle proprietà dello stream



Calcoli in. Selezionare Radianti o Gradi per specificare l'unità di misura da utilizzare nelle espressioni CLEM trigonometriche.

Importa data/ora come. Specificare se i campi di data/ora verranno memorizzati nel formato di data/ora o se verranno importati come variabili stringa.

Formato data. Selezionare un formato di data da utilizzare per i campi di archiviazione delle date o per le stringhe interpretate come date dalle funzioni di data CLEM.

Formato ora. Selezionare un formato di ora da utilizzare per i campi di archiviazione dell'ora o per le stringhe interpretate come orari dalle funzioni di ora CLEM.

Cambia giorni/min. Per i formati di ora, specificare se le differenze di ora negative devono essere interpretate come riferite al giorno o all'ora precedenti.

Formato di visualizzazione numeri. È possibile scegliere il formato di visualizzazione standard (####.###), scientifico (#.###E+##) o valuta (\$###.##).

Numero di decimali (standard, scientifico, valuta). Per i formati di visualizzazione numerici, specifica il numero di decimali utilizzati per la visualizzazione o la stampa di numeri reali. Questa opzione viene specificata separatamente per ogni formato di visualizzazione.

Simbolo decimale. Selezionare la virgola (,) o il punto (.) come separatore decimale.

Simbolo di raggruppamento. Per i formati di visualizzazione numerici, selezionare il simbolo utilizzato per raggruppare valori, per esempio, il punto 3.000. Le opzioni disponibili sono Nessuno, Punto, Virgola, Spazio e Definito in base alle opzioni internazionali (in questo caso verrà utilizzato il valore di default per le opzioni internazionali correnti).

Data di riferimento (1 gen). Selezionare gli anni di riferimento (sempre 1 gen) che verranno utilizzati dalle funzioni di data CLEM che operano con una singola data.

Date di 2 cifre a partire da. Specificare l'anno di riferimento per l'aggiunta delle cifre del secolo negli anni a due cifre. Per esempio, se si specifica 1930 come anno di riferimento, si presupporrà che il 05/11/02 sia nell'anno 2002. La medesima impostazione utilizzerà il 20esimo secolo per le date successive all'anno 1930. Di conseguenza, si presupporrà che il 05/11/73 sia nell'anno 1973.

Codifica. Specificare il metodo di default dello stream per la codifica di testo. (*Nota:* si applica solo al nodo di input Testo variabile e al nodo di esportazione File piatto. L'impostazione non viene utilizzata da altri nodi. La maggior parte dei file di dati includono informazioni di codifica incorporate). È possibile scegliere l'impostazione di default del sistema o UTF-8. L'impostazione di default del sistema è specificata nel Pannello di controllo di Windows oppure, se si esegue il programma in modalità distribuita, sul computer server. [Per ulteriori informazioni, vedere l'argomento Supporto Unicode in IBM SPSS Modeler in l'appendice B a pag. 264.](#)

Numero massimo di righe visualizzate nell'anteprima dei dati. Specificare il numero di righe da visualizzare quando viene richiesta un'anteprima dei dati di un nodo. [Per ulteriori informazioni, vedere l'argomento Anteprima dei dati nei nodi a pag. 50.](#)

Numero di membri massimo per campi nominali. Selezionare questa opzione per specificare il numero massimo di membri per i campi nominali (insieme) dopo il quale il tipo di campo diventa **Senza tipo**. Tale opzione è utile quando si lavora con campi nominali di grandi dimensioni. *Nota:* Quando il livello di misurazione di un campo è impostato su **Senza tipo**, il suo ruolo viene impostato automaticamente su **Nessuno**, e ciò implica che i campi non sono disponibili per la fase di modeling.

Limita dimensione insieme per modelli neurali, Kohonen e K-Means. Selezionare questa opzione per specificare un numero massimo di membri per i campi nominali utilizzati nei modelli di reti neurali, reti Kohonen e K-Means. La dimensione di default è 20, dopo la quale il campo verrà ignorato e verrà visualizzato un avviso che fornisce informazioni sul campo.

Valutazione insieme di regole. Specifica la modalità di valutazione dei modelli di insiemi di regole. Per default, gli insiemi di regole utilizzano l'impostazione Confronto per combinare le previsioni delle singole regole e definire la previsione finale. Per assicurarsi che gli insiemi di regole utilizzino per default la regola del primo risultato positivo, selezionare Primo risultato positivo. [Per](#)

ulteriori informazioni, vedere l'argomento [Insiemi di modelli dell'Insieme di regole in il capitolo 6 in IBM SPSS Modeler 14.2 Nodi Modelli](#). Si noti che questa opzione non si applica ai modelli Elenco decisionale, che utilizzano sempre il primo risultato positivo definito dall'algoritmo.

Aggiorna nodi di input durante l'esecuzione. Selezionare questa opzione per aggiornare automaticamente tutti i nodi di input durante l'esecuzione dello stream corrente. Questa operazione è analoga alla selezione del pulsante Aggiorna su un nodo Input, ma consente di aggiornare automaticamente tutti i nodi, esclusi i nodi Input utente, per lo stream corrente.

Nota: se si seleziona questa opzione, la cache dei nodi downstream verrà svuotata, anche se i dati non sono stati modificati. Durante l'esecuzione dello stream viene tuttavia eseguito un solo svuotamento. Sarà quindi possibile utilizzare comunque le cache a valle come archivio temporaneo per una singola esecuzione. Per esempio, si supponga che sia stata impostata una cache midstream dopo una complessa operazione di creazione di un nuovo campo e che downstream rispetto al nodo Nuovo campo siano disponibili svariati grafici e report. Durante l'esecuzione dello stream, la cache del nodo Nuovo campo verrà svuotata e riempita nuovamente, ma solo per il primo grafico o report. I nodi finali successivi leggeranno i dati dalla cache del nodo Nuovo campo.

Visualizza etichette di valori e campi nell'output. Visualizza etichette di valori e campi in tabelle, grafici e altre forme di output. Se non esistono etichette, vengono visualizzati i nomi dei campi e i valori dei dati. Le etichette sono disattivate per default; tuttavia, è possibile attivarle o disattivarle singolarmente altrove in IBM® SPSS® Modeler. È inoltre possibile scegliere di visualizzare etichette nella finestra di output utilizzando un pulsante di attivazione disponibile sulla barra degli strumenti.

Figura 5-16

Icona della barra degli strumenti utilizzata per attivare le etichette di campi e valori



Salva come default. Le opzioni specificate vengono applicate solo allo stream corrente. Fare clic su questo pulsante per impostare queste opzioni come default per tutti gli stream.

Impostazione delle opzioni relative al layout dello stream

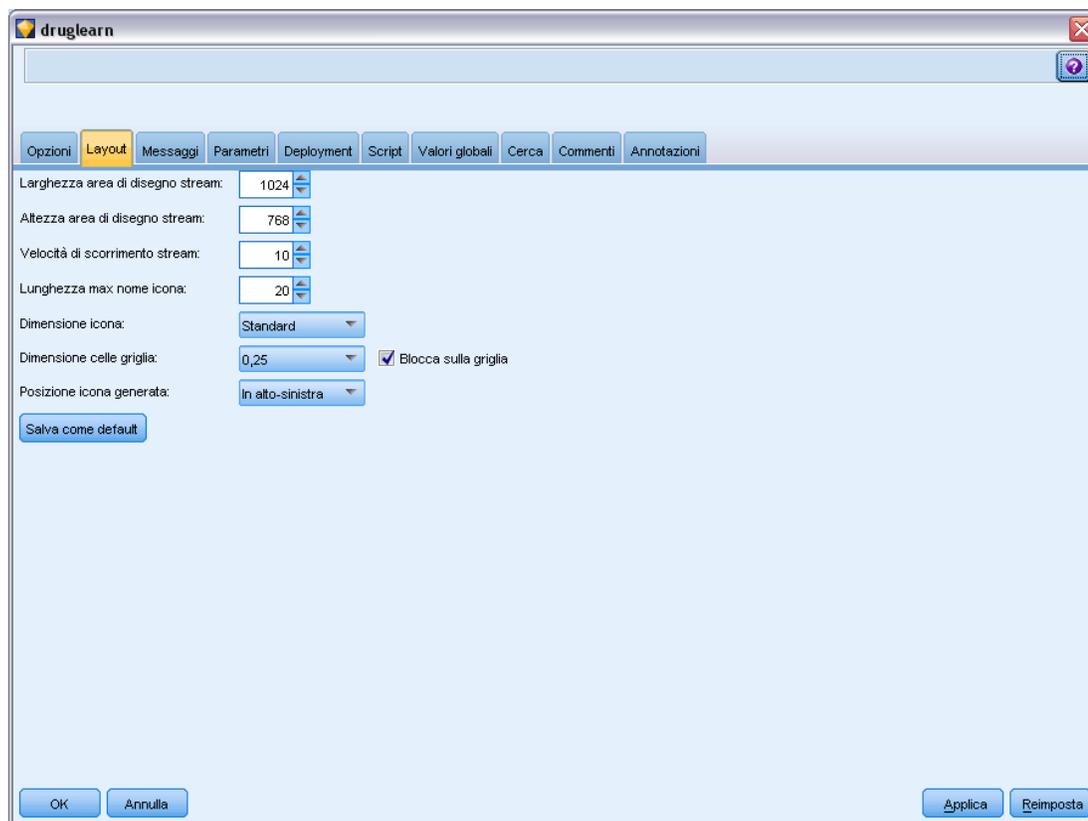
Nella scheda Layout della finestra di dialogo delle proprietà dello stream è possibile specificare numerose opzioni relative alla visualizzazione e all'utilizzo dell'area di disegno dello stream.

Per impostare le opzioni relative al layout

- ▶ Nel menu File fare clic su Proprietà stream. In alternativa, selezionare lo stream dalla scheda Stream del riquadro dei manager, quindi fare clic con il pulsante destro del mouse e scegliere Proprietà stream dal menu di scelta rapida).
- ▶ Fare clic sulla scheda Layout.

In alternativa, nel menu Strumenti, fare clic su:
Proprietà stream > Layout

Figura 5-17
Scheda Layout della finestra di dialogo delle proprietà dello stream



Larghezza area di disegno stream. Specificare la larghezza in pixel dell'area di disegno dello stream.

Altezza area di disegno stream. Specificare l'altezza in pixel dell'area di disegno dello stream.

Velocità di scorrimento stream. Specificare la velocità di scorrimento dell'area di disegno dello stream per controllare la velocità a cui il riquadro dell'area di disegno dello stream scorre quando un nodo viene trascinato da un punto all'altro dell'area di disegno. Più alti sono i valori, maggiore è la velocità di scorrimento.

Lunghezza max nome icona. Specificare un numero massimo di caratteri per i nomi dei nodi dell'area di disegno dello stream.

Dimensione icona. Specificare se nell'area di disegno dello stream verranno visualizzate icone dei nodi grandi o piccole.

Dimensione celle griglia. Selezionare una dimensione delle celle della griglia dall'elenco. Il valore specificato verrà utilizzato per l'allineamento dei nodi nell'area di disegno dello stream mediante una griglia invisibile. La dimensione di default per le celle della griglia è 0,25.

Blocca sulla griglia. Selezionare questa opzione per allineare le icone su una griglia invisibile (opzione di default).

Posizione icona generata. Scegliere il punto dell'area di disegno in cui posizionare le icone relative ai nodi generati dagli insiemi di modelli. L'impostazione predefinita è in alto a sinistra.

Salva come default. Le opzioni specificate vengono applicate solo allo stream corrente. Fare clic su questo pulsante per impostare queste opzioni come default per tutti gli stream.

Visualizzazione dei messaggi relativi alle operazioni dello stream

I messaggi relativi alle operazioni dello stream, per esempio l'esecuzione, l'ottimizzazione e il tempo richiesto dalla creazione dei modelli e dalla valutazione possono essere facilmente visualizzati mediante la scheda Messaggi della finestra di dialogo delle proprietà dello stream. I messaggi di errore vengono inoltre riportati nella tabella.

Per visualizzare i messaggi relativi allo stream

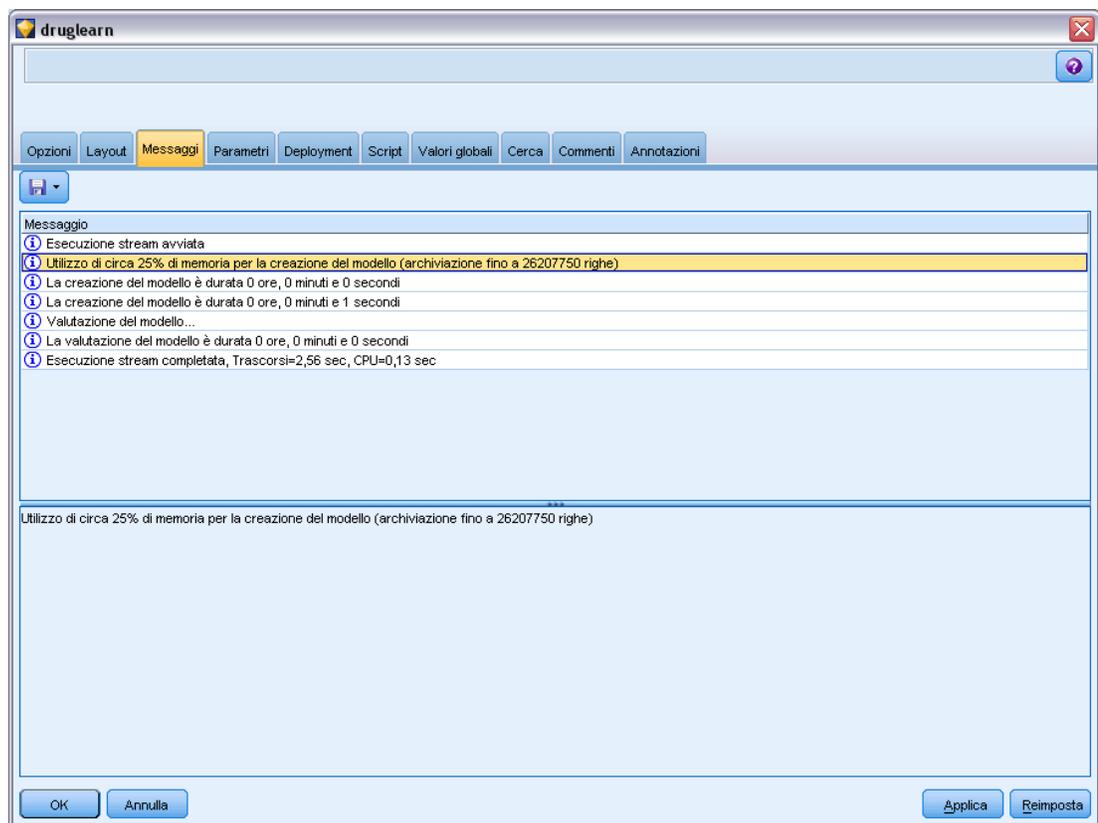
- ▶ Nel menu File fare clic su Proprietà stream. In alternativa, selezionare lo stream dalla scheda Stream del riquadro dei manager, quindi fare clic con il pulsante destro del mouse e scegliere Proprietà stream dal menu di scelta rapida).
- ▶ Fare clic sulla scheda Messaggi.

In alternativa, nel menu Strumenti, fare clic su:

Proprietà stream > Messaggi

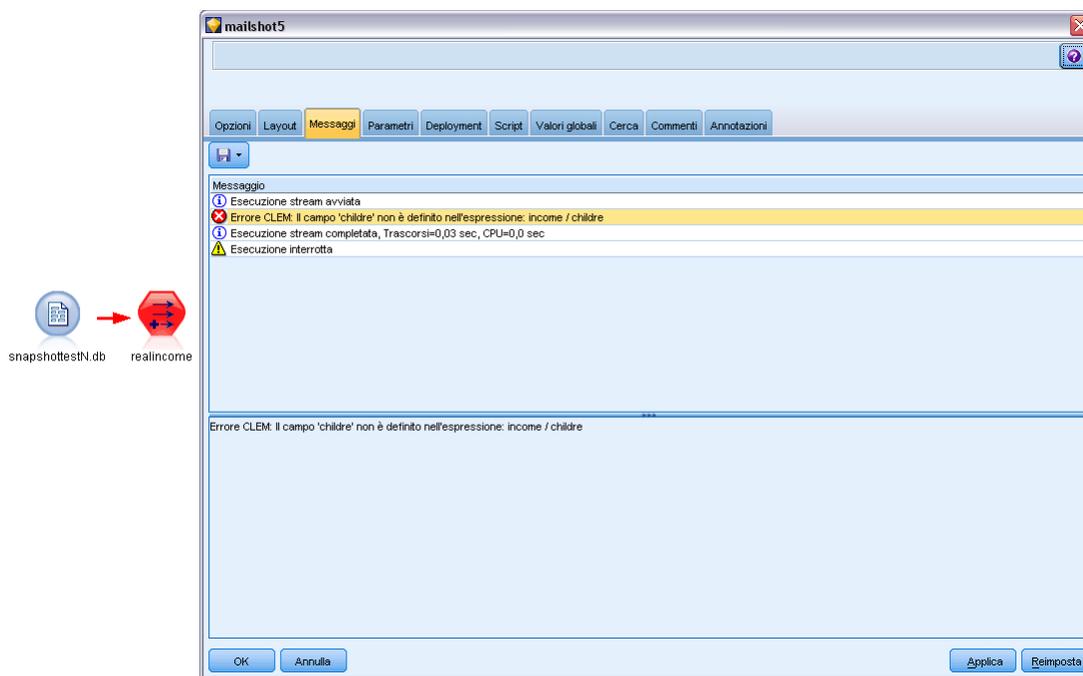
Figura 5-18

Scheda Messaggi della finestra di dialogo delle proprietà dello stream



Oltre ai messaggi relativi alle operazioni di sistema, nella scheda sono visualizzati anche i messaggi di errore. Se l'esecuzione dello stream viene interrotta a causa di un errore, verrà visualizzata la scheda Messaggi di questa finestra di dialogo con il messaggio di errore specifico. Inoltre, il nodo con gli errori sarà evidenziato in rosso nell'area di disegno dello stream.

Figura 5-19
Stream in esecuzione con segnalazione degli errori



Se le opzioni di ottimizzazione SQL e registrazione sono attivate nella finestra di dialogo Opzioni utente, vengono visualizzate anche le informazioni sul codice SQL generato. [Per ulteriori informazioni, vedere l'argomento Impostazione delle opzioni di ottimizzazione in il capitolo 12 a pag. 233.](#)

È possibile salvare i messaggi relativi a uno stream facendo clic su Salva messaggi nell'elenco a discesa del pulsante di salvataggio (a sinistra, subito sotto la scheda Messaggi). È inoltre possibile cancellare tutti i messaggi per uno stream specifico facendo clic su Cancella tutti i messaggi nello stesso elenco.

Impostazione di parametri per stream e sessione

È possibile definire parametri da utilizzare in espressioni CLEM e negli script. Si tratta a tutti gli effetti di variabili definite dall'utente che vengono salvate e mantenute con lo stream, la sessione o il Supernodo corrente e a cui è possibile accedere dall'interfaccia utente nonché attraverso script. Per esempio, durante il salvataggio di uno stream, qualsiasi parametro impostato per tale stream viene salvato contemporaneamente. In questo modo i parametri vengono distinti dalle variabili di script locali, utilizzabili solo nello script in cui sono dichiarate. In ambiente di scripting i parametri fanno spesso parte di un'espressione CLEM e il relativo valore è specificato nello script stesso.

L'ambito di un parametro dipende dalla posizione in cui è impostato:

- i parametri impostati per uno stream possono essere impostati in uno script dello stream o nella finestra di dialogo delle proprietà dello stream e sono disponibili a tutti i nodi dello stream. Vengono visualizzati nell'elenco Parametri nel generatore di espressioni.
- I parametri impostati per una sessione possono essere impostati in uno script autonomo oppure nella finestra di dialogo dei parametri della sessione e sono disponibili per tutti gli stream utilizzati nella sessione corrente, ovvero per tutti gli stream elencati nella scheda Stream del riquadro dei manager.

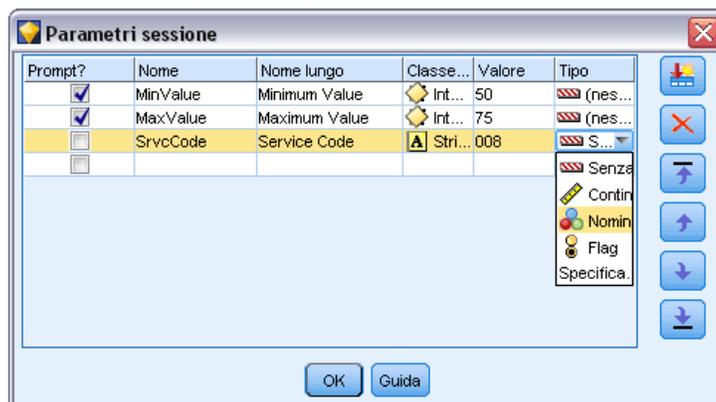
È inoltre possibile impostare parametri per Supernodi. In questo caso, sono visibili solo ai nodi incapsulati all'interno del Supernodo. [Per ulteriori informazioni, vedere l'argomento Definizione dei parametri di un Supernodo in il capitolo 9 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)

[Per informazioni sull'impostazione e l'utilizzo dei parametri negli script, vedere Parametri stream, sessione e Supernodo nel il capitolo 3.](#)

Per impostare parametri per stream e sessione tramite l'interfaccia utente

- ▶ Per impostare i parametri degli stream, nel menu principale fare clic su:
Strumenti > Proprietà stream > Parametri
- ▶ Per impostare parametri di sessione, fare clic su Imposta parametri sessione nel menu Strumenti.

Figura 5-20
Impostazione dei parametri per la sessione



Prompt?. Selezionare questa casella per far sì che il sistema chieda all'utente di inserire il valore di questo parametro al momento dell'esecuzione.

Nome. In questa colonna sono visualizzati i nomi dei parametri. È possibile creare un nuovo parametro inserendo nel campo un nome. Per esempio, per creare un parametro per la temperatura minima, è possibile digitare `valminimo`. Non includere il prefisso `$P-` che indica un parametro nelle espressioni CLEM. Tale nome è inoltre quello visualizzato nel Generatore espressioni CLEM.

Nome lungo. Elenca il nome descrittivo di ogni parametro creato.

Classe archiviazione. Selezionare un tipo di archiviazione dall'elenco. La classe di archiviazione indica la modalità di memorizzazione dei valori dei dati nel parametro. Per esempio, quando si utilizzano valori contenenti zero iniziali che si desidera mantenere (per esempio 008), è necessario selezionare Stringa come tipo di archiviazione. In caso contrario, gli zeri verranno rimossi dal valore. I tipi di archiviazione disponibili sono stringa, intero, reale, ora, data e timestamp. Per i parametri data, tenere presente che è necessario specificare i valori con la notazione standard ISO, come illustrato di seguito.

Valore. Elenca il valore corrente di ogni parametro. Modificare il parametro in base alle esigenze specifiche. Si noti che per specificare i parametri data è necessario utilizzare la notazione standard ISO (ovvero, YYYY-MM-DD). Le date specificate in altri formati non sono accettate.

Tipo (facoltativo). Se si pianifica il deployment dello stream in un'applicazione esterna, selezionare un livello di misurazione dall'elenco. In caso contrario, si consiglia di non modificare la colonna *Tipo*. Se si desidera specificare dei vincoli per il valore del parametro quale un limite superiore e inferiore per un intervallo numerico, selezionare Specifica dall'elenco.

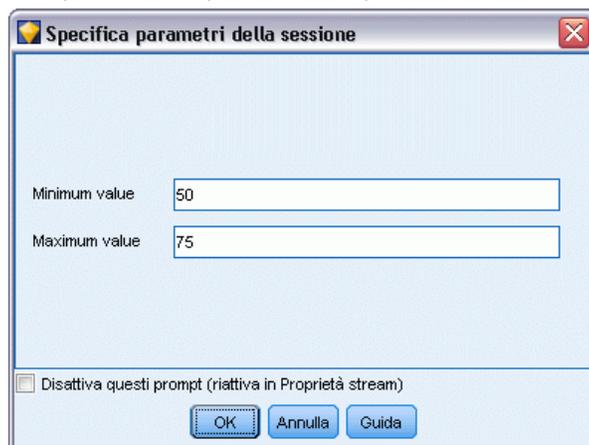
Si noti che è possibile impostare le opzioni nome lungo, classe archiviazione e tipo per i parametri solo tramite l'interfaccia utente e non tramite script.

Fare clic sulle frecce a destra per spostare il parametro selezionato più in alto o in basso nell'elenco dei parametri. Per rimuovere il parametro selezionato, utilizzare il pulsante di eliminazione contrassegnato da una *X*.

Definizione dei prompt di runtime per i valori dei parametri

In presenza di stream in cui è necessario inserire valori diversi per lo stesso parametro a seconda dei casi, è possibile specificare dei prompt di runtime per uno o più valori di parametri di stream o di sessione.

Figura 5-21
Prompt di runtime per i valori dei parametri



Parametri. (Facoltativo) Inserire un valore per il parametro o lasciare il valore di default se già indicato.

Disattiva questi prompt. Selezionare questa casella per non visualizzare questi prompt quando si esegue lo stream. È possibile fare in modo che vengano visualizzati nuovamente selezionando la casella di controllo Prompt? nella finestra di dialogo delle proprietà dello stream o della sessione in cui sono stati definiti i parametri. [Per ulteriori informazioni, vedere l'argomento Impostazione di parametri per stream e sessione a pag. 57.](#)

Definizione di vincoli per i valori di un tipo di parametro

È possibile rendere disponibili dei vincoli per i valori di un parametro durante il deployment degli stream a un'applicazione esterna che legge gli stream di modellazione dei dati. Questa finestra di dialogo consente di specificare i valori disponibili per un utente esterno che esegue lo stream. A seconda del tipo di dati, i vincoli relativi ai valori cambiano dinamicamente nella finestra di dialogo. Le opzioni illustrate sono identiche a quelle disponibili per i valori relativi al nodo Tipo.

Figura 5-22
Definizione dei valori disponibili per un parametro



Tipo. Visualizza il livello di misurazione selezionato. Questo valore si può modificare per indicare il modo in cui si intende utilizzare il parametro in IBM® SPSS® Modeler.

Classe archiviazione. Visualizza il tipo di archiviazione, se noto. I tipi di archiviazione non sono influenzati dal livello di misurazione (continuo, nominale o flag) scelto in SPSS Modeler. È possibile modificare il tipo di archiviazione nella scheda Parametri principale.

La parte inferiore della finestra di dialogo viene modificata dinamicamente, a seconda del livello di misurazione selezionato nel campo Tipo.

Livelli di misurazione continui

Ultimo. Specificare un limite inferiore per i valori dei parametri.

Primo. Specificare un limite superiore per i valori dei parametri.

Etichette. È possibile definire le etichette per qualsiasi valore di un campo intervallo. Fare clic sul pulsante Etichette per aprire una finestra di dialogo nella quale è possibile definire le etichette dei valori.

Livelli di misurazione nominali

Valori. Questa opzione consente di specificare i valori di un parametro che verrà utilizzato come campo nominale. I valori non verranno forzati nello stream di SPSS Modeler, ma verranno utilizzati in un elenco a discesa per le applicazioni di deployment esterne. I pulsanti freccia e quello contrassegnato da una X consentono di modificare l'ordine dei valori esistenti o di eliminarli.

Livelli di misurazione flag

Vero. Specificare un valore flag per il parametro quando la condizione viene soddisfatta.

Falso. Specificare un valore flag per il parametro quando la condizione non viene soddisfatta.

Etichette. È possibile definire le etichette per i valori di un campo flag.

Opzioni di deployment degli stream

La scheda Deployment della finestra di dialogo delle proprietà dello stream consente di specificare le opzioni per il deployment dello stream come scenario in IBM® SPSS® Collaboration and Deployment Services ai fini dell'aggiornamento dei modelli, della pianificazione automatica dei processi o di ulteriore uso da parte di IBM® SPSS® Decision Management o Predictive Applications 5.x. Per poter essere sottoposti a deployment, tutti gli stream necessitano di un ramo di calcolo del punteggio designato; gli altri requisiti e opzioni dipendono dal tipo di deployment. [Per ulteriori informazioni, vedere l'argomento Archiviazione e deployment di oggetti di IBM SPSS Collaboration and Deployment Services Repository in il capitolo 9 a pag. 158.](#)

Visualizzazione dei valori globali degli stream

Mediante la scheda Valori globali della finestra di dialogo delle proprietà dello stream è possibile visualizzare i valori globali impostati per lo stream corrente. Per creare i valori globali viene utilizzato un nodo Globali, che consente di determinare le statistiche quali la media, la somma o la deviazione standard per i campi selezionati. [Per ulteriori informazioni, vedere l'argomento Nodo Globali in il capitolo 6 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)

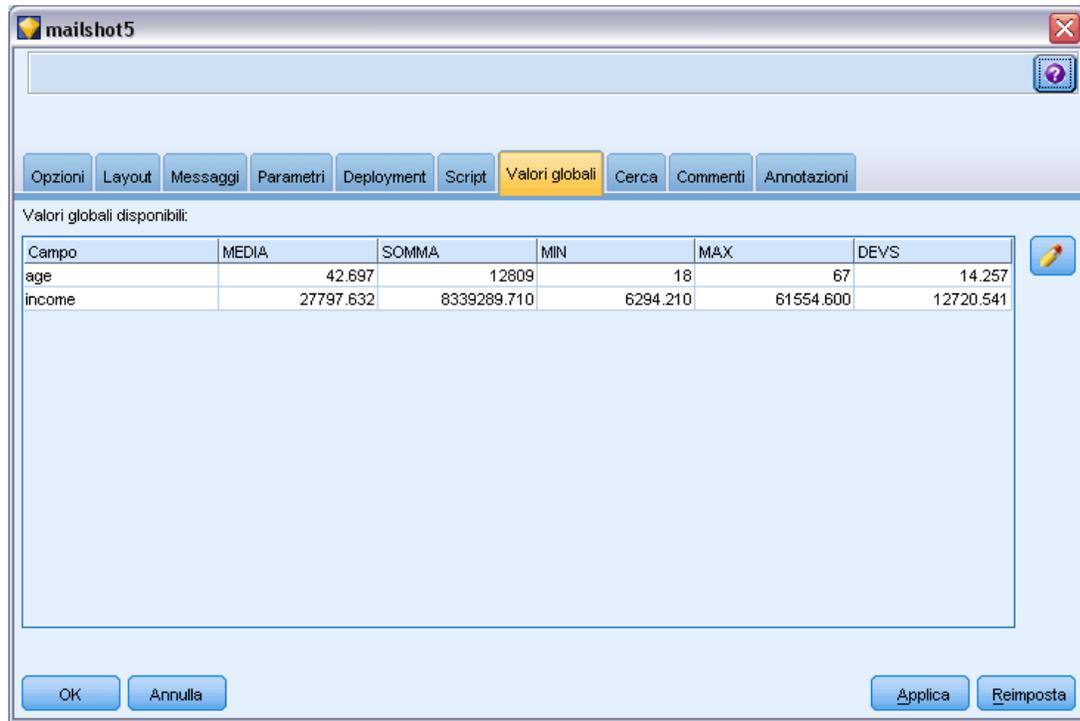
Dopo avere eseguito il nodo Globali, sarà possibile utilizzare i valori ottenuti in molti modi diversi nelle operazioni dello stream. [Per ulteriori informazioni, vedere l'argomento Funzioni globali in il capitolo 8 a pag. 152.](#)

Per visualizzare i valori globali per uno stream

- ▶ Nel menu File fare clic su Proprietà stream. In alternativa, selezionare lo stream dalla scheda Stream del riquadro dei manager, quindi fare clic con il pulsante destro del mouse e scegliere Proprietà stream dal menu di scelta rapida).
- ▶ Fare clic sulla scheda Valori globali.

In alternativa, nel menu Strumenti, fare clic su:
Proprietà stream > Valori globali

Figura 5-23
Visualizzazione dei valori globali disponibili per lo stream



Valori globali disponibili Nella tabella sono elencati i valori globali disponibili. Non è possibile modificare i valori globali qui, ma è possibile cancellare tutti i valori globali per uno stream mediante il pulsante Cancellare tutti i valori disponibile a destra della tabella.

Ricerca di nodi in uno stream

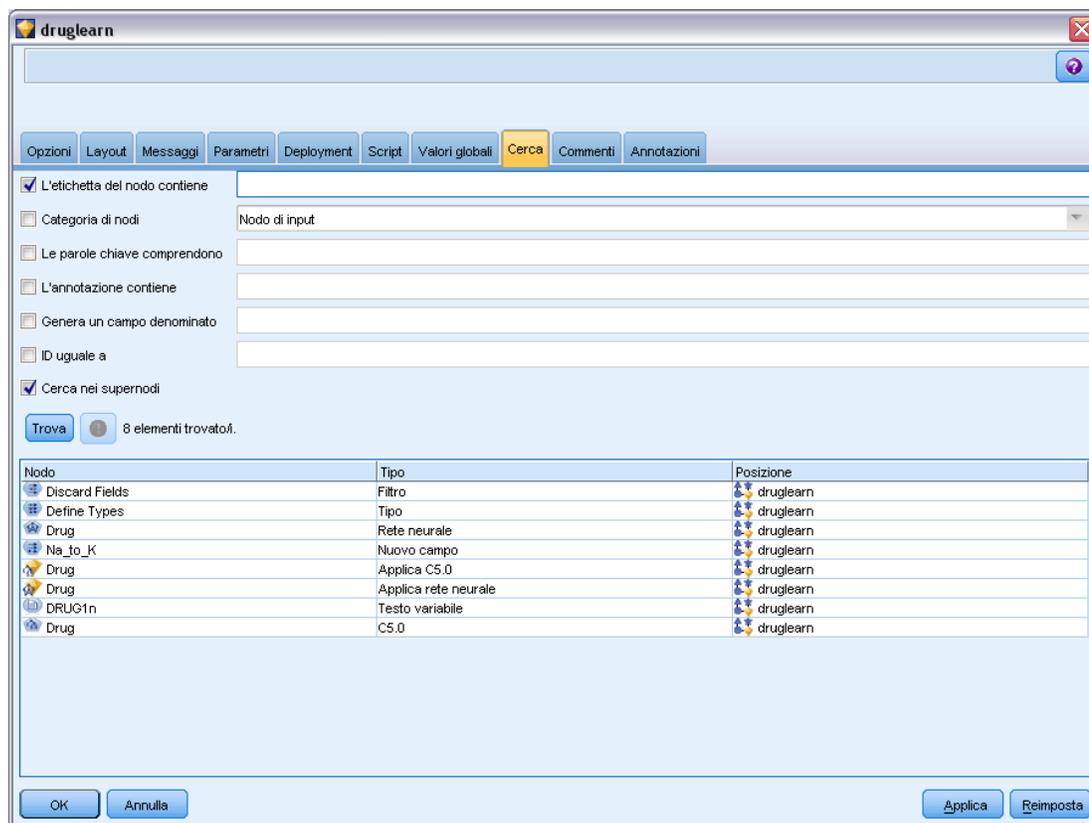
È possibile effettuare ricerche di nodi in uno stream specificando una serie di criteri di ricerca, per esempio nome, categoria e identificatore del nodo. Questa funzione risulta particolarmente utile per stream complessi contenenti un numero elevato di nodi.

Per effettuare ricerche di nodi in uno stream

- ▶ Nel menu File fare clic su Proprietà stream. In alternativa, selezionare lo stream dalla scheda Stream del riquadro dei manager, quindi fare clic con il pulsante destro del mouse e scegliere Proprietà stream dal menu di scelta rapida).
- ▶ Fare clic sulla scheda Cerca.

In alternativa, nel menu Strumenti, fare clic su:
Proprietà stream > Ricerca

Figura 5-24
Ricerca di nodi in uno stream



È possibile specificare più di un'opzione per limitare la ricerca. Utilizzando come criterio di ricerca l'ID del nodo (campo ID uguale a), tuttavia, vengono escluse le altre opzioni.

L'etichetta del nodo contiene. Selezionare questa casella di controllo e immettere l'etichetta di un nodo (completa o parte di essa) per eseguire la ricerca di quel determinato nodo. Le ricerche non distinguono le maiuscole dalle minuscole, mentre le parole multiple vengono considerate come un testo unico.

Categoria di nodi. Selezionare questa casella di controllo e fare clic su una categoria nell'elenco per cercare un determinato tipo di nodo. Nodo di elaborazione indica un nodo della scheda Oper su record o Oper su campi della palette dei nodi; Applica nodo modello fa riferimento a un insieme di modelli.

Le parole chiave comprendono. Selezionare questa casella di controllo e immettere una o più parole chiave complete per cercare i nodi che presentano quel testo nel campo Parole chiave nella scheda Annotazioni della finestra di dialogo del nodo. Il testo delle parole chiave inserite deve essere una corrispondenza esatta. Separare le parole chiave multiple con punto e virgola in modo da cercare alternative (per esempio digitando `proton;neutron` si troveranno tutti i nodi con una di queste due parole chiave. [Per ulteriori informazioni, vedere l'argomento Annotazioni a pag. 77.](#)

L'annotazione contiene. Selezionare questa casella di controllo e immettere una o più parole per cercare i nodi che contengono il testo in questione nell'area di testo principale della scheda Annotazioni della finestra di dialogo del nodo. Le ricerche non distinguono le maiuscole dalle minuscole, mentre le parole multiple vengono considerate come un testo unico. [Per ulteriori informazioni, vedere l'argomento Annotazioni a pag. 77.](#)

Genera un campo denominato. Selezionare questa casella di controllo e immettere il nome di un campo generato (ad esempio, \$C-Cura). Questa opzione consente di cercare nodi Modelli che generano uno specifico campo. Immettere solo un nome di campo, che deve essere una corrispondenza esatta.

ID uguale a. Selezionare questa casella di controllo e immettere l'ID di un nodo per cercare uno specifico nodo con tale identificatore (selezionando questa opzione vengono deselectionate tutte quelle precedenti). Gli ID dei nodi vengono assegnati dal sistema alla creazione del nodo e possono essere utilizzati come riferimenti al nodo per operazioni legate agli script o di automazione. Immettere solo un ID di nodo, che deve essere una corrispondenza esatta. [Per ulteriori informazioni, vedere l'argomento Annotazioni a pag. 77.](#)

Cerca nei supernodi. Questa casella di controllo è selezionata di default. La ricerca viene pertanto effettuata nei nodi sia all'interno sia all'esterno dei Supernodi. Deselezionare la casella di controllo se si desidera che la ricerca venga effettuata solo nei nodi al di fuori dei Supernodi, al livello più elevato dello stream.

Trova. Dopo aver specificato tutte le operazioni desiderate, fare clic su questo pulsante per avviare la ricerca.

I nodi che corrispondono alle opzioni specificate vengono elencati nella parte inferiore della finestra di dialogo. Selezionare un nodo dell'elenco per evidenziarlo nell'area di disegno dello stream.

Rinomina degli stream

Nella scheda Annotazioni della finestra di dialogo delle proprietà dello stream è possibile aggiungere annotazioni descrittive e creare un nome personalizzato per lo stream. Queste opzioni sono particolarmente utili durante la creazione di report per gli stream aggiunti al riquadro dei progetti. [Per ulteriori informazioni, vedere l'argomento Annotazioni a pag. 77.](#)

Descrizioni stream

Per ogni stream che viene creato, IBM® SPSS® Modeler produce una descrizione contenente informazioni sul contenuto dello stream. Questa opzione può essere utile se si sta cercando di capire quali funzioni vengono svolte da uno stream ma non è stato installato SPSS Modeler; per esempio, quando si accede a uno stream mediante IBM® SPSS® Collaboration and Deployment Services.

Figura 5-25
Sezione iniziale della descrizione dello stream

Documento di stream	newschancart_commented
Creato il:	27 gennaio 2006 1.17
Creato da:	llanders
Ultimo salvataggio in data:	28 maggio 2010 20.51
Ultimo salvataggio di:	djones

Contenuti

[Descrizione & Commento](#)
[Informazioni sul calcolo del punteggio](#)
[Informazioni di modellazione](#)

Descrizione & Commento

Commento
This stream predicts whether customers are likely to subscribe to our interactive news service, based on characteristics of customers who have already subscribed
This is the source data about our existing customers
This is where we tell the stream what we are trying to predict

La descrizione dello stream viene visualizzata sotto forma di documento HTML formato da una serie di sezioni.

Informazioni generali sullo stream

Questa sezione contiene il nome dello stream e i dettagli relativi alla creazione e all'ultimo salvataggio dello stream.

Descrizione e commenti

Questa sezione include qualsiasi:

- Annotazioni stream (vedere [Annotazioni a pag. 77](#))
- Commenti non collegati a nodi specifici
- Commenti collegati a nodi dei rami di modellazione e calcolo del punteggio dello stream

Informazioni sul calcolo del punteggio

Questa sezione contiene informazioni, raggruppate sotto diverse intestazioni, relative al ramo del calcolo del punteggio dello stream.

- **Commenti.** Include i commenti legati solo ai nodi nel ramo di calcolo del punteggio.
- **Input.** Elenca i campi di input con i relativi tipi di archiviazione (per esempio, stringa, numero intero, reale, e così via).
- **Output.** Elenca i campi di output, inclusi i campi aggiuntivi generati dal nodo Modelli, nonché i relativi tipi di archiviazione.
- **Parametri.** Elenca i parametri relativi al ramo di calcolo del punteggio dello stream, che possono essere visualizzati o modificati ogni volta che viene calcolato il punteggio del modello. Questi parametri vengono identificati quando si fa clic sul pulsante Parametri di calcolo del punteggio nella scheda Deployment della finestra di dialogo Proprietà stream.
- **Nodo Modello.** Mostra il nome e il tipo di modello (per esempio Rete neurale, C&R Tree e così via). Questo è l'insieme di modelli selezionato per il campo nodo Modello nella scheda Deployment della finestra di dialogo delle proprietà dello stream.
- **Dettagli del modello.** Mostra i dettagli dell'insieme di modelli identificato sotto l'intestazione precedente. Dove possibile, vengono inclusi l'importanza dei predittori e i grafici di valutazione relativi al modello.

Informazioni sul modello

Contiene informazioni relative al ramo di modellazione dello stream.

- **Commenti.** Elenca gli eventuali commenti o annotazioni collegati ai nodi nel ramo di modellazione.
- **Input.** Elenca i campi di input e i relativi ruoli nel ramo di modellazione (sotto forma di valore del ruolo dei campi, per esempio Input, Obiettivo, Suddivisione e così via).
- **Parametri.** Elenca i parametri relativi al ramo di modellazione dello stream, che possono essere visualizzati o modificati ogni volta che il modello viene aggiornato. Questi parametri vengono identificati quando si fa clic sul pulsante Parametri di creazione del modello nella scheda Deployment della finestra di dialogo delle proprietà dello stream.
- **Nodo Modelli.** Mostra il nome e il tipo del nodo Modelli utilizzato per generare o aggiornare il modello.

Anteprima delle descrizioni degli stream

Per visualizzare il contenuto della descrizione di uno stream in un browser Web, è possibile fare clic su un'opzione nella finestra di dialogo delle proprietà dello stream. Il contenuto della descrizione dipende dalle opzioni specificate nella scheda Deployment della finestra di dialogo. [Per ulteriori informazioni, vedere l'argomento Opzioni di deployment degli stream in il capitolo 9 a pag. 186.](#)

Per visualizzare la descrizione di uno stream:

- ▶ Nel menu principale di IBM® SPSS® Modeler, fare clic su:
Strumenti > Proprietà stream > Deployment
- ▶ Impostare il tipo di deployment, il nodo designato per il calcolo del punteggio e i parametri di calcolo del punteggio.
- ▶ Se il tipo di deployment è Aggiornamento modelli, è possibile, se lo si desidera, selezionare un:
 - Nodo Modelli e i parametri di creazione del modello
 - Insieme di modelli sul ramo di calcolo del punteggio dello stream
- ▶ Fare clic sul pulsante Anteprima descrizione stream.

Esportazione di descrizioni stream

Il contenuto della descrizione dello stream può essere esportato in un file HTML.

Per esportare la descrizione di uno stream:

- ▶ Nel menu principale, fare clic su:
File > Esporta descrizione stream
- ▶ Immettere un nome per il file HTML e fare clic su Salva.

Esecuzione degli stream

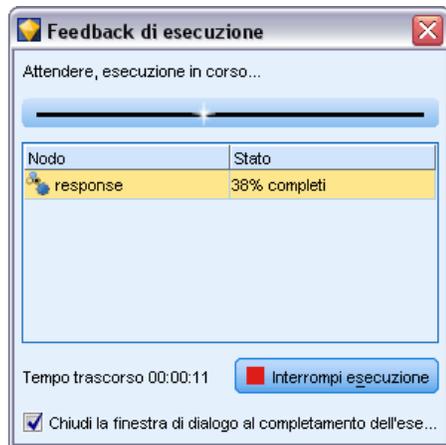
Dopo avere specificato le opzioni richieste per gli stream e avere connesso i nodi, è possibile eseguire lo stream mediante l'elaborazione dei dati attraverso i nodi dello stream. Esistono diversi modi per eseguire uno stream all'interno di IBM® SPSS® Modeler. È possibile:

- Fare clic su Esegui nel menu Strumenti.
- Fare clic su uno dei pulsanti Esegui... sulla barra degli strumenti. Questi pulsanti consentono di eseguire l'intero stream o semplicemente il nodo terminale selezionato. [Per ulteriori informazioni, vedere l'argomento IBM SPSS Modeler Barra degli strumenti in il capitolo 3 a pag. 19.](#)
- Eseguire un singolo stream di dati facendo clic con il pulsante destro del mouse su un nodo terminale e scegliendo Esegui dal menu di scelta rapida.
- Eseguire parte di uno stream di dati facendo clic con il pulsante destro del mouse su un nodo non terminale e scegliendo Esegui da qui dal menu di scelta rapida. Questa operazione determina l'esecuzione soltanto delle operazioni successive al nodo selezionato.

Per interrompere l'esecuzione di uno stream, è possibile fare clic sul pulsante di colore rosso Interrompi disponibile sulla barra degli strumenti oppure fare clic su Interrompi esecuzione nel menu Strumenti.

Se l'esecuzione di uno stream richiede oltre 3 secondi, viene visualizzata la finestra di dialogo Feedback di esecuzione che mostra l'avanzamento dell'esecuzione.

Figura 5-26
Finestra di dialogo *Feedback di esecuzione*



Per alcuni nodi vengono visualizzate informazioni aggiuntive sull'esecuzione dello stream. Per visualizzarle, è necessario selezionare la riga corrispondente nella finestra di dialogo. La prima riga è selezionata automaticamente.

Utilizzo dei modelli

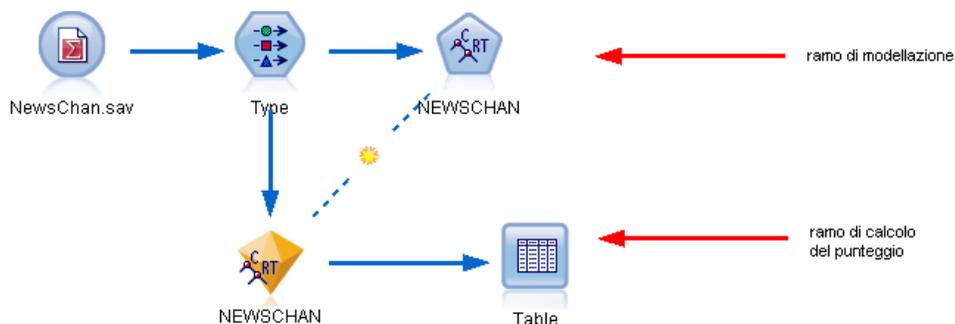
Se in uno stream è presente un nodo Modelli (cioè un nodo contenuto nella scheda Modelli o Modelli database della palette dei nodi) quando si esegue lo stream viene creato un **insieme di modelli**. Un insieme di modelli è un contenitore per un **modello**, ovvero per l'insieme di regole, formule o equazioni che consentono di generare previsioni mediante i dati di input e che costituiscono il fulcro dell'analisi predittiva.

Figura 5-27
Insieme di modelli



Quando si esegue un nodo Modelli, nell'area di disegno dello stream viene inserito un insieme di modelli corrispondente, rappresentato da un'icona che raffigura un diamante dorato. È possibile aprire l'insieme di modelli per esaminarne il contenuto e visualizzare i dettagli del modello. Per visualizzare le previsioni, collegare ed eseguire uno o più nodi terminali, il cui output presenta le previsioni in formato leggibile.

Figura 5-28
Rami di modellazione e di calcolo del punteggio in uno stream



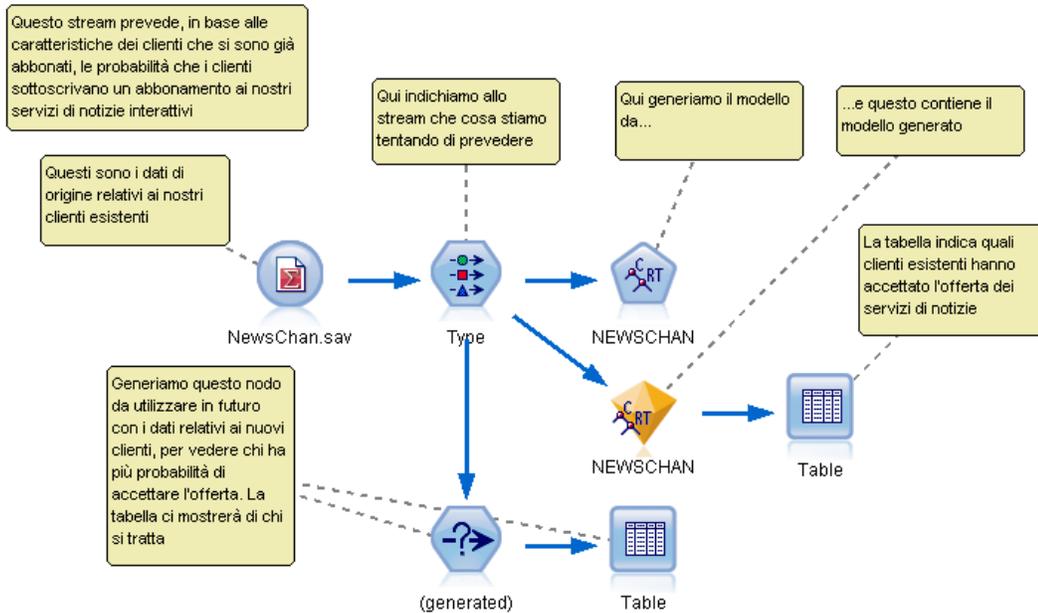
Un tipico stream di modellazione è composto da due rami. Il **ramo di modellazione** contiene il nodo Modelli con i nodi di input e di elaborazione che lo precedono. Il **ramo di calcolo del punteggio** viene creato quando si esegue il nodo di modellazione e contiene l'insieme di modelli e il nodo o i nodi terminali da utilizzare per visualizzare le previsioni.

Per ulteriori informazioni, fare riferimento al documento *IBM® SPSS® Modeler Modeling Nodes guide*.

Aggiunta di commenti e annotazioni a nodi e stream

Può essere necessario descrivere uno stream ad altri utenti della propria azienda. A questo scopo è possibile allegare commenti esplicativi a stream, nodi e insiemi di modelli.

Figura 5-29
Stream con commenti aggiunti



Gli altri utenti possono visualizzare questi commenti sullo schermo oppure è possibile stampare un'immagine dello stream che contenga i commenti.

È possibile elencare tutti i commenti relativi a uno stream o a un Supernodo, cambiare l'ordine dei commenti nell'elenco, modificare il testo di commento e cambiare il colore di primo piano o di sfondo di un commento. [Per ulteriori informazioni, vedere l'argomento Elenco di commenti su stream a pag. 75.](#)

È inoltre possibile aggiungere note sotto forma di annotazioni di testo a stream, nodi e insiemi, per mezzo della scheda Annotazioni della finestra di dialogo delle proprietà di uno stream o di un nodo o da una finestra di insiemi di modelli. Queste note sono visibili solo quando la scheda Annotazioni è aperta; tuttavia, le annotazioni agli stream possono anche essere visualizzate come commenti su schermo. [Per ulteriori informazioni, vedere l'argomento Annotazioni a pag. 77.](#)

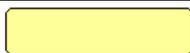
Commenti

I commenti hanno la forma di caselle di testo in cui è possibile inserire testo e commenti senza limitazioni. Un commento può essere indipendente (non collegato a nessun oggetto dello stream) oppure collegato a uno o più nodi o insiemi di modelli dello stream. I commenti indipendenti sono utilizzati solitamente per descrivere le finalità generali dello stream; i commenti collegati descrivono il nodo o l'insieme a cui sono associati. A nodi e insiemi può essere associato più di un commento, mentre lo stream può avere un numero illimitato di commenti indipendenti.

Nota: le annotazioni agli stream possono essere visualizzate anche come commenti su schermo, ma questi non possono essere collegati né a nodi, né a insiemi di modelli. [Per ulteriori informazioni, vedere l'argomento Conversione di annotazioni in commenti a pag. 77.](#)

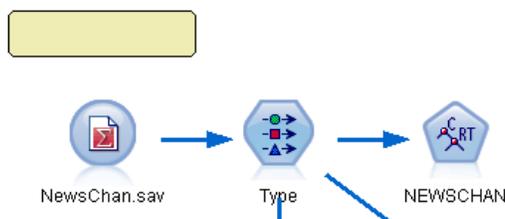
L'aspetto della casella di testo cambia a indicare la modalità corrente del commento (o dell'annotazione visualizzata come commento), come illustrato dalla seguente tabella.

Tabella 5-1
Modalità caselle di testo commenti e annotazioni

Casella di testo commento	Casella di testo annotazione	Mode	Indica	Ottenuto da...
		Modifica	Il commento può essere modificato.	Creare un nuovo commento o annotazione o fare doppio clic su un commento o un'annotazione esistente.
		Ultimo selezionato	Il commento può essere spostato, ridimensionato o eliminato.	Fare clic sullo sfondo dello stream dopo la modifica o fare clic su un commento o un'annotazione esistente.
		Visualizzato	La modifica è terminata.	Dopo avere apportato le modifiche, fare clic su un altro nodo, commento o annotazione.

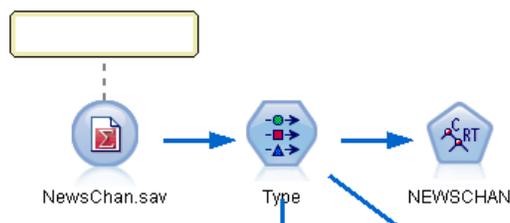
Quando si crea un nuovo commento indipendente, viene inizialmente visualizzato nell'angolo superiore sinistro dell'area di disegno dello stream.

Figura 5-30
Nuovo commento indipendente



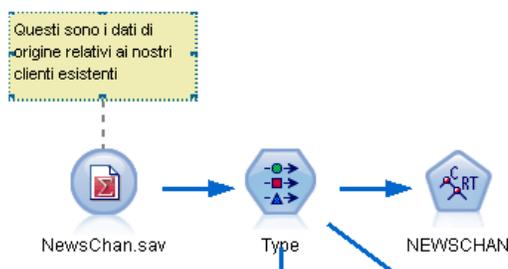
Quando si associa un commento a un nodo o a un insieme, questo viene visualizzato inizialmente al di sopra dell'oggetto dello stream a cui è associato.

Figura 5-31
Nuovo commento associato a nodo



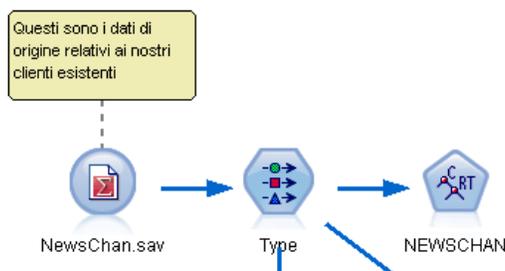
Quando la casella di testo è bianca significa che è possibile immettere il testo. Una volta digitato il testo, fare clic all'esterno della casella di testo. Lo sfondo del commento diventa giallo indicando che l'inserimento del testo è concluso. Il commento rimane selezionato consentendo all'utente di spostarlo, ridimensionarlo o eliminarlo.

Figura 5-32
Commento in modalità di modifica



Quando si fa clic nuovamente, il bordo diventa una linea continua per indicare che la modifica è terminata.

Figura 5-33
Commento completato



Se si fa doppio clic su un commento, la casella di testo entra in modalità di modifica, lo sfondo diventa bianco e il testo del commento può essere modificato.

È inoltre possibile associare commenti ai Supernodi. [Per ulteriori informazioni, vedere l'argomento Annotazione e ridenominazione dei Supernodi in il capitolo 9 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)

Operazioni eseguibili sui commenti

Sui commenti è possibile eseguire una serie di operazioni. È possibile:

- Aggiungere un commento indipendente
- Associare un commento a un nodo o a un insieme
- Modificare un commento
- Ridimensionare un commento
- Spostare un commento
- Disconnettere un commento
- Eliminare un commento
- Visualizzare o nascondere tutti i commenti a uno stream

Per aggiungere un commento indipendente

- ▶ Controllare che nello stream non vi siano selezioni attive.

- ▶ Eseguire una delle seguenti operazioni:
 - Nel menu principale, fare clic su:
Inserisci > Nuovo commento
 - Fare clic con il pulsante destro del mouse sullo sfondo dello stream e scegliere Nuovo commento nel menu di scelta rapida.
 - Fare clic sul pulsante Nuovo commento nella barra degli strumenti.
- ▶ Inserire il testo del commento (o incollarlo dagli Appunti).
- ▶ Fare clic su un nodo dello stream per salvare il commento.

Per associare un commento a un nodo o a un insieme

- ▶ Selezionare uno o più nodi o insiemi nell'area di disegno dello stream.
- ▶ Eseguire una delle seguenti operazioni:
 - Nel menu principale, fare clic su:
Inserisci > Nuovo commento
 - Fare clic con il pulsante destro del mouse sullo sfondo dello stream e scegliere Nuovo commento nel menu di scelta rapida.
 - Fare clic sul pulsante Nuovo commento nella barra degli strumenti.
- ▶ Immettere il testo del commento.
- ▶ Fare clic su un altro nodo dello stream per salvare il commento.
In alternativa è possibile:
- ▶ Inserire un commento indipendente (vedere sezione precedente).
- ▶ Eseguire una delle seguenti operazioni:
 - Selezionare il commento, premere F2 e selezionare il nodo o l'insieme.
 - Selezionare il nodo o l'insieme, premere F2 e selezionare il commento.
 - (Solo per mouse a tre pulsanti) Portare il puntatore del mouse sul commento, tenere premuto il pulsante centrale e trascinare il puntatore sul nodo o sull'insieme, quindi rilasciare il pulsante.

Per associare un commento a un nodo o a un insieme aggiuntivo

Se un commento è già allegato a un nodo o a un insieme o se si trova attualmente a livello dello stream e si desidera associarlo a un altro nodo o insieme, eseguire una delle seguenti operazioni:

- Selezionare il commento, premere F2 e selezionare il nodo o l'insieme.
- Selezionare il nodo o l'insieme, premere F2 e selezionare il commento.
- (Solo per mouse a tre pulsanti) Portare il puntatore del mouse sul commento, tenere premuto il pulsante centrale e trascinare il puntatore sul nodo o sull'insieme, quindi rilasciare il pulsante.

Per modificare un commento esistente

- ▶ Eseguire una delle seguenti operazioni:
 - Fare doppio clic sulla casella di testo del commento.
 - Selezionare la casella di testo e premere Invio.
 - Fare clic con il pulsante destro sulla casella di testo per visualizzare il menu, quindi scegliere Modifica.
- ▶ Modificare il testo del commento. Durante la modifica, è possibile utilizzare i tasti di scelta rapida standard di Windows (per esempio Ctrl+C per copiare il testo). Le altre opzioni disponibili durante la modifica sono elencate nel menu di scelta rapida del commento.
- ▶ Fare clic una volta all'esterno della casella di testo per visualizzare i comandi di ridimensionamento e poi un'altra volta per completare il commento.

Per ridimensionare la casella di testo di un commento

- ▶ Selezionare il commento per visualizzare i comandi di ridimensionamento.
- ▶ Fare clic e trascinare un comando per ridimensionare la casella.
- ▶ Fare clic all'esterno della casella di testo per salvare la modifica.

Per spostare un commento esistente

Se si desidera spostare un commento, ma non gli eventuali oggetti allegati, eseguire le seguenti operazioni:

- Posizionare il puntatore del mouse sul commento, tenere premuto il pulsante sinistro e trascinare il commento nella nuova posizione.
- Selezionare il commento, tenere premuto il tasto Alt e spostare il commento con i tasti freccia.

Se si desidera spostare un commento insieme agli eventuali nodi o insiemi a cui il commento è associato:

- ▶ Selezionare tutti gli oggetti che si desidera spostare.
- ▶ Eseguire una delle seguenti operazioni:
 - Posizionare il puntatore del mouse su uno degli oggetti, tenere premuto il pulsante sinistro e trascinare gli oggetti nella nuova posizione.
 - Selezionare uno degli oggetti, tenere premuto il tasto Alt e spostare gli oggetti utilizzando i tasti freccia.

Per disconnettere un commento da un nodo o da un insieme

- ▶ Selezionare uno o più commenti da disconnettere.

- ▶ Eseguire una delle seguenti operazioni:
 - Premere F3.
 - Fare clic con il pulsante destro del mouse su un commento selezionato e scegliere Disconnetti dal menu.

Per eliminare un commento

- ▶ Selezionare uno o più commenti da eliminare.
- ▶ Eseguire una delle seguenti operazioni:
 - Premere il tasto Canc.
 - Fare clic con il pulsante destro del mouse su un commento selezionato e scegliere Elimina dal menu.

Se il commento era associato a un nodo o a un insieme, viene eliminata anche la linea di connessione.

Se il commento era in origine un'annotazione di stream o di Supernodo che è stata convertita in un commento indipendente, il commento viene eliminato dall'area di disegno, a differenza del testo che viene mantenuto nella scheda Annotazioni dello stream o del Supernodo.

Per visualizzare o nascondere i commenti a uno stream

- ▶ Eseguire una delle seguenti operazioni:
 - Nel menu principale, fare clic su:
Visualizza > Commenti
 - Fare clic sul pulsante Mostra/nascondi commenti nella barra degli strumenti.

Elenco di commenti su stream

È possibile visualizzare un elenco di tutti i commenti che sono stati apportati a uno stream o Supernodo specifici.

Nell'elenco è possibile

- Modificare l'ordine dei commenti
- Modificare il testo del commento
- Cambiare il colore di primo piano o di sfondo di un commento

Elenco di commenti

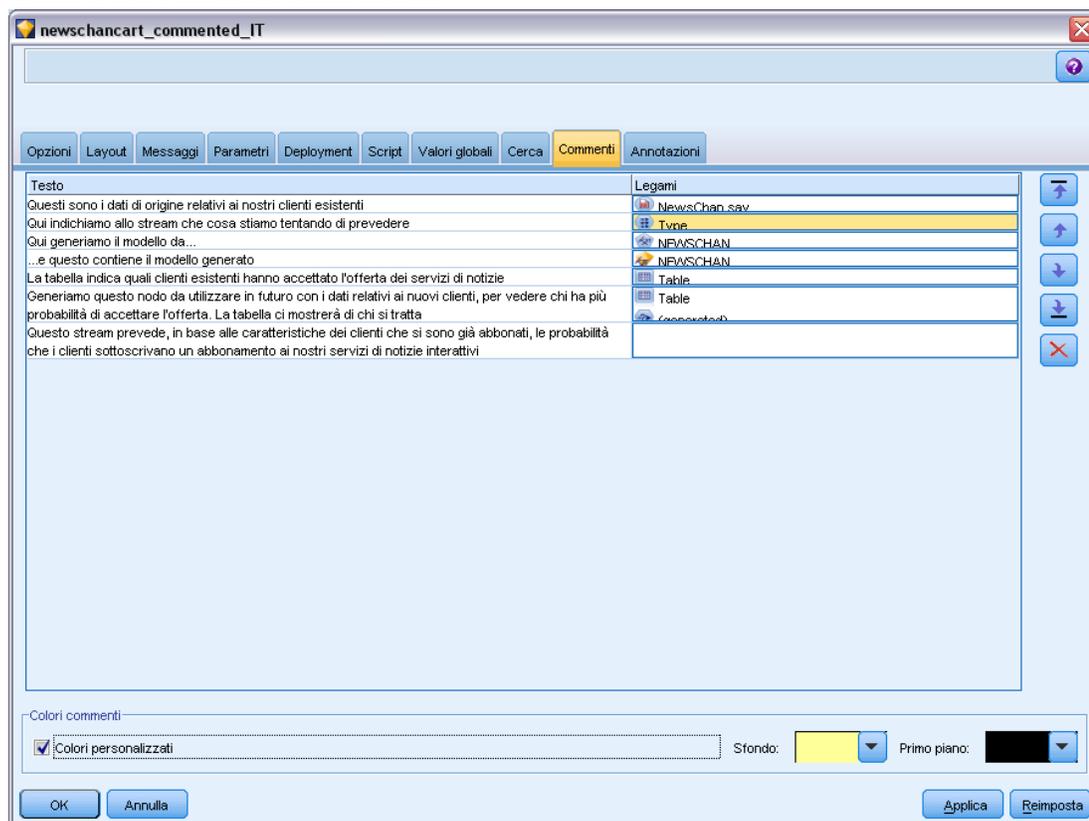
Per elencare tutti i commenti relativi a uno stream, procedere in uno dei seguenti modi:

- Nel menu principale, fare clic su:
Strumenti > Proprietà stream > Commenti

- Fare clic con il pulsante destro del mouse su uno stream nel riquadro dei manager e scegliere Proprietà stream, quindi Commenti.
- Fare clic con il pulsante destro del mouse sullo sfondo di uno stream nell'area di disegno e scegliere Proprietà stream, quindi Commenti.

Figura 5-34

Elenco di tutti i commenti di uno stream



Testo. Testo del commento. Fare doppio clic sul testo per cambiare il campo in una casella di testo modificabile.

Collegamenti. Nome del nodo a cui è collegato il commento. Se questo campo è vuoto, il commento si applica allo stream.

Pulsanti di posizionamento. Consentono di spostare un commento selezionato verso l'alto o verso il basso nell'elenco.

Colori commenti. Per cambiare il colore di primo piano o di sfondo di un commento, selezionare il commento, selezionare la casella di controllo Colori personalizzati, quindi selezionare un colore dall'elenco Sfondo o Primo piano (o da entrambi). Fare clic su Applica, quindi fare clic sullo sfondo dello stream per visualizzare l'effetto della modifica. Fare clic su OK per salvare la modifica.

Conversione di annotazioni in commenti

Le annotazioni apportate a stream e Supernodi possono essere convertite in commenti.

Nel caso degli stream, l'annotazione viene convertita in un commento "autonomo", vale a dire non collegato ad alcun nodo, nell'area di disegno dello stream.

Quando un'annotazione di un Supernodo viene convertita in commento, questo non è collegato al Supernodo nell'area di disegno dello stream, ma è visibile quando si ingrandisce il Supernodo.

Per convertire un'annotazione di stream in commento

- ▶ Fare clic su Proprietà stream nel menu Strumenti. In alternativa, è possibile fare clic con il pulsante destro del mouse su uno stream nel riquadro dei manager e scegliere Proprietà stream.
- ▶ Fare clic sulla scheda Annotazioni.
- ▶ Selezionare la casella di controllo Mostra annotazione come commento.
- ▶ Fare clic su OK.

Per convertire un'annotazione di Supernodo in commento

- ▶ Fare doppio clic sull'icona del Supernodo nell'area di disegno.
- ▶ Fare clic sulla scheda Annotazioni.
- ▶ Selezionare la casella di controllo Mostra annotazione come commento.
- ▶ Fare clic su OK.

Annotazioni

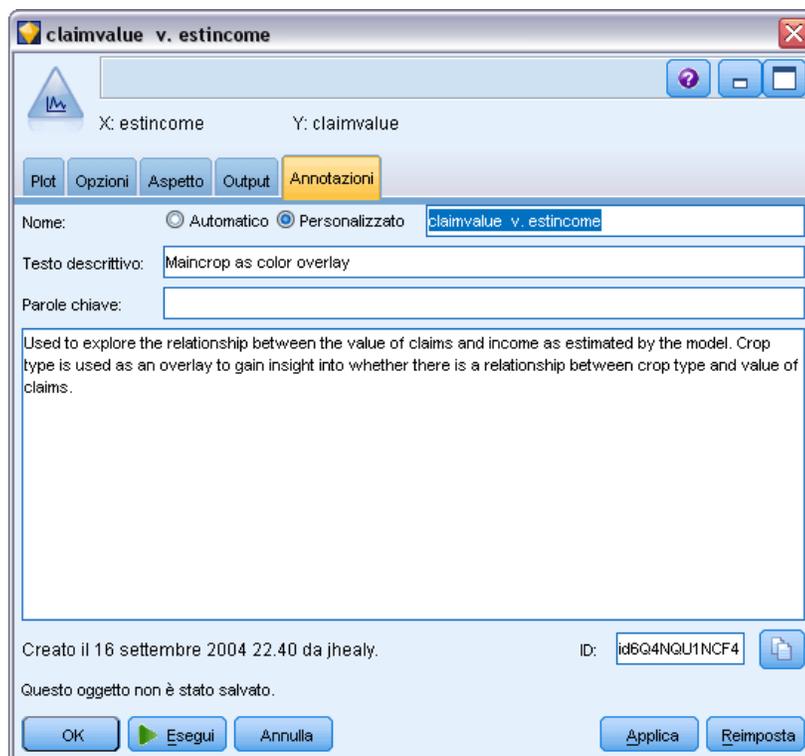
È possibile annotare nodi, stream e modelli in diversi modi. È possibile aggiungere annotazioni descrittive e specificare un nome personalizzato. Queste opzioni sono particolarmente utili durante la creazione di report per gli stream aggiunti al riquadro dei progetti. Per i nodi e gli insiemi di modelli, è inoltre possibile aggiungere un testo descrittivo per distinguere i nodi simili nell'area di disegno dello stream.

Aggiunta di annotazioni

Se si modifica un nodo o un insieme di modelli, verrà aperta una finestra di dialogo a schede contenente una scheda Annotazioni, nella quale è possibile impostare numerose opzioni per le annotazioni. È inoltre possibile aprire direttamente la scheda Annotazioni.

- ▶ Per annotare un nodo o un insieme, fare clic con il pulsante destro del mouse sul nodo o sull'insieme nell'area di disegno dello stream e fare clic su Rinomina e annota. Verrà visualizzata la scheda Annotazioni della finestra di dialogo di modifica.
- ▶ Per annotare uno stream, fare clic su Proprietà stream nel menu Strumenti. In alternativa, è possibile fare clic con il pulsante destro del mouse su uno stream nel riquadro dei manager e scegliere Proprietà stream. Fare clic sulla scheda Annotazioni.

Figura 5-35
Opzioni della scheda Annotazioni



Nome. Selezionare Personalizzato per modificare il nome creato automaticamente o per creare un nome univoco per il nodo da visualizzare nell'area di disegno dello stream.

Testo descrittivo. (Nodi e insiemi di modelli soltanto) Inserire il testo descrittivo visualizzato nell'area di disegno dello stream, che può risultare utile quando si utilizzano molti nodi simili.

Parole chiave. Specificare le parole chiave da utilizzare nei report dei progetti e per la ricerca di nodi in uno stream o per la registrazione di oggetti archiviati nel repository (vedere Informazioni su IBM SPSS Collaboration and Deployment Services Repository a pag. 156). Se sono presenti più parole chiave, è possibile separarle con un punto e virgola, per esempio reddito; tipo raccolto; valore domanda. Gli spazi all'inizio o alla fine di una parola chiave vengono eliminati, per esempio reddito ; tipo raccolto produrrà gli stessi risultati di reddito;tipo raccolto. Gli spazi all'interno delle parole chiave, invece, non vengono troncati. Per esempio, tipo di raccolto con uno spazio e tipo di raccolto con due spazi non si equivalgono.

Nell'area di testo principale è possibile inserire annotazioni più complesse relative alle operazioni o alle scelte eseguite nel nodo. Per esempio, se si condividono e riutilizzano gli stream, è utile inserire note sulle scelte, quali lo scarto di un campo con molti valori vuoti mediante un nodo Filtro. L'annotazione del nodo comporta la memorizzazione di queste informazioni con il nodo. È inoltre possibile includere le annotazioni in un report progetto creato nel riquadro dei progetti. [Per ulteriori informazioni, vedere l'argomento Introduzione ai progetti in il capitolo 11 a pag. 211.](#)

Mostra annotazione come commento. (Annotazioni di stream e Supernodi soltanto) Selezionare questa casella per convertire l'annotazione in un commento libero che sarà visibile nell'area di disegno dello stream. [Per ulteriori informazioni, vedere l'argomento Aggiunta di commenti e annotazioni a nodi e stream a pag. 69.](#)

ID. Visualizza un ID univoco che può essere utilizzato per fare riferimento al nodo a scopo di script o automazione. Questo valore viene generato automaticamente alla creazione del nodo e non subisce variazioni. Si noti inoltre che per evitare confusione con la lettera "O", negli ID dei nodi non vengono utilizzati zero. Utilizzare il pulsante Copia a destra per copiare e incollare l'ID in script o altrove in base alle proprie esigenze. [Per ulteriori informazioni, vedere l'argomento Riferimenti a nodi in il capitolo 3 in IBM SPSS Modeler 14.2 Guida per script e automazione.](#)

Salvataggio degli stream di dati

È possibile salvare gli stream creati in modo da poterli riutilizzare in seguito.

Per salvare uno stream

- ▶ Fare clic su Salva stream o Salva stream con nome nel menu File.
- ▶ Nella finestra di dialogo Salva passare alla cartella in cui si desidera salvare il file di stream.
- ▶ Inserire un nome per lo stream nella casella di testo Nome file.
- ▶ Selezionare Aggiungi a progetto se si desidera aggiungere lo stream salvato al progetto corrente.

Facendo clic su Salva, lo stream verrà memorizzato nella directory specificata con l'estensione **.str*.

File di backup automatici. Ogni volta che viene salvato uno stream, la versione precedente del file viene automaticamente conservata come backup, con un trattino aggiunto alla fine del nome del file, come in *mystream.str-*. Per ripristinare la versione di backup, basta eliminare il trattino e riaprire il file.

Salvataggio degli stati

Oltre agli stream, è possibile salvare anche gli **stati**, che includono il diagramma dello stream correntemente visualizzato e tutti gli insiemi di modelli creati (elencati nella scheda Modelli del riquadro dei manager).

Per salvare uno stato

- ▶ Nel menu File, fare clic su:
State > Salva stato o Salva stato come
- ▶ Nella finestra di dialogo Salva, passare alla cartella in cui si desidera salvare il file di stato.

Facendo clic su Salva, lo stato verrà memorizzato nella directory specificata con l'estensione **.cst*.

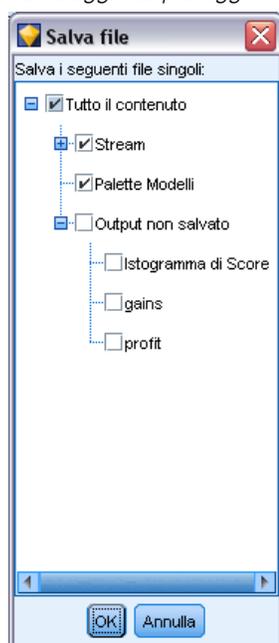
Salvataggio dei nodi

È inoltre possibile salvare un singolo nodo facendo clic con il pulsante destro del mouse sul nodo nell'area di disegno dello stream e fare clic su Salva nodo dal menu di scelta rapida. Utilizzare l'estensione di file **.nod*.

Salvataggio di più oggetti dello stream

Quando si esce da IBM® SPSS® Modeler con più oggetti non salvati, come stream, progetti o insiemi di modelli, verrà richiesto di salvare tali oggetti prima della chiusura completa del software. Se si decide di salvare gli elementi, verrà visualizzata una finestra di dialogo, contenente opzioni per il salvataggio di ogni oggetto.

Figura 5-36
Salvataggio di più oggetti



- ▶ È sufficiente selezionare le caselle di controllo relative agli oggetti che si desidera salvare.
- ▶ Per salvare ogni oggetto nel percorso richiesto, fare clic su OK.

Verrà quindi visualizzata una finestra di dialogo Salva standard per ogni oggetto. Al termine del salvataggio, l'applicazione verrà chiusa, come richiesto in origine.

Salvataggio dell'output

Le tabelle, i grafici e i report generati dai nodi Output di IBM® SPSS® Modeler possono essere salvati in formato oggetto di output (**.cou*).

- ▶ Quando si visualizza l'output che si desidera salvare, nei menu della finestra di output fare clic su: File > Salva

- ▶ Specificare un nome e una posizione per il file di output.
- ▶ Se si desidera, selezionare **Aggiungi file a progetto** nella finestra di dialogo **Salva**, per includere il file nel progetto corrente. [Per ulteriori informazioni, vedere l'argomento Introduzione ai progetti in il capitolo 11 a pag. 211.](#)

In alternativa, è possibile fare clic con il pulsante destro del mouse su uno qualsiasi degli oggetti di output elencati nel riquadro dei manager e selezionare **Salva** dal menu di scelta rapida.

Crittografia e decrittografia di informazioni

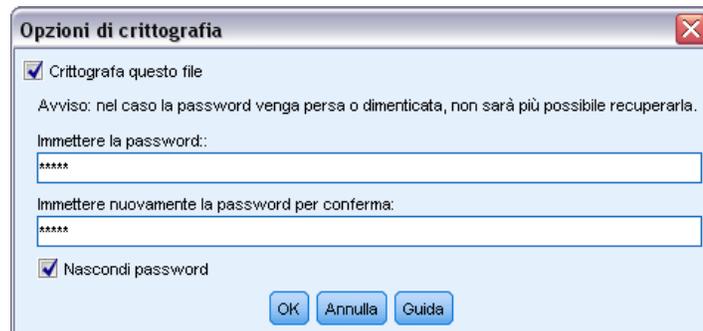
Quando si salva uno stream, un nodo, un progetto, un file di output o un insieme di modelli, è possibile crittografarlo per impedirne l'uso non autorizzato. A questo scopo, è necessario selezionare un'opzione aggiuntiva al momento del salvataggio e aggiungere una password all'elemento che si sta salvando. È possibile crittografare qualsiasi elemento salvato fornendogli una protezione aggiuntiva. Questa funzione non corrisponde alla crittografia SSL utilizzata quando si passano file tra IBM® SPSS® Modeler e IBM® SPSS® Modeler Server.

Quando si tenta di aprire un elemento crittografato, il sistema chiede di immettere la password. Dopo aver immesso la password corretta, l'elemento viene decrittografato automaticamente e si apre come di consueto.

Per crittografare un elemento

- ▶ Nella finestra di dialogo **Salva**, fare clic su **Opzioni** per l'elemento da crittografare. Viene visualizzata la finestra di dialogo **Opzioni di crittografia**.

Figura 5-37
Opzioni di crittografia al salvataggio di un file



- ▶ Selezionare **Crittografa questo file**.
- ▶ Se lo si desidera, per aumentare la protezione selezionare **Nascondi password**. Tutti i caratteri immessi verranno visualizzati come una serie di punti.
- ▶ Immettere la password. *Avviso:* se si dimentica la password, non è possibile aprire il file o il modello.
- ▶ Se si seleziona **Nascondi password**, immettere nuovamente la password per confermarla.
- ▶ Scegliere **OK** per ritornare alla finestra di dialogo **Salva**.

Nota: se si salva una copia di un elemento protetto mediante crittografia, il nuovo elemento viene automaticamente salvato in formato crittografato utilizzando la password originale, a meno che non si modifichino le impostazioni della finestra di dialogo Opzioni di crittografia.

Caricamento di file

È possibile ricaricare numerosi oggetti salvati in IBM® SPSS® Modeler:

- Stream (.str)
- Stati (.cst)
- Modelli (.gm)
- Palette Modelli (.gen)
- Nodi (.nod)
- Output (.cou)
- Progetti (.cpj)

Apertura di nuovi file

È possibile caricare gli stream direttamente dal menu File.

- ▶ Nel menu File, fare clic su Apri stream.

Tutti gli altri tipi di file possono essere aperti mediante i comandi dei sottomenu disponibili nel menu File. Per esempio, per caricare un modello, nel menu File fare clic su:

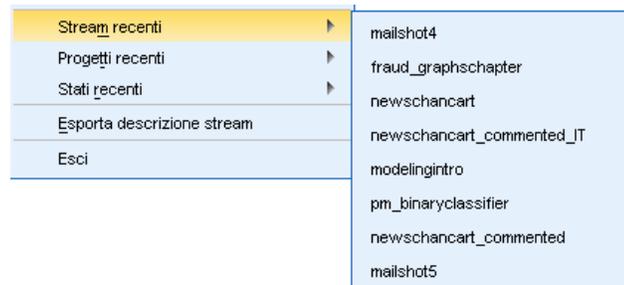
Modelli > Apri modello o Carica palette modelli

Apertura di file utilizzati di recente

Per caricare rapidamente i file utilizzati di recente, utilizzare le opzioni disponibili nella parte inferiore del menu File.

Figura 5-38

Apertura di opzioni utilizzate di recente dal menu File



Selezionare Stream recenti, Progetti recenti o Stati recenti per espandere un elenco di file utilizzati di recente.

Mappatura degli stream di dati

Lo strumento per la mappatura dei dati consente di connettere una nuova sorgente dati a uno stream preesistente. Lo strumento per la mappatura non solo imposterà la connessione, ma consentirà anche di specificare in quale modo i campi della nuova sorgente sostituiranno quelli dello stream esistente. Anziché ricreare un intero stream di dati per una nuova sorgente dati, sarà sufficiente connettersi a uno stream esistente.

Lo strumento per la mappatura consente di unire due frammenti di stream avendo la certezza che tutti i nomi dei campi essenziali corrispondano. La mappatura dei dati consiste semplicemente nella creazione di un nuovo nodo Filtro, che crea una corrispondenza dei campi appropriati rinominandoli.

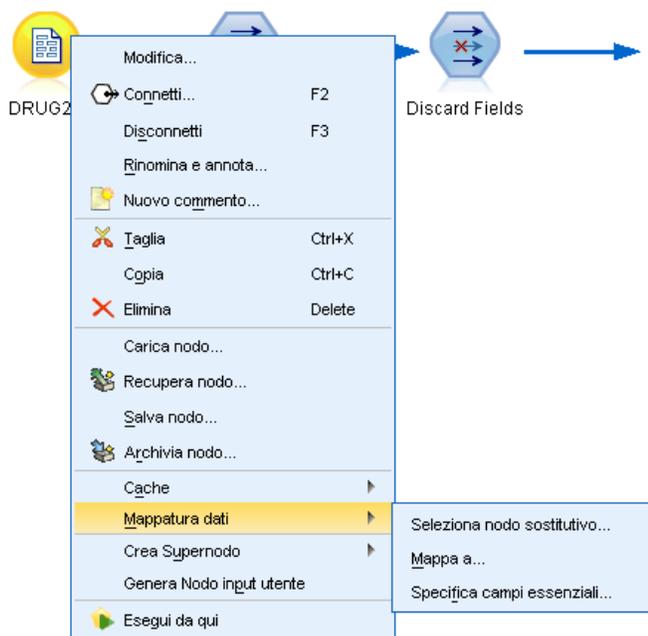
Esistono due metodi equivalenti per mappare i dati:

Seleziona nodo sostitutivo. Questo metodo parte dal nodo che deve essere sostituito. Innanzitutto, fare clic con il pulsante destro del mouse sul nodo da sostituire; quindi, utilizzando l'opzione Mappatura dati > Seleziona nodo sostitutivo del menu di scelta rapida, selezionare il nodo con il quale sostituirlo.

Mappa a. Questo metodo parte dal nodo che deve essere inserito nello stream. Innanzitutto, fare clic con il pulsante destro del mouse sul nodo da inserire; quindi, utilizzando l'opzione Mappatura dati > Mappa a del menu di scelta rapida, selezionare il nodo in cui inserirlo. Questo metodo è particolarmente utile per la mappatura su un nodo finale. *Nota:* non è possibile eseguire la mappatura su nodi Unione o Accodamento. In questi casi è sufficiente connettere lo stream al nodo Unione nel modo normale.

Figura 5-39

Selezione delle opzioni di mappatura dei dati



La mappatura dei dati è strettamente integrata con la creazione degli stream. Se si tenta di connettersi a un nodo che ha già una connessione, sarà possibile sostituire la connessione o la mappatura al nodo specifico.

Mappatura dei dati su un modello

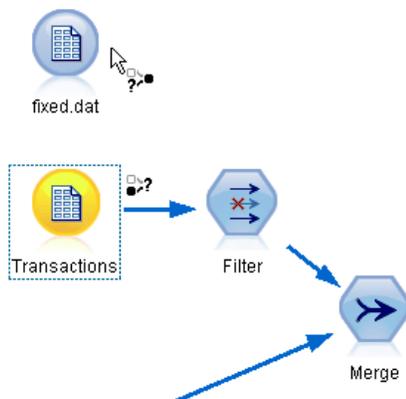
Per sostituire l'origine dati di uno stream di modelli con un nuovo nodo di input che consente di inserire i dati in IBM® SPSS® Modeler, è consigliabile utilizzare l'opzione **Seleziona nodo sostitutivo** dal menu di scelta rapida **Mappatura dati**. L'opzione è disponibile per tutti i nodi a eccezione dei nodi **Unione**, **Aggregazione** e tutti i nodi finali. Se si utilizza lo strumento per la mappatura dei dati per eseguire questa operazione, si avrà la certezza che esista una corrispondenza esatta tra i campi delle operazioni dello stream esistenti e della nuova sorgente dati. I passaggi seguenti forniscono una panoramica del processo di mappatura dei dati.

Passo 1: specificare i campi essenziali nel nodo di input originale. Affinché le operazioni dello stream vengano eseguite correttamente, è necessario specificare i campi essenziali. [Per ulteriori informazioni, vedere l'argomento Definizione dei campi essenziali a pag. 86.](#)

Passo 2: aggiungere la nuova sorgente dati nell'area di disegno dello stream. Utilizzando uno dei nodi di input, inserire i nuovi dati sostitutivi.

Passo 3: sostituire il nodo di input del modello. Utilizzando l'opzione **Mappatura dati** del menu di scelta rapida per il nodo di input del modello, fare clic su **Seleziona nodo sostitutivo**, quindi selezionare il nodo di input per i dati sostitutivi.

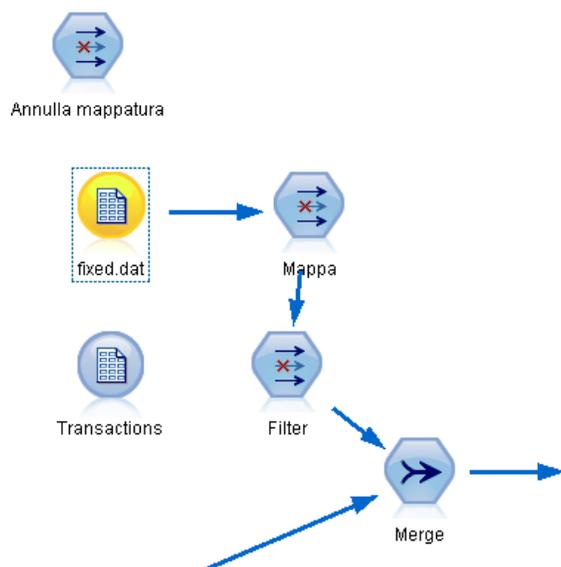
Figura 5-40
Selezione di un nodo di input sostitutivo



Passo 4: controllare i campi mappati. Nella finestra di dialogo visualizzata, controllare che venga eseguita correttamente la mappatura dei campi dalla sorgente dati sostitutiva allo stream. I campi essenziali che non sono stati mappati vengono visualizzati in rosso. Questi campi vengono utilizzati nelle operazioni dello stream e devono essere sostituiti con un campo simile nella nuova sorgente dati affinché le operazioni downstream possano essere eseguite correttamente. [Per ulteriori informazioni, vedere l'argomento Analisi dei campi mappati a pag. 87.](#)

Dopo che si è controllato che tutti i campi essenziali siano stati mappati correttamente, la sorgente dati precedente verrà disconnessa e la nuova sorgente dati verrà connessa allo stream mediante un nodo Filtro denominato *Mappa*. Questo nodo Filtro gestisce la mappatura effettiva dei campi nello stream. Nell'area di disegno dello stream viene inoltre inserito un nodo Filtro denominato *Annulla mappatura*, che consente di invertire la mappatura del nome di campo aggiungendolo allo stream. I campi mappati verranno annullati, ma sarà necessario modificare i nodi downstream per riselectare i campi e le sovrapposizioni.

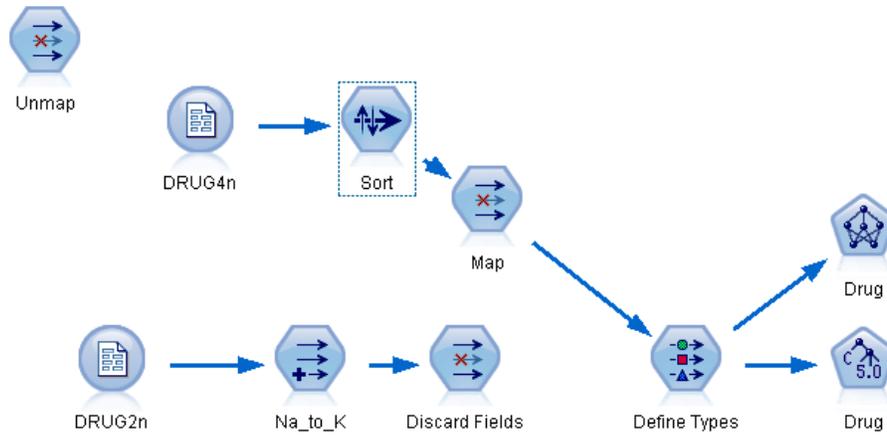
Figura 5-41
Nuova sorgente dati mappata sullo stream



Mappatura tra stream

Questo metodo di mappatura dei dati è simile alla connessione dei nodi, ma non richiede l'impostazione iniziale dei campi essenziali. Con questo metodo, si esegue semplicemente la connessione da uno stream a un altro utilizzando l'opzione *Mappa* a del menu di scelta rapida *Mappatura dati*. Questo tipo di mappatura dei dati è utile per la mappatura su nodi finali e per le operazioni di copia e incolla tra stream. *Nota:* se si utilizza l'opzione *Mappa* a, non è possibile eseguire la mappatura ai nodi Unione, Accodamento e a tutti i nodi di input.

Figura 5-42
Mappatura di uno stream dal relativo nodo Ordina al nodo Tipo di un altro stream



Per mappare i dati tra stream

- ▶ Fare clic con il pulsante destro del mouse sul nodo da utilizzare per la connessione al nuovo stream.
- ▶ Nel menu, fare clic su:
Mappatura dati > Mappa a
- ▶ Utilizzare il cursore per selezionare un nodo di destinazione nello stream obiettivo.
- ▶ Nella finestra di dialogo visualizzata, controllare che i campi corrispondano e fare clic su OK.

Definizione dei campi essenziali

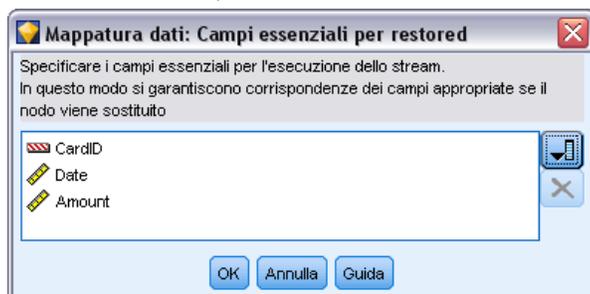
Quando si esegue la mappatura su uno stream esistente, i campi essenziali vengono in genere specificati dall'autore dello stream. Questi campi essenziali indicano se un campo specifico viene utilizzato nelle operazioni downstream. Per esempio, è possibile che lo stream esistente crei un modello che utilizza un campo denominato *Tasso di abbandono*. In questo stream, *Tasso di abbandono* è un campo essenziale perché non è possibile creare il modello senza di esso. Allo stesso modo, i campi utilizzati nei nodi di manipolazione quali un nodo Nuovo campo sono necessari per derivare il nuovo campo. L'impostazione esplicita di questi campi come campi essenziali garantisce che ad essi vengano mappati i campi corretti della nuova sorgente dati. Se non si mappano i campi obbligatori, verrà visualizzato un messaggio di errore. Se si stabilisce che alcune manipolazioni o alcuni nodi di output non sono necessari, è possibile eliminare i nodi dallo stream e rimuovere i campi appropriati dall'elenco dei campi essenziali.

Per impostare i campi essenziali

- ▶ Fare clic con il pulsante destro del mouse sul nodo di input dello stream di modelli che verrà sostituito.

- Nel menu, fare clic su:
Mappatura dati > Specifica campi essenziali

Figura 5-43
Definizione dei campi essenziali



- È possibile aggiungere o rimuovere campi dall'elenco mediante il pulsante di selezione dei campi. Per aprire lo strumento per la selezione dei campi, fare clic sull'icona a destra dell'elenco di campi.

Analisi dei campi mappati

Dopo avere selezionato il punto nel quale uno stream di dati o una sorgente dati verranno mappati a un altro, verrà aperta una finestra di dialogo nella quale è possibile selezionare i campi da mappare o verificare che la mappatura di default di sistema sia corretta. Se i campi essenziali per lo stream o la sorgente dati sono stati impostati e non corrispondono, verranno visualizzati in rosso. I campi non mappati dalla sorgente dati passeranno attraverso il nodo Filtro senza essere modificati. Si noti tuttavia che è possibile mappare anche i campi non essenziali.

Figura 5-44
Selezione dei campi per la mappatura



Originali. Elenca tutti i campi del modello o dello stream esistente, ovvero tutti i campi a valle disponibili. I campi della nuova sorgente dati verranno mappati a questi campi.

Mappati. Elenca i campi selezionati per la mappatura ai campi del modello. Si tratta dei campi il cui nome potrebbe essere modificato in modo da corrispondere a quello dei campi originali utilizzati nelle operazioni dello stream. Se si fa clic sulla cella di un campo, verrà attivato un elenco dei campi disponibili.

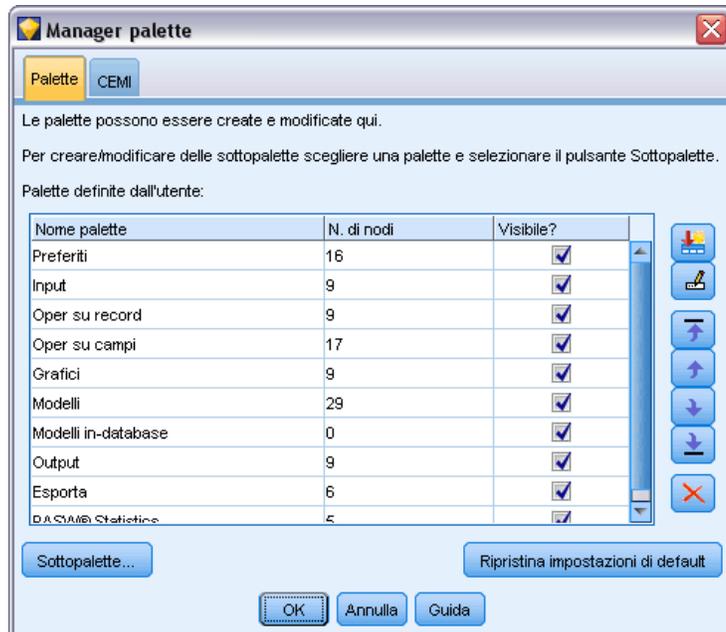
Se non si è certi dei campi da mappare, può essere utile analizzare in modo approfondito la sorgente dati prima di procedere alla mappatura. Per esempio, è possibile utilizzare la scheda Tipo nel nodo di input per esaminare un riepilogo dei dati di input.

Suggerimenti e tasti di scelta rapida

Per rendere più rapido e semplice l'uso di Clementine, si consiglia di familiarizzare con i tasti di scelta rapida e i suggerimenti indicati di seguito.

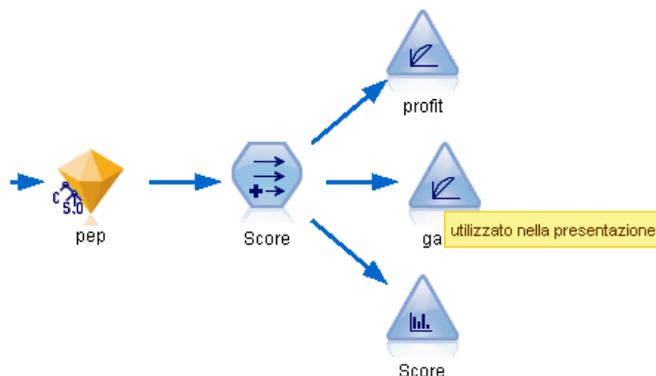
- **Creare stream rapidamente facendo doppio clic.** Fare semplicemente doppio clic su un nodo della palette per aggiungerlo e connetterlo allo stream corrente.
- **Utilizzare le combinazioni dei tasti di scelta rapida per selezionare i nodi a valle.** Premere Ctrl+Q e Ctrl+W per attivare/disattivare la selezione di tutti i nodi a valle.
- **Utilizzare i tasti di scelta rapida per connettere e disconnettere i nodi.** Quando si seleziona un nodo nell'area di disegno, premere F2 per iniziare una connessione, premere Tab per passare al nodo richiesto e quindi Maiusc+Barra spaziatrice per completare la connessione. Utilizzare F3 per disconnettere tutti gli input e gli output al nodo selezionato.
- **Personalizzare la scheda Palette nodi con i nodi preferiti.** Nel menu Strumenti, fare clic su Gestisci palette per aprire una finestra di dialogo e aggiungere, rimuovere o spostare i nodi presenti nella Palette nodi.

Figura 5-45
Manager palette



- **Rinominare i nodi e aggiungere le descrizioni.** In ogni finestra di dialogo di nodo è disponibile una scheda Annotazioni in cui è possibile specificare un nome personalizzato per i nodi nell'area di disegno nonché descrizioni che consentano di organizzare lo stream. È inoltre possibile includere annotazioni più complesse per tenere traccia dell'avanzamento, salvare i dettagli del processo e indicare le decisioni aziendali necessarie o già raggiunte.

Figura 5-46
Descrizione e nome di nodo personalizzato



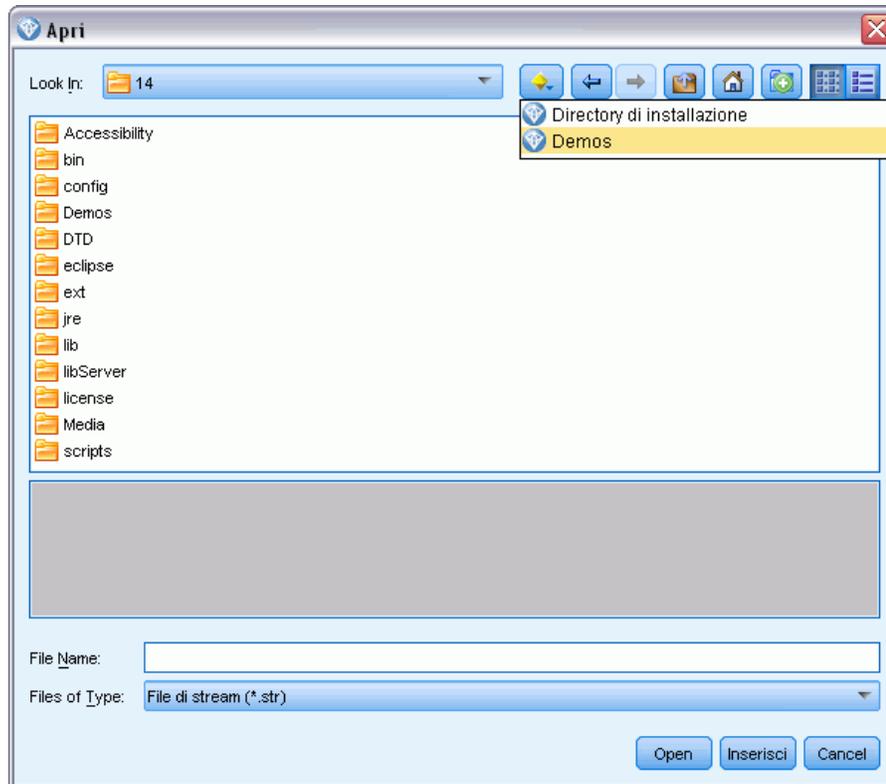
- Inserire automaticamente i valori in un'espressione CLEM.** Tramite il generatore di espressioni, a cui è possibile accedere da diverse finestre di dialogo quali quelle dei nodi Nuovo campo e Riempimento, è possibile inserire automaticamente i valori dei campi in un'espressione CLEM. Fare clic sul pulsante dei valori nel generatore di espressioni per scegliere tra i valori di campo esistenti.

Figura 5-47
Pulsante dei valori



- Cercare file rapidamente.** Quando si cercano i file in una finestra di dialogo Apri, utilizzare l'elenco File (fare clic sul pulsante rombo giallo) per accedere alle directory utilizzate in precedenza nonché alle directory di default di IBM® SPSS® Modeler. Utilizzare i pulsanti avanti e indietro per scorrere le directory a cui si è effettuato l'accesso.

Figura 5-48
 Selezione della cartella Demos dall'elenco delle directory utilizzate di recente



- **Ridurre lo spazio della finestra di output.** È possibile chiudere ed eliminare rapidamente l'output utilizzando il pulsante con la X rossa nell'angolo in alto a destra di tutte le finestre di output. In questo modo nella scheda Output del riquadro dei manager resteranno visualizzati solo i risultati interessanti.

Per il software è disponibile una gamma completa di tasti di scelta rapida. [Per ulteriori informazioni, vedere l'argomento Accesso facilitato mediante la tastiera in l'appendice A a pag. 252.](#)

Operazioni che è possibile effettuare

- Trascinare e selezionare un gruppo di nodi nell'area di disegno dello stream utilizzando il mouse.
- Copiare e incollare nodi da uno stream a un altro.
- Accedere alla Guida in linea da ogni finestra di dialogo e finestra di output.
- Visualizzare gli argomenti della Guida in linea relativi a CRISP-DM (Cross-Industry Standard Process for Data Mining). Fare clic su Guida in linea di CRISP-DM nel menu ?.

Gestione dei valori mancanti

Panoramica sui valori mancanti

Durante la fase Data Preparation del data mining è in genere opportuno sostituire i valori mancanti. I **valori mancanti** sono valori dell'insieme di dati sconosciuti, non raccolti o inseriti non correttamente e che in genere non sono validi per i relativi campi. Per esempio il campo *Sesso* dovrebbe contenere valori *M* e *F*. Se invece nel campo sono presenti i valori *Y* o *Z*, si può presumere che tali valori non siano validi e pertanto verranno interpretati come valori vuoti. Anche un valore negativo per il campo *Età* è privo di significato e va interpretato come un valore vuoto. Accade di frequente che in un questionario vengano inseriti valori chiaramente errati di questo tipo oppure che i campi vengano lasciati vuoti, a indicare che l'utente non ha voluto rispondere. A volte può essere utile analizzare i valori vuoti in modo più approfondito, per determinare se il rifiuto di fornire una risposta (per esempio relativa alla propria età) possa essere utile per prevedere un risultato specifico.

Alcune tecniche di modellazione consentono di gestire meglio i dati mancanti rispetto ad altre. Per esempio, C5.0 e Apriori sono particolarmente utili per i valori dichiarati esplicitamente come mancanti in un nodo Tipo. Altre tecniche non funzionano altrettanto bene con i valori mancanti, richiedono tempi di addestramento maggiori e i modelli risultanti sono meno precisi.

Esistono vari tipi di valori mancanti riconosciuti da IBM® SPSS® Modeler:

- **Valori nulli o mancanti di sistema.** Sono valori diversi da stringhe che vengono lasciati vuoti nel database o nel file di origine e che non sono specificati in modo esplicito come “mancanti” in un nodo Tipo o in un nodo di input. I valori mancanti di sistema vengono visualizzati come \$null\$. In SPSS Modeler le stringhe vuote non vengono considerate valori nulli, ma possono essere considerate valori nulli da alcuni database.
- **Stringhe vuote e spazi bianchi.** I valori di stringhe vuote e spazi bianchi (stringhe prive di caratteri visibili) sono trattati in modo diverso rispetto ai valori nulli. Nella maggior parte dei casi, le stringhe vuote vengono considerate equivalenti agli spazi bianchi. Per esempio, se si seleziona l'opzione che consente di trattare gli spazi bianchi come valori vuoti in un nodo Tipo o di input, questa impostazione risulta valida anche per le stringhe vuote.
- **Valori vuoti o mancanti definiti dall'utente.** Sono valori quali sconosciuto, 99 o -1 che vengono definiti in modo esplicito come mancanti in un nodo Tipo o in un nodo di input. Se lo si desidera, si può anche decidere di considerare i valori nulli e gli spazi bianchi come valori vuoti: in questo modo è possibile evidenziarli in modo da poter essere elaborati diversamente ed esclusi dalla maggior parte dei calcoli. Per esempio, è possibile utilizzare la funzione @BLANK per trattare questi valori, insieme ad altri tipi di valori mancanti, come valori vuoti. [Per ulteriori informazioni, vedere l'argomento Utilizzo della finestra di dialogo Valori in il capitolo 4 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)

Figura 6-1
Definizione dei valori mancanti per una variabile continua

Letture di dati misti. Quando si leggono campi con archiviazione numerica (numeri interi, reali, ora, timestamp o data), gli eventuali valori non numerici vengono impostati come *valori nulli* o *valori mancanti di sistema* poiché, a differenza di altre applicazioni, Clementine non consente di utilizzare tipi di archiviazione misti all'interno dello stesso campo. Per evitare questo inconveniente, è necessario modificare opportunamente il tipo di archiviazione nel nodo di input o nell'applicazione esterna per fare in modo che tutti i campi contenenti dati misti siano letti come stringhe. [Per ulteriori informazioni, vedere l'argomento Impostazione dell'archiviazione e della formattazione dei campi in il capitolo 2 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)

Letture di stringhe vuote da Oracle. Durante la lettura o la scrittura in un database Oracle, tenere presente che, a differenza di SPSS Modeler e della maggioranza degli altri database, Oracle gestisce e archivia i valori di stringhe vuote come valori nulli. Questo significa che gli stessi dati estratti da un database Oracle potrebbero comportarsi in modo diverso rispetto a quando vengono estratti da un file o da un altro database e i dati potrebbero restituire risultati diversi.

Gestione dei valori mancanti

Il tipo di trattamento scelto per i valori mancanti dipende da conoscenze specifiche di business o di dominio. La rimozione dei valori vuoti dall'insieme di dati in uso consente di ridurre i tempi di addestramento e di ottenere una maggiore precisione. La presenza dei valori vuoti può d'altra

parte determinare nuove opportunità aziendali o fornire informazioni aggiuntive. Per stabilire la tecnica migliore da utilizzare, è necessario considerare i seguenti aspetti dei dati con cui si lavora:

- Dimensione dell'insieme di dati
- Numero dei campi contenenti valori vuoti
- Quantità di informazioni mancanti

In linea generale, è possibile adottare due approcci:

- Escludere i campi o i record con valori mancanti
- Assegnare, sostituire o forzare i valori mancanti mediante una serie di metodi diversi

Entrambi gli approcci possono essere in gran parte automatizzati mediante il nodo Esplora. Per esempio, è possibile generare un nodo Filtro che escluda i campi che presentano troppi valori mancanti per essere utili ai fini della modellazione e generare un Supernodo che assegni i valori mancanti ad alcuni o tutti i campi restanti. Proprio in questi casi viene dimostrata la vera efficacia dell'esplorazione, che consente non solo di valutare lo stato corrente dei dati, ma anche di agire sulla base di tale valutazione. [Per ulteriori informazioni, vedere l'argomento Preparazione dei dati per l'analisi \(Esplorazione dei dati\) in il capitolo 8 in Guida alle applicazioni di IBM SPSS Modeler 14.2.](#)

Gestione dei record con valori mancanti

Se la maggior parte dei valori mancanti è concentrata in un numero ridotto di record, è possibile semplicemente escludere quei record. Per esempio, le banche in genere dispongono di informazioni dettagliate sui clienti che hanno richiesto prestiti, ma se per l'approvazione dei prestiti al personale della banca vengono applicate condizioni meno restrittive, è probabile che nei relativi dati siano presenti molti campi vuoti. In questo caso, è possibile gestire i valori mancanti in due modi:

- È possibile utilizzare un nodo Selezione per rimuovere i record relativi al personale.
- Se l'insieme di dati è grande, è possibile scartare tutti i record con valori vuoti. [Per ulteriori informazioni, vedere l'argomento Selezione di record con dati mancanti in il capitolo 6 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)

Gestione dei campi con valori mancanti

Se la maggior parte dei valori mancanti si concentra in un numero ristretto di campi, è possibile gestire questi valori a livello di campo anziché a livello di record. Ciò consente inoltre di verificare l'importanza relativa di campi specifici prima di scegliere un approccio per la gestione dei valori mancanti. Se un campo è irrilevante per la creazione di modelli, probabilmente non vale la pena di conservarlo, indipendentemente dal suo numero di valori mancanti.

Per esempio, è possibile che una società di ricerche di mercato raccolga i dati mediante un questionario generale che include 50 domande. Due domande sono relative all'età e alle opinioni politiche, ovvero informazioni che la maggior parte delle persone preferisce non fornire. In questo caso, i campi *Età* e *Tendenze politiche* contengono numerosi valori mancanti.

Livello di misurazione dei campi

Per stabilire il metodo da utilizzare, è opportuno considerare anche il livello di misurazione dei campi contenenti i valori mancanti.

Campi numerici. Per i tipi di campi numerici, per esempio *Continuo*, è sempre consigliabile eliminare tutti i valori non numerici prima di creare un modello, perché molti modelli non funzioneranno se i campi numerici contengono valori vuoti.

Campi categoriali. Per i campi categoriali, per esempio i campi *nominali* e *flag*, la modifica dei valori mancanti non è necessaria, ma consentirà comunque di aumentare la precisione del modello. Per esempio, un modello che utilizza il campo *Sesso* funzionerà anche se sono presenti valori senza significato, come *Y* e *Z*, ma la rimozione di tutti i valori diversi da *M* e *F* aumenterà la precisione del modello.

Screening o rimozione di campi

I campi con troppi valori mancanti possono essere sottoposti a screening in diversi modi:

- I nodi Esplora possono essere utilizzati per filtrare i campi in base alla qualità. [Per ulteriori informazioni, vedere l'argomento Filtraggio dei campi con dati mancanti in il capitolo 6 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)
- Si può utilizzare un nodo Selezione funzioni per escludere i campi i cui valori mancanti superano una percentuale specifica e per ordinare i campi in base all'importanza rispetto a un obiettivo specificato. [Per ulteriori informazioni, vedere l'argomento Nodo Selezione funzioni in il capitolo 4 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)
- Anziché rimuovere i campi, è possibile utilizzare un nodo Tipo per impostare il ruolo dei campi su Nessuno. In questo modo, i campi verranno mantenuti nell'insieme di dati, ma saranno esclusi dai processi di modellazione.

Assegnazione o riempimento dei valori mancanti

Se il numero dei valori mancanti è ridotto, può essere utile inserire valori sostitutivi al posto dei valori vuoti. Questa operazione può essere eseguita dal report del nodo Esplora, che consente di specificare secondo necessità le opzioni per i campi specifici e di generare un Supernodo che assegna i valori mediante diversi metodi. Questo è il metodo più flessibile, che consente inoltre di specificare la gestione di grandi quantità di campi in un unico nodo.

Per l'immissione di valori mancanti sono disponibili i seguenti metodi:

Fissi. Sostituisce con un valore fisso (la media di un campo, il valore intermedio dell'intervallo o una costante specificata dall'utente).

Casuale. Sostituisce con un valore casuale in base a una distribuzione normale o uniforme.

Espressione. Consente di specificare un'espressione personalizzata. Per esempio, è possibile sostituire valori con una variabile globale creata dal nodo Globali.

Algoritmo. Sostituisce con un valore previsto da un modello basato sull'algoritmo C&RT. Per ogni campo immesso utilizzando questo metodo, ci sarà un modello C&RT separato e un nodo Riempimento che sostituisce i valori vuoti e nulli con il valore previsto dal modello. Viene quindi utilizzato un nodo Filtro per rimuovere i campi di previsione generati dal modello.

Per ulteriori informazioni, vedere l'argomento [Assegnazione dei valori mancanti in il capitolo 6 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)

In alternativa, per forzare i valori per campi specifici, è possibile utilizzare un nodo Tipo per fare in modo che i tipi di campi includano solo valori validi e quindi impostare su Forza la colonna *Controllo* dei campi per i quali è necessario sostituire i valori mancanti. Per ulteriori informazioni, vedere l'argomento [Nodo Tipo in il capitolo 4 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)

Funzioni CLEM per i valori mancanti

Sono disponibili numerose funzioni per la gestione dei valori mancanti. Per scartare o sostituire i valori mancanti, nei nodi Selezione e Riempimento vengono spesso utilizzate le seguenti funzioni:

- `count_nulls(LIST)`
- `@BLANK(FIELD)`
- `@NULL(FIELD)`
- `undef`

La combinazione delle funzioni @ e della funzione @FIELD consente di individuare i valori vuoti o nulli in uno o più campi. È possibile limitarsi a contrassegnare i campi che contengono valori vuoti o nulli, riempirli con valori sostitutivi o utilizzarli per molte altre operazioni.

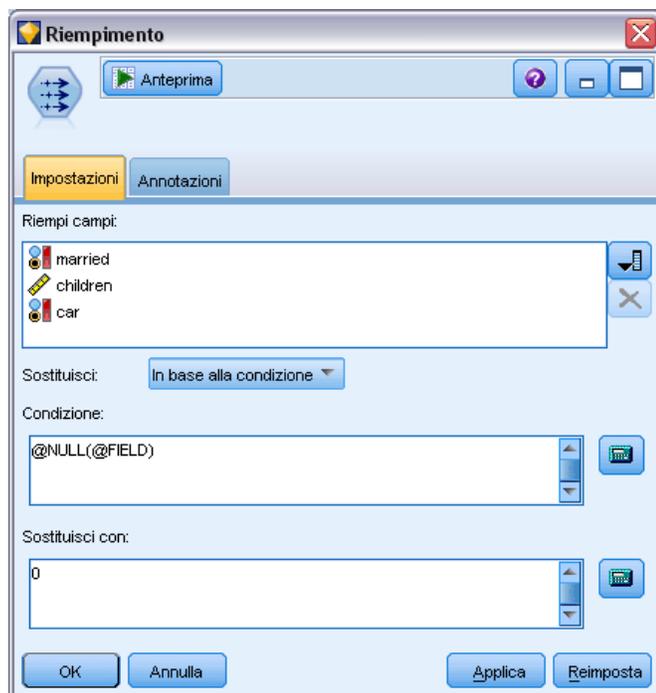
All'interno di un elenco di campi, i valori nulli si possono contare nel modo seguente:

```
count_nulls(['possessocarta' 'possessocarta2' 'possessocarta3'])
```

Quando si utilizzano funzioni che accettano come input un elenco di campi, è possibile utilizzare le funzioni speciali @FIELDS_BETWEEN e @FIELDS_MATCHING come nell'esempio seguente:

```
count_nulls(@FIELDS_MATCHING('carta*'))
```

Figura 6-2
 Utilizzo di un nodo Riempimento per sostituire i valori vuoti con 0 nel campo selezionato



La funzione `undef` consente di riempire i campi con il valore mancante di sistema, visualizzato come `$null$`. Per esempio, mediante un'istruzione condizionale simile alla seguente è possibile sostituire qualsiasi valore numerico:

```
if not(Age > 17) or not(Age < 66) then undef else Age endif
```

Tutti i valori che non rientrano nell'intervallo verranno sostituiti da un valore mancante di sistema, visualizzato con `$null$`. Con la funzione `not()`, è possibile rilevare tutti gli altri valori numerici, inclusi quelli negativi. [Per ulteriori informazioni, vedere l'argomento Funzioni per la gestione di valori vuoti e nulli in il capitolo 8 a pag. 153.](#)

Nota sullo scarto di record

Se si scartano i record utilizzando un nodo Seleziona, tenere presente che la sintassi utilizza la logica tripolare e include automaticamente i valori nulli nelle istruzioni di selezione. Per escludere i valori nulli (valori mancanti di sistema) in un'espressione di selezione è necessario specificarlo in modo esplicito utilizzando `and not` nell'espressione. Per esempio, per selezionare e includere tutti i record nei quali il farmaco prescritto è CuraC, verrà utilizzata l'istruzione di selezione seguente:

```
Drug = 'drugC' and not(@NULL(Drug))
```

Nelle versioni precedenti in questo tipo di situazioni i valori nulli venivano esclusi.

Creazione di espressioni CLEM

Informazioni su CLEM

Control Language for Expression Manipulation (CLEM) è un linguaggio efficace per l'analisi e la manipolazione dei flussi di dati negli stream di IBM® SPSS® Modeler. I data miner fanno ampio ricorso al linguaggio CLEM nelle operazioni stream per eseguire attività sia semplici, come la derivazione dei profitti dai dati relativi ai costi e alle entrate, sia complesse, come la trasformazione dei dati dei registri Web in un insieme di campi e record contenenti informazioni utilizzabili.

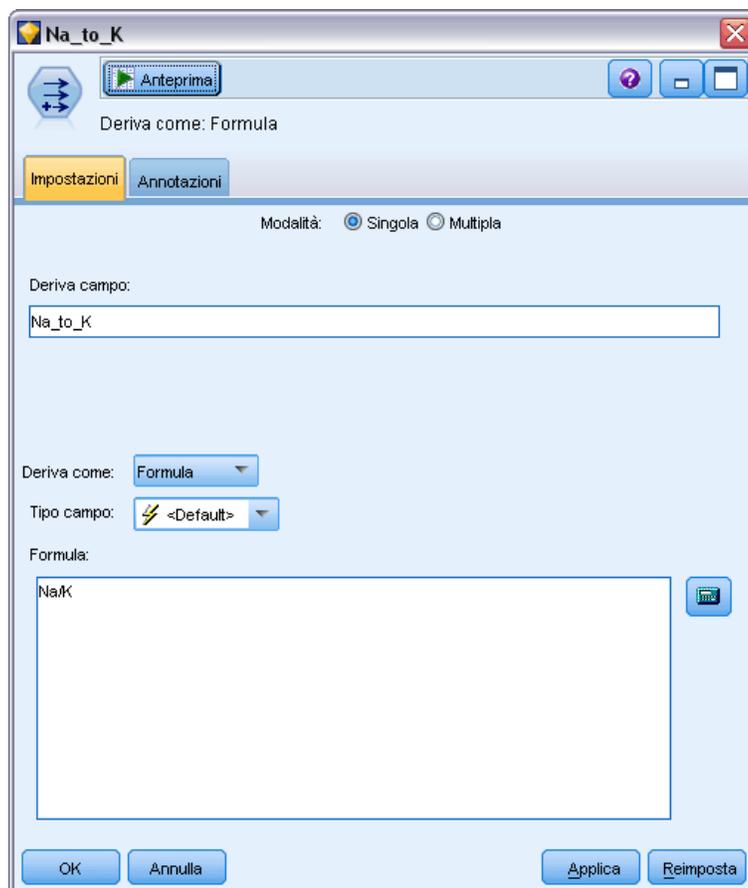
CLEM è utilizzato all'interno di SPSS Modeler per:

- Confrontare e valutare le condizioni nei campi dei record.
- Derivare valori per nuovi campi.
- Derivare nuovi valori per campi esistenti.
- Analizzare le sequenze di record.
- Inserire dati di record in report.

Script. È inoltre possibile utilizzare un sottoinsieme del linguaggio CLEM quando gli script sono eseguiti nell'interfaccia utente, in modo da automatizzare numerose operazioni di manipolazione di dati dello stesso tipo. [Per ulteriori informazioni, vedere l'argomento Panoramica sugli script in il capitolo 2 in IBM SPSS Modeler 14.2 Guida per script e automazione.](#)

Le espressioni CLEM sono indispensabili per la preparazione dei dati in SPSS Modeler e possono essere utilizzate in una vasta gamma di nodi, da operazioni su record e campi (Selezione, Bilanciamento, Riempimento) a plot e output (Analisi, Report, Tabella). Per esempio, è possibile utilizzare il linguaggio CLEM in un nodo Nuovo campo per creare un nuovo campo in base a una formula, quale un rapporto.

Figura 7-1
Nodo Nuovo campo che crea un nuovo campo in base a una formula



È inoltre possibile utilizzare le espressioni CLEM per operazioni di ricerca e sostituzione globali. Per esempio, è possibile utilizzare l'espressione @NULL(@FIELD) in un nodo Riempimento per sostituire **valori mancanti di sistema** con il valore intero 0. Per sostituire **valori mancanti definiti dall'utente**, denominati anche vuoti, si utilizza invece la funzione @BLANK.

Figura 7-2
Nodo Riempimento tramite il quale i valori mancanti di sistema vengono sostituiti con 0



È inoltre possibile creare espressioni CLEM più complesse. Per esempio, è possibile derivare nuovi campi in base a un insieme di regole condizionali.

Figura 7-3
Nodo Nuovo campo condizionale tramite il quale vengono confrontati i valori di un campo con i valori del campo precedente

Nuovo campo

Anteprima

Deriva come: Condizionale

Impostazioni Annotazioni

Modalità: Singola Multipla

Deriva campo:
ValueCategory1

Deriva come: Condizionale

Tipo campo: <Default>

Se:
CardID = @OFFSET(CardID,1)

Allora:
@OFFSET(ValueCategory,1)

Altrimenti:
'exclude'

OK Annulla Applica Reimposta

Esempi CLEM

Di seguito sono riportate alcune espressioni di esempio che illustrano la sintassi corretta nonché i tipi di espressione validi con CLEM.

Espressioni semplici

Le formule possono essere molto semplici, come quella riportata di seguito, che consente di creare un nuovo campo in base ai valori dei campi *Dopo* e *Prima*:

(Dopo - Prima) / Prima * 100.0

Si osservi che i nomi dei campi sono privi di virgolette quando viene fatto riferimento ai valori del campo.

Analogamente, la seguente espressione restituisce il logaritmo di ciascun valore per il campo *salario*.

```
log(salario)
```

Espressioni complesse

Le espressioni possono anche essere lunghe e complesse. La seguente espressione restituisce *vero* se il valore di due campi (*\$KX-Kohonen* e *\$KY-Kohonen*) rientra nell'intervallo specificato. Si noti che qui i nomi dei campi sono racchiusi tra virgolette singole, perché contengono caratteri speciali.

```
('KX-Kohonen' >= -0.2635771036148072 and 'KX-Kohonen' <= 0.3146203637123107  
and 'KY-Kohonen' >= -0.18975617885589602 and  
'KY-Kohonen' <= 0.17674794197082522) -> T
```

Per diverse funzioni, ad esempio le funzioni stringa, è necessario immettere vari parametri utilizzando la sintassi corretta. La funzione *subscr* dell'esempio seguente restituisce il primo carattere di un campo *ID_prodotto*, per indicare se un elemento è biologico, geneticamente modificato o convenzionale. I risultati di un'espressione vengono descritti da *-> `result`*.

```
subscr(1,ID_prodotto) -> `c`
```

Analogamente, la seguente espressione è:

```
stripchar('3','123') -> `12`
```

È importante osservare che i caratteri sono sempre racchiusi tra accenti gravi.

Combinazione di funzioni in un'espressione

Spesso le espressioni CLEM sono costituite da una combinazione di funzioni. Nella funzione seguente, la combinazione di *subscr* e *lowertoupper* restituisce il primo carattere di *ID_prodotto* e lo converte in una lettera maiuscola.

```
lowertoupper(subscr(1,ID_prodotto)) -> `C`
```

Questa stessa espressione può essere scritta in forma ridotta, vale a dire:

```
lowertoupper(ID_prodotto(1)) -> `C`
```

Di seguito è riportata un'altra combinazione comune di funzioni:

```
locchar_back(`n`, (length(pagina_web)), web_page)
```

Questa espressione consente di individuare il carattere *`n`* nei valori del campo *pagina_web* partendo dall'ultimo carattere del valore del campo. Se si include anche la funzione *length*, l'espressione può calcolare dinamicamente la lunghezza del valore corrente, anziché utilizzare un numero statico come 7, che risulterebbe non valido per i valori con meno di sette caratteri.

Funzioni speciali

Sono disponibili numerose funzioni speciali, precedute dal simbolo @. Le funzioni di utilizzo più comune sono:

@BLANK('ID referente') -> T

Spesso le funzioni speciali vengono utilizzate insieme. Si tratta di un metodo di utilizzo comune per evidenziare i valori vuoti in più campi alla volta.

@BLANK(@FIELD)-> T

Ulteriori esempi sono disponibili nella documentazione di CLEM. [Per ulteriori informazioni, vedere l'argomento Panoramica di riferimento su CLEM in il capitolo 8 a pag. 119.](#)

Valori e tipi di dati

Le espressioni CLEM sono simili a formule costruite in base a valori, nomi di campi, operatori e funzioni. La più semplice delle espressioni CLEM valide è un valore o un nome di campo. Esempi di valori validi sono:

3
1.79
'banana'

Esempi di nomi di campi sono:

ID_Prodotto
'\$P-NextField'

dove *Prodotto* è il nome di un campo derivante da un insieme di dati di market basket, '\$P-NextField' è il nome di un parametro e il valore dell'espressione è il valore del campo denominato. I nomi di campi iniziano in genere con una lettera, ma possono contenere anche cifre e caratteri di sottolineatura (_). Se si racchiude il nome tra virgolette, è possibile utilizzare nomi che non seguono queste regole. I valori CLEM possono essere dei seguenti tipi:

- Stringhe: per esempio, "c1", "Type 2", "a piece of free text"
- Interi: per esempio, 12, 0, -189
- Numeri reali: per esempio 12.34, 0.0, -0.0045
- Campi di data/ora: per esempio 05/12/2002, 12/05/2002, 12/05/02

È possibile utilizzare anche i seguenti elementi:

- Codici carattere: per esempio, `a` or 3
- Elenchi di elementi: per esempio, [1 2 3], ['Type 1' 'Type 2']

In genere i codici carattere e gli elenchi non vengono utilizzati come valori di campo, ma come argomenti di funzioni CLEM.

Regole per l'inserimento tra virgolette

Sebbene il software sia flessibile nella determinazione dei campi, dei valori, dei parametri e delle stringhe utilizzati in un'espressione CLEM, di seguito sono suggerite alcune "procedure consigliate" da seguire durante la creazione di espressioni.

- **Stringhe:** utilizzare sempre le virgolette doppie quando si scrivono stringhe ("Type 2" o "value"). È inoltre possibile utilizzare le virgolette singole, ma il rischio è di creare confusione con i campi racchiusi tra virgolette.
- **Caratteri:** utilizzare l'accento grave, ovvero ` . Si osservi, per esempio, il carattere d nella funzione `stripchar(`d`,`drugA`)`. L'unica eccezione è rappresentata dall'utilizzo di un intero per fare riferimento a un carattere specifico in una stringa. Si osservi, per esempio, il carattere 5 nella funzione `lowertoupper("drugA"(5))` → "A". *Nota:* il codice Unicode dell'accento grave è 0060.
- **Campi:** in genere, nelle espressioni CLEM i campi non sono racchiusi tra virgolette (`subscr(2,arrayID)` → CHAR). Utilizzare le virgolette singole solo se necessarie per racchiudere spazi o altri caratteri speciali ('Order Number'). I campi racchiusi tra virgolette ma non definiti nell'insieme di dati verranno interpretati come stringhe.
- **Parametri:** utilizzare sempre le virgolette singole ('\$P-threshold').

Espressioni e condizioni

Le espressioni CLEM possono restituire un risultato, utilizzato quando si derivano nuovi valori, per esempio:

```
Weight * 2.2
Age + 1
sqrt(Signal-Echo)
```

oppure possono restituire *vero* o *falso*, utilizzato quando la selezione viene effettuata in base a una condizione, per esempio:

```
Drug = "drugA"
Age < 16
not(PowerFlux) and Power > 2000
```

Nelle espressioni CLEM è possibile combinare in modo arbitrario operatori e funzioni, per esempio:

```
sqrt(abs(Signal)) * max(T1, T2) + Baseline
```

Le parentesi e la precedenza tra gli operatori determinano in quale ordine le espressioni verranno valutate. In questo esempio l'ordine di valutazione è il seguente:

- Viene valutato `abs(Signal)` e al suo risultato si applica `sqrt`.
- Viene valutato `max(T1, T2)`.
- I due risultati vengono moltiplicati: `x` ha la precedenza su `+`.
- Infine al risultato viene aggiunto `Baseline`.

Di seguito è specificato l'ordine decrescente di precedenza, ovvero dalle operazioni eseguite per prime a quelle eseguite per ultime:

- Argomenti della funzione
- Chiamate della funzione
- xx
- x / mod div rem
- + -
- > < >= <= /= == === !=

Se si desidera ignorare la precedenza o se si hanno dei dubbi sull'ordine di valutazione, è possibile utilizzare le parentesi in modo esplicito, per esempio

```
sqrt(abs(Signal)) * (max(T1, T2) + Baseline)
```

Parametri stream, sessione e Supernodo

È possibile definire parametri da utilizzare in espressioni CLEM e negli script. Si tratta a tutti gli effetti di variabili definite dall'utente che vengono salvate e mantenute con lo stream, la sessione o il Supernodo corrente e a cui è possibile accedere dall'interfaccia utente nonché attraverso script. Per esempio, durante il salvataggio di uno stream, qualsiasi parametro impostato per tale stream viene salvato contemporaneamente. In questo modo i parametri vengono distinti dalle variabili di script locali, utilizzabili solo nello script in cui sono dichiarate. In ambiente di scripting i parametri fanno spesso parte di un'espressione CLEM e il relativo valore è specificato nello script stesso.

L'ambito di un parametro dipende dalla posizione in cui è impostato:

- i parametri impostati per uno stream possono essere impostati in uno script dello stream o nella finestra di dialogo delle proprietà dello stream e sono disponibili a tutti i nodi dello stream. Vengono visualizzati nell'elenco Parametri nel generatore di espressioni.
- I parametri impostati per una sessione possono essere impostati in uno script autonomo oppure nella finestra di dialogo dei parametri della sessione e sono disponibili per tutti gli stream utilizzati nella sessione corrente, ovvero per tutti gli stream elencati nella scheda Stream del riquadro dei manager.

È inoltre possibile impostare parametri per Supernodi. In questo caso, sono visibili solo ai nodi incapsulati all'interno del Supernodo. [Per ulteriori informazioni, vedere l'argomento Definizione dei parametri di un Supernodo in il capitolo 9 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)

Utilizzo dei parametri nelle espressioni CLEM

I parametri vengono rappresentati nelle espressioni CLEM con \$P-pname dove pname rappresenta il nome del parametro. Quando vengono utilizzati nelle espressioni CLEM, i parametri devono essere racchiusi tra virgolette singole, per esempio '\$P-scale'.

I parametri disponibili possono essere facilmente visualizzati mediante il Generatore di espressioni. Per visualizzare i parametri correnti:

- ▶ In qualsiasi finestra di dialogo che accetta espressioni CLEM, fare clic sul pulsante Generatore espressioni.
- ▶ Selezionare Parametri dall'elenco Campi.

È quindi possibile selezionare dall'elenco i parametri da inserire nell'espressione CLEM. [Per ulteriori informazioni, vedere l'argomento Selezione di campi, parametri e variabili globali a pag. 113.](#)

Utilizzo di stringhe

Esistono svariate operazioni disponibili per le stringhe, tra cui:

- Conversione di una stringa in caratteri maiuscoli o minuscoli: `uppertolower(CHAR)`.
- Rimozione di caratteri specifici, per esempio ``ID_`` o ``$``, da una variabile stringa: `stripchar(CHAR,STRING)`.
- Determinazione della lunghezza (numero di caratteri) per una variabile stringa: `length(STRING)`.
- Controllo dell'ordine alfabetico dei valori di stringa: `alphabefore(STRING1, STRING2)`.
- Rimozione dello spazio bianco iniziale o finale dai valori: `trim(STRING)`, `trim_start(STRING)` o `trimend(STRING)`.
- Estrazione dei primi o ultimi *n* caratteri da una stringa: `startstring(LENGTH, STRING)` o `endstring(LENGTH, STRING)`. Si supponga, per esempio, di disporre di un campo denominato *articolo* che combina il nome di un prodotto con un codice ID di 4 cifre (ACME CAMERA-D109). Per creare un nuovo campo contenente solo il codice di 4 cifre, specificare la formula seguente in un nodo Nuovo campo:

```
endstring(4, articolo)
```

- Corrispondenza di uno schema specifico: `STRING matches PATTERN`. Per esempio, per selezionare le persone il cui titolo professionale include "market", è possibile specificare quanto segue in un nodo Seleziona:

```
posizione matches "*"market*"
```

- Sostituzione di tutte le istanze di una sottostringa all'interno di una stringa: `replace(SUBSTRING, NEWSUBSTRING, STRING)`. Per esempio, per sostituire tutte le istanze di un carattere non supportato, come una barra verticale (|), con un punto e virgola prima del text mining, utilizzare la funzione `replace` in un nodo Riempimento. In Riempi campi,

selezionare tutti i campi in cui il carattere può essere presente. Per la condizione Sostituisci: selezionare Sempre e specificare la condizione seguente per Sostituisci con:

```
replace(';',',',@FIELD)
```

- Derivazione di un campo flag in base alla presenza di una specifica sottostringa. Per esempio, sarebbe possibile utilizzare una funzione di stringa in un nodo Nuovo campo per generare un campo flag separato per ciascuna risposta con un'espressione quale:

```
hassubstring(museums,"museum_of_design")
```

Per ulteriori informazioni, vedere l'argomento Funzioni stringa in il capitolo 8 a pag. 136.

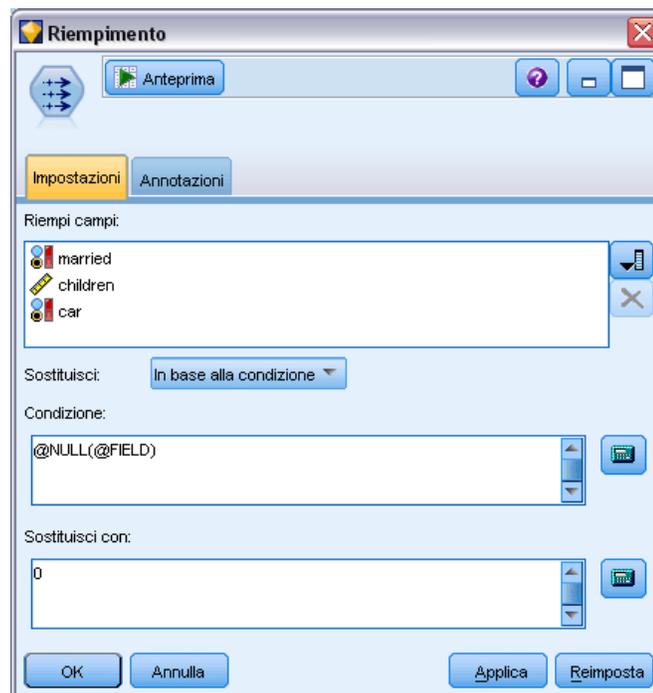
Gestione di vuoti e valori mancanti

La sostituzione dei vuoti o dei valori mancanti è una comune attività di preparazione dei dati per i data miner. In CLEM sono disponibili diversi strumenti per automatizzare la gestione dei vuoti. Il nodo Riempimento è l'ambito più comune per la gestione dei vuoti. È tuttavia possibile utilizzare le seguenti funzioni in qualsiasi nodo che accetti le espressioni CLEM:

- @BLANK(FIELD) può essere utilizzata per determinare i record i cui valori sono vuoti per un determinato campo, quale *Età*.
- @NULL(FIELD) può essere utilizzata per determinare i record i cui valori sono valori mancanti di sistema per i campi specificati. In IBM® SPSS® Modeler, i valori mancanti di sistema sono visualizzati come valori \$null\$.

Figura 7-4

Nodo Riempimento tramite il quale i valori mancanti di sistema vengono sostituiti con 0



Per ulteriori informazioni, vedere l'argomento [Funzioni per la gestione di valori vuoti e nulli](#) in il capitolo 8 a pag. 153.

Utilizzo di numeri

In IBM® SPSS® Modeler sono disponibili numerose operazioni standard sui valori numerici, quali:

- Calcolo del seno dell'angolo specificato: $\sin(\text{NUM})$
- Calcolo del logaritmo naturale dei campi numerici: $\log(\text{NUM})$.
- Calcolo della somma di due numeri: $\text{NUM1} + \text{NUM2}$

Per ulteriori informazioni, vedere l'argomento [Funzioni numeriche](#) in il capitolo 8 a pag. 132.

Utilizzo di date e ore

I formati di data e ora possono variare a seconda della sorgente dati e delle impostazioni internazionali. I formati di data e ora sono specifici di ogni stream e sono impostati nella finestra di dialogo delle proprietà dello stream. I seguenti sono esempi di funzioni comunemente utilizzate per i campi di data/ora.

Calcolo del tempo trascorso

È possibile calcolare facilmente il tempo trascorso da una data di riferimento tramite una famiglia di funzioni simili a quella seguente, che restituisce il tempo, in mesi, dalla data di riferimento alla data rappresentata dalla stringa della data `DATE`, come numero reale. Tale numero corrisponde a una cifra approssimativa basata su un mese formato da 30 giorni.

```
date_in_months(Date)
```

Confronto di valori Data/Ora

È possibile confrontare i valori dei campi di data/ora nei record tramite funzioni simili a quella seguente, che restituisce il valore *vero*, se la stringa della data `DATE1` rappresenta una data antecedente a quella rappresentata dalla stringa della data `DATE2`. In caso contrario, questa funzione restituirà un valore 0.

```
date_before(Date1, Date2)
```

Calcolo di differenze

È anche possibile calcolare la differenza tra due ore e due date tramite funzioni quali:

```
date_weeks_difference(Date1, Date2)
```

Questa funzione restituisce il tempo, in settimane, dalla data rappresentata dalla stringa della data DATE1 alla data rappresentata dalla stringa della data DATE2, come numero reale. Tale numero è basato su una settimana formata da sette giorni. Se DATE2 è antecedente a DATE1, questa funzione restituisce un numero negativo.

Data odierna

È possibile aggiungere la data corrente all'insieme di dati tramite la funzione @TODAY. La data odierna viene aggiunta come stringa al campo specificato o al nuovo campo con il formato di data selezionato nella finestra di dialogo delle proprietà dello stream. [Per ulteriori informazioni, vedere l'argomento Funzioni di data e ora in il capitolo 8 a pag. 142.](#)

Riepilogo di più campi

Il linguaggio CLEM include una serie di funzioni che restituiscono statistiche riepilogative di più campi. Tali funzioni possono essere particolarmente utili per l'analisi di dati di sondaggi in cui è possibile che risposte multiple a una singola domanda siano state archiviate in più campi. [Per ulteriori informazioni, vedere l'argomento Utilizzo di dati a risposta multipla a pag. 109.](#)

Funzioni di confronto

È possibile confrontare valori di più campi utilizzando le funzioni min_n e max_n, per esempio:

```
max_n(['importocarta1' 'importocarta2' 'importocarta3' 'importocarta4'])
```

È inoltre possibile utilizzare una serie di funzioni di conteggio per ottenere conteggi di valori in base a determinati criteri, anche quando tali valori sono archiviati in più campi. Per esempio, per contare il numero di carte che sono state possedute per oltre cinque anni:

```
count_greater_than(5, ['possessocarta' 'possessocarta2' 'possessocarta3'])
```

Per contare valori nulli nello stesso insieme di campi:

```
count_nulls(['possessocarta' 'possessocarta2' 'possessocarta3'])
```

Si noti che nell'esempio viene conteggiato il numero di carte possedute e non il numero di possessori. [Per ulteriori informazioni, vedere l'argomento Funzioni di confronto in il capitolo 8 a pag. 129.](#)

Funzioni numeriche

È possibile ottenere i dati statistici di più campi utilizzando le funzioni sum_n, mean_n e sdev_n, per esempio:

```
sum_n(['saldocarta1' 'saldocarta2' 'saldocarta3'])
```

```
mean_n(['saldocarta1' 'saldocarta2' 'saldocarta3'])
```

[Per ulteriori informazioni, vedere l'argomento Funzioni numeriche in il capitolo 8 a pag. 132.](#)

Creazione di elenchi di campi

Quando si utilizza una qualsiasi funzione che accetta un elenco di campi come input, è possibile utilizzare come input le funzioni speciali @FIELDS_BETWEEN(start, end) e @FIELDS_MATCHING(pattern). Per esempio, supponendo che l'ordine dei campi sia quello mostrato nell'esempio sum_n precedente, la seguente espressione sarebbe equivalente:

```
sum_n(@FIELDS_BETWEEN(saldocarta1, saldocarta3))
```

In alternativa, per contare il numero di valori nulli in tutti i campi che iniziano con "carta":

```
count_nulls(@FIELDS_MATCHING('carta*'))
```

Per ulteriori informazioni, vedere l'argomento Campi speciali in il capitolo 8 a pag. 154.

Utilizzo di dati a risposta multipla

Per analizzare dati a risposta multipla è possibile utilizzare una serie di funzioni di confronto, quali:

- value_at
- first_index / last_index
- first_non_null / last_non_null
- first_non_null_index / last_non_null_index
- min_index / max_index

Si supponga, per esempio, una domanda a risposta multipla sulla prima, seconda e terza più importante ragione per decidere su un determinato acquisto (come prezzo, consiglio personale, commenti su pubblicazioni specializzate, fornitore locale, altro). In tal caso è possibile determinare l'importanza del prezzo derivando l'indice del campo in cui è stato incluso per primo:

```
first_index("prezzo", [Motivo1 Motivo2 Motivo3])
```

Analogamente, si supponga di avere chiesto ai clienti di classificare tre auto in ordine di probabilità di acquisto e di avere codificato le risposte in tre campi separati, come segue:

ID cliente	auto1	auto2	auto3
101	1	3	2
102	3	2	1
103	2	3	1

In tal caso è possibile determinare l'indice del campo dell'auto preferita (classificata al numero 1 o con il rango inferiore) utilizzando la funzione min_index:

```
min_index(['auto1' 'auto2' 'auto3'])
```

Per ulteriori informazioni, vedere l'argomento Funzioni di confronto in il capitolo 8 a pag. 129.

Riferimento agli insiemi di risposte multiple

La funzione speciale @MULTI_RESPONSE_SET può essere utilizzata per fare riferimento a tutti i campi in un insieme di risposte multiple. Per esempio, se i tre campi *auto* dell'esempio precedente venissero inclusi in un insieme di risposte multiple denominato *classifica_auto*, l'espressione seguente restituirebbe lo stesso risultato:

```
max_index(@MULTI_RESPONSE_SET("classifica_auto"))
```

Per ulteriori informazioni, vedere l'argomento [Modifica di insiemi di risposte multiple](#) in [capitolo 4 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output](#).

Generatore di espressioni

È possibile immettere le espressioni CLEM manualmente oppure utilizzare il Generatore di espressioni, che visualizza un elenco completo di funzioni e operatori CLEM, nonché i campi di dati dello stream corrente, consentendo di creare espressioni in modo rapido senza la necessità di memorizzare il nome esatto di campi o funzioni. Inoltre il Generatore esegue controlli automatici e aggiunge le virgolette appropriate a campi e valori, facilitando la creazione di espressioni sintatticamente corrette.

Figura 7-5
Finestra di dialogo Generatore espressioni

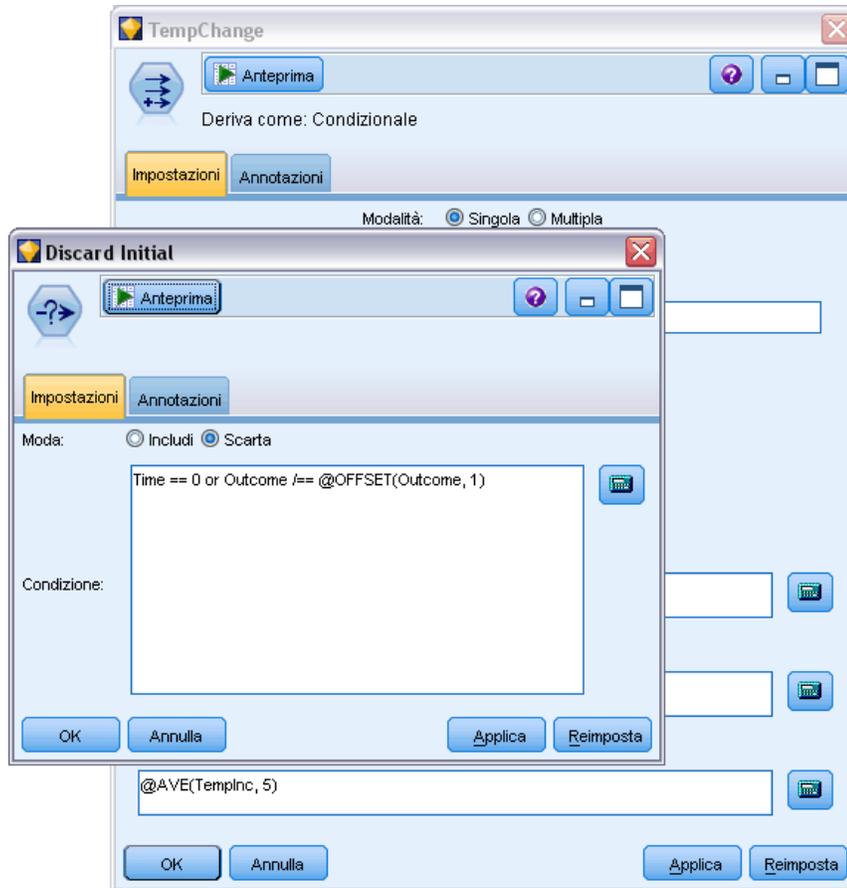


Nota: il Generatore di espressioni non è supportato nelle impostazioni di script o di parametri.

Accesso al Generatore di espressioni

Il Generatore di espressioni è disponibile in tutti i nodi in cui vengono utilizzate espressioni CLEM, inclusi i nodi Seleziona, Bilanciamento, Nuovo campo, Riempimento, Plot, Plot multiplo, Analisi, Report e Tabella. Per aprire il Generatore di espressioni, fare clic sul pulsante del calcolatore immediatamente a destra del campo della formula.

Figura 7-6
Nodi visualizzati con il pulsante del Generatore di espressioni



Creazione di espressioni

Il Generatore di espressioni non solo fornisce elenchi completi di campi, funzioni e operatori, ma consente anche di accedere ai valori dei dati, se i dati sono istanziati.

Per creare un'espressione tramite il Generatore di espressioni

- ▶ Digitare nel campo dell'espressione, utilizzando gli elenchi di funzioni e campi come riferimenti.
or
- ▶ Selezionare i campi e le funzioni richiesti dagli elenchi scorrevoli.
- ▶ Fare doppio clic o fare clic sul pulsante con la freccia gialla per aggiungere il campo o la funzione al campo dell'espressione.
- ▶ Utilizzare i pulsanti degli operandi al centro della finestra di dialogo per inserire le operazioni nell'espressione.

Selezione di funzioni

Nell'elenco di funzioni sono visualizzate tutte le funzioni e gli operatori CLEM disponibili. Scorrere l'elenco per selezionare una funzione o, per effettuare la ricerca in modo più semplice, utilizzare l'elenco a discesa per visualizzare un sottoinsieme di funzioni o di operatori. Le funzioni disponibili sono state raggruppate in categorie, per semplificare la ricerca.

Figura 7-7
Elenco a discesa delle funzioni



Due categorie possono risultare particolarmente utili:

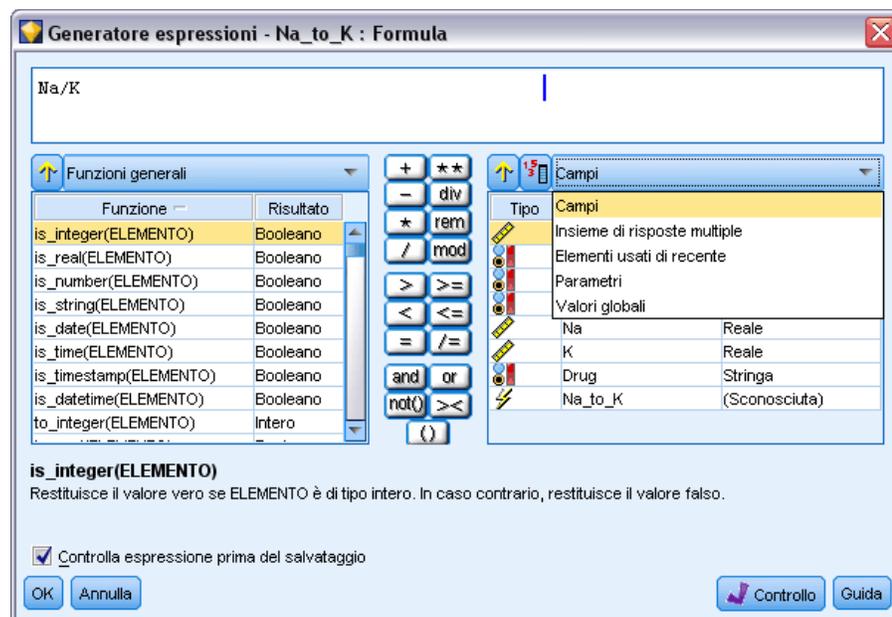
- **Tutte le funzioni** contiene un elenco completo delle funzioni CLEM disponibili.
- **Elementi usati di recente** contiene un elenco delle funzioni CLEM utilizzate nella sessione corrente.

Dopo avere selezionato un gruppo di funzioni, fare doppio clic per inserire le funzioni nel campo dell'espressione nel punto indicato dalla posizione del cursore.

Selezione di campi, parametri e variabili globali

Nell'elenco dei campi sono visualizzati tutti i campi disponibili in questo punto dello stream di dati. Scorrere l'elenco per selezionare un campo. Fare doppio clic o utilizzare il pulsante con la freccia gialla per aggiungere un campo all'espressione.

Figura 7-8
Generatore di espressioni: Elenco di campi



Per ulteriori informazioni, vedere l'argomento [Parametri stream, sessione e Supernodo](#) a pag. 104.

Oltre ai campi, è possibile scegliere tra i seguenti elementi:

Insiemi di risposte multiple. Per maggiori informazioni, consultare il manuale *Nodi di input, di elaborazione e di output in IBM SPSS Modeler*.

Elementi usati di recente contiene un elenco di campi, insieme a risposta multipla, parametri e valori globali utilizzati all'interno della sessione corrente.

Parametri. Per ulteriori informazioni, vedere l'argomento [Parametri stream, sessione e Supernodo](#) a pag. 104.

Valori globali. Per maggiori informazioni, consultare il manuale *Nodi di input, di elaborazione e di output in IBM SPSS Modeler*.

Visualizzazione o selezione di valori

I valori dei campi possono essere visualizzati da svariati punti del sistema, inclusi il Generatore di espressioni e i report del nodo Esplora, e nel corso della modifica dei valori futuri in un nodo Intervalli di tempo. Si noti che per utilizzare questa funzionalità è necessario che i dati vengano completamente istanziati in un nodo Tipo o di input, in modo che l'archiviazione, i tipi e i valori

siano noti. Per ulteriori informazioni, vedere l'argomento Utilizzo della finestra di dialogo Valori in il capitolo 4 in *IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output*.

Figura 7-9

Elenco Campi con i valori visualizzati per il campo selezionato



- Per visualizzare i valori dal Generatore di espressioni o da un nodo Intervalli di tempo, selezionare il campo richiesto e fare clic sul pulsante di selezione dei valori per aprire una finestra di dialogo contenente i valori del campo selezionato. È quindi possibile selezionare un valore e fare clic su Inserisci per incollarlo nell'espressione o elenco corrente.

Figura 7-10

Pulsante di selezione dei valori

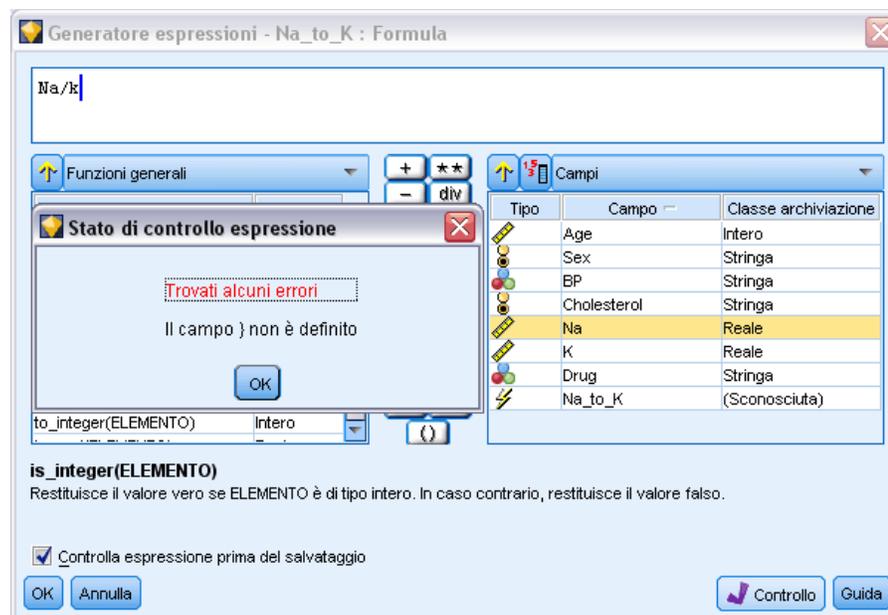


Per i campi flag e nominali, vengono elencati tutti i valori definiti. Per i campi continui (intervallo numerico), vengono visualizzati i valori minimi e massimi.

Controllo di espressioni CLEM

Per convalidare l'espressione, fare clic su Controlla nel Generatore di espressioni (angolo inferiore destro). Le espressioni non controllate sono visualizzate in rosso. Se vengono rilevati degli errori, viene visualizzato un messaggio che ne comunica la causa.

Figura 7-11
Espressione CLEM non valida



Vengono controllati i seguenti elementi:

- Inserimento corretto tra virgolette di valori e nomi di campi
- Utilizzo corretto di parametri e variabili globali
- Utilizzo valido di operatori
- Esistenza di campi di riferimento
- Esistenza e definizione di valori globali di riferimento

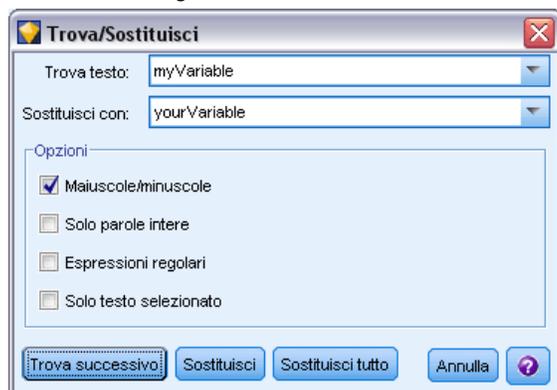
Se vengono rilevati errori di sintassi, provare a ricreare l'espressione utilizzando gli elenchi e i pulsanti degli operatori anziché la digitazione manuale. In questo modo le virgolette appropriate per campi e valori verranno aggiunte automaticamente.

Trova e sostituisci

La finestra di dialogo Trova/Sostituisci è disponibile ogni volta che è possibile modificare il testo di script o di espressioni, compreso l'editor di script, il generatore di espressioni CLEM e quando si definisce un modello nel nodo Report. Quando si modifica un testo in una di queste aree, premere Ctrl-F per accedere alla finestra di dialogo, assicurandosi che il cursore sia posizionato in un'area di testo. In un nodo Riempimento, per esempio, è possibile accedere alla finestra

di dialogo da qualsiasi area di testo della scheda Impostazioni oppure dal campo di testo nel Generatore di espressioni.

Figura 7-12
Finestra di dialogo Trova/Sostituisci



- ▶ Con il cursore posizionato in un'area di testo, premere Ctrl+F per accedere alla finestra di dialogo Trova/Sostituisci.
- ▶ Immettere il testo da cercare oppure sceglierne uno dall'elenco a discesa degli elementi cercati di recente.
- ▶ Se necessario, immettere il testo sostitutivo.
- ▶ Fare clic su Trova successivo per avviare la ricerca.
- ▶ Fare clic su Sostituisci per sostituire la selezione corrente oppure scegliere Sostituisci tutto per aggiornare tutte le istanze o quelle selezionate.
- ▶ Al termine di ogni operazione, la finestra di dialogo si chiude. Premere F3 da qualsiasi area di testo per ripetere l'ultima operazione di ricerca oppure premere Ctrl+F per accedere nuovamente alla finestra di dialogo.

Opzioni di ricerca

Mauscole/minuscole. Specifica se l'operazione di ricerca fa distinzione tra maiuscole e minuscole, per esempio se *miavar* corrisponde a *miaVar*. Il testo sostitutivo viene sempre inserito esattamente come viene digitato, indipendentemente da questa impostazione.

Solo parole intere. Specifica se l'operazione di ricerca cerca le occorrenze che sono parole intere. Se questa opzione è selezionata, la ricerca di *palla* non consentirà di trovare per esempio *pallavolo* o *Palladio*.

Espressioni regolari. Specifica se è utilizzata la sintassi delle espressioni regolari (vedere la sezione seguente). Quando questa opzione è selezionata, l'opzione Solo parole intere è disattivata e il relativo valore viene ignorato.

Solo testo selezionato. Controlla l'ambito della ricerca quando si utilizza l'opzione Sostituisci tutto.

Sintassi delle espressioni regolari

Le espressioni regolari consentono di cercare caratteri speciali, quali tabulazioni o caratteri di nuova riga, classi o intervalli di caratteri quali *a - d*, cifre e caratteri diversi da cifre, nonché limiti, per esempio l'inizio o la fine di una riga. Sono supportati i seguenti tipi di espressioni.

Corrispondenze di caratteri

Caratteri	Corrispondenze
x	Il carattere x
\\	Il carattere barra rovesciata
\On	Il carattere con valore ottale 0n (0 <= n <= 7)
\Onn	Il carattere con valore ottale 0nn (0 <= n <= 7)
\Omnn	Il carattere con valore ottale 0mnn (0 <= m <= 3, 0 <= n <= 7)
\xhh	Il carattere con valore esadecimale 0xhh
\uhhhh	Il carattere con valore esadecimale 0xhhhh
\t	Il carattere di tabulazione (' <u>0009'</u>)
\n	Il carattere di nuova riga (avanzamento riga) (' <u>000A'</u>)
\r	Il carattere di ritorno a capo (' <u>000D'</u>)
\f	Il carattere di avanzamento carta (' <u>000C'</u>)
\a	Il carattere di avviso (campanello) (' <u>0007'</u>)
\e	Il carattere di escape (' <u>001B'</u>)
\cx	Il carattere di controllo corrispondente a x

Corrispondenze di classi di caratteri

Classi di caratteri	Corrispondenze
[abc]	a, b o c (classe semplice)
[^abc]	Qualsiasi carattere, eccetto a, b o c (sottrazione)
[a-zA-Z]	a-z oppure A-Z, incluse (intervallo)
[a-d[m-p]]	a-d oppure m-p (unione). In alternativa è possibile specificare [a-dm-p].
[a-z&&[def]]	a-z + d, e oppure f (intersezione)
[a-z&&[^bc]]	a-z, eccetto b e c (sottrazione). In alternativa è possibile specificare [a-dz].
[a-z&&[^m-p]]	a-z, eccetto m-p (sottrazione). In alternativa è possibile specificare [a-lq-z].

Classi di caratteri predefinite

Classi di caratteri predefinite	Corrispondenze
.	Qualsiasi carattere (può corrispondere o meno a terminazioni di riga)
\d	Qualsiasi cifra: [0-9]
\D	Una non cifra: [^0-9]
\s	Uno spazio vuoto: [\t\n\r\b\f\r]
\S	Uno spazio non vuoto: [^\s]
\w	Un carattere alfanumerico: [a-zA-Z_0-9]
\W	Un carattere diverso da alfanumerico: [^\w]

Corrispondenze di limiti

Corrispondenze di limiti	Corrispondenze
<code>^</code>	L'inizio di una riga
<code>\$</code>	La fine di una riga
<code>\b</code>	Un limite di parola
<code>\B</code>	Un limite diverso da un limite di parola
<code>\A</code>	L'inizio dell'input
<code>\Z</code>	La fine dell'input ma per la terminazione finale, se disponibile
<code>\z</code>	La fine dell'input

Riferimenti al linguaggio CLEM

Panoramica di riferimento su CLEM

Questa sezione illustra il linguaggio CLEM (Control Language for Expression Manipulation), un potente strumento per l'analisi e la manipolazione dei dati utilizzati negli stream di IBM® SPSS® Modeler. È possibile utilizzare CLEM all'interno di nodi per eseguire svariate attività, dalla valutazione di condizioni o derivazione di valori fino all'inserimento di dati in report. [Per ulteriori informazioni, vedere l'argomento Informazioni su CLEM in il capitolo 7 a pag. 97.](#)

È inoltre possibile utilizzare un sottoinsieme del linguaggio CLEM quando gli script sono eseguiti nell'interfaccia utente, in modo da automatizzare numerose operazioni di manipolazione di dati dello stesso tipo. [Per ulteriori informazioni, vedere l'argomento Espressioni CLEM negli script in il capitolo 3 in IBM SPSS Modeler 14.2 Guida per script e automazione.](#)

Le espressioni CLEM sono costituite da valori, nomi di campi, operatori e funzioni. Tramite la sintassi corretta è possibile creare una vasta gamma di potenti operazioni sui dati. [Per ulteriori informazioni, vedere l'argomento Esempi CLEM in il capitolo 7 a pag. 100.](#)

Tipi di dati CLEM

Nel linguaggio CLEM è possibile utilizzare i seguenti tipi di dati:

- Interi
- Numeri reali
- Caratteri
- Stringhe
- Elenchi
- Campi
- Data/ora

Regole per l'inserimento tra virgolette

Sebbene IBM® SPSS® Modeler sia flessibile quando l'utente determina campi, valori, parametri e stringhe utilizzati in un'espressione CLEM, di seguito sono suggerite alcune "procedure consigliate" da seguire durante la creazione di espressioni.

- Stringhe: utilizzare sempre virgolette doppie per la scrittura di stringhe, per esempio "Type 2". È inoltre possibile utilizzare le virgolette singole, ma il rischio è di creare confusione con i campi racchiusi tra virgolette.
- Campi: utilizzare virgolette singole solo se necessarie per racchiudere spazi o altri caratteri speciali, per esempio 'Order Number'. I campi racchiusi tra virgolette ma non definiti nell'insieme di dati verranno interpretati come stringhe.

- Parametri: utilizzare sempre virgolette singole quando si specificano parametri, per esempio '\$P-threshold'.
- Caratteri: utilizzare sempre virgolette singole a barra rovesciata (```), per esempio `stripchar(`d`, "drugA")`.

Per ulteriori informazioni, vedere l'argomento [Valori e tipi di dati in il capitolo 7 a pag. 102](#). Queste regole verranno illustrate in modo più dettagliato negli argomenti riportati di seguito.

Interi

I numeri interi vengono rappresentati come sequenze di cifre decimali. Se lo si desidera, è possibile inserire un segno meno (-) prima del numero intero per indicare un numero negativo, per esempio 1234, 999, -77.

Il linguaggio CLEM è in grado di gestire numeri interi a precisione arbitraria. La dimensione massima dei numeri interi dipende dalla piattaforma in uso. Se i valori hanno dimensioni superiori a quelle visualizzabili in un campo di numeri interi, l'impostazione del tipo di campo su Real consente in genere di ripristinare il valore.

Numeri reali

Con *numero reale* si intende un numero a virgola mobile. I numeri reali sono rappresentati da una o più cifre seguite da un punto decimale, seguito a sua volta da una o più cifre. Nel linguaggio CLEM, vengono utilizzati numeri reali con precisione doppia.

Se lo si desidera, è possibile inserire un segno meno (-) prima del numero reale per indicare un numero negativo, per esempio 1.234, 0.999, -77.001. Per esprimere un numero reale con la notazione esponenziale, utilizzare il formato `<numero>` e `<esponente>`, per esempio `1234.0e5`, `1.7e-` e `2`. Quando le stringhe di numeri vengono lette e convertite automaticamente in numeri dall'applicazione IBM® SPSS® Modeler, vengono comunque accettati i numeri senza cifre iniziali prima del punto decimale e senza cifre dopo il punto, per esempio `999.` o `.11`. Questi formati non sono tuttavia validi nelle espressioni CLEM.

Nota: quando si fa riferimento a numeri reali in espressioni CLEM, è necessario utilizzare il punto come separatore decimale indipendentemente da eventuali impostazioni dello stream corrente o dalle impostazioni internazionali. Per esempio, specificare

`Na > 0.6`

anziché

`Na > 0.6`

Questa regola è valida anche se nella finestra di dialogo delle proprietà dello stream è stata selezionata la virgola come simbolo decimale ed è conforme alla linea guida generale secondo cui la sintassi del codice deve essere indipendente da specifiche impostazioni internazionali o convenzioni.

Caratteri

I caratteri sono in genere visualizzati come CHAR e nelle espressioni CLEM consentono di eseguire test sulle stringhe. Per esempio, la funzione `isuppercode` consente di stabilire se il primo carattere di una stringa è maiuscolo. L'espressione CLEM seguente utilizza un carattere per indicare che il test deve essere eseguito sul primo carattere della stringa:

```
isuppercode(subscrs(1,"MiaStringa"))
```

Per esprimere il codice (anziché la posizione) di un carattere specifico in un'espressione CLEM, utilizzare virgolette singole a barra rovesciata nel formato `<carattere>`, per esempio, ``A``, ``Z``.

Nota: Non esiste un tipo di archiviazione CHAR per un campo, pertanto se un campo deriva o viene riempito con un'espressione che produce un tipo CHAR, il risultato verrà convertito in una stringa.

Stringhe

È in genere opportuno racchiudere le stringhe tra virgolette doppie. Per esempio le stringhe "c35product2" e "referrerID". Per indicare caratteri speciali in una stringa, utilizzare una barra rovesciata, per esempio "\$65443". (Per indicare una barra rovesciata, utilizzare una doppia barra rovesciata, `\\`). È possibile racchiudere una stringa tra virgolette singole, ma in questo caso il risultato (per esempio 'referrerID') non è distinguibile da un campo tra virgolette. [Per ulteriori informazioni, vedere l'argomento Funzioni stringa a pag. 136.](#)

Elenchi

Un elenco rappresenta una sequenza ordinata di elementi, che possono essere di tipo eterogeneo. Gli elenchi sono racchiusi tra parentesi quadre (`[]`). Per esempio gli elenchi `[1 2 4 16]` e `["abc" "def"]`. Gli elenchi non vengono utilizzati come valori dei campi di IBM® SPSS® Modeler, bensì per fornire argomenti alle funzioni, quali `member` e `oneof`.

Campi

I nomi inclusi nelle espressioni CLEM non corrispondenti a nomi di funzioni vengono interpretati come nomi di campi. È possibile scriverli come `Power`, `val27`, `state_flag` e così via, ma se il nome inizia con una cifra o include caratteri non alfabetici, quali gli spazi (ad eccezione del carattere di sottolineatura), racchiudere il nome tra virgolette singole, per esempio, `'Power Increase'`, `'2nd answer'`, `'#101'`, `'$P-NextField'`.

Nota: I campi racchiusi tra virgolette ma non definiti nell'insieme di dati verranno interpretati come stringhe.

Date

I calcoli di data sono basati su una data di riferimento, specificata nella finestra di dialogo delle proprietà dello stream. La data di riferimento di default è 1 gennaio 1900. [Per ulteriori informazioni, vedere l'argomento Impostazione delle opzioni per gli stream in il capitolo 5 a pag. 51.](#)

Il linguaggio CLEM supporta i formati di data seguenti.

Formato	Esempi
DDMMYY	150163
MMDDYY	011563
YYMMDD	630115
YYYYMMDD	19630115
YYYYDDD	Anno a quattro cifre seguito da numero a tre cifre rappresentante il giorno dell'anno, per esempio 2000032 rappresenta il 32esimo giorno del 2000 oppure l'1 febbraio 2000.
DAY	Giorno della settimana nelle opzioni internazionali correnti, per esempio Monday, Tuesday, ..., ecc., in inglese.
MONTH	Mese nelle opzioni internazionali correnti, per esempio January, February,
DD/MM/YY	15/01/63
DD/MM/YYYY	15/01/1963
MM/DD/YY	01/15/63
MM/DD/YYYY	01/15/1963
DD-MM-YY	15-01-63
DD-MM-YYYY	15-01-1963
MM-DD-YY	01-15-63
MM-DD-YYYY	01-15-1963
DD.MM.YY	15.01.63
DD.MM.YYYY	15.01.1963
MM.DD.YY	01.15.63
MM.DD.YYYY	01.15.1963
DD-MON-YY	15-JAN-63, 15-jan-63, 15-Jan-63
DD/MON/YY	15/JAN/63, 15/jan/63, 15/Jan/63
DD.MON.YY	15.JAN.63, 15.jan.63, 15.Jan.63
DD-MON-YYYY	15-JAN-1963, 15-jan-1963, 15-Jan-1963
DD/MON/YYYY	15/JAN/1963, 15/jan/1963, 15/Jan/1963
DD.MON.YYYY	15.JAN.1963, 15.jan.1963, 15.Jan.1963
MON YYYY	Jan 2004

Formato	Esempi
q Q YYYY	Data rappresentata come una cifra (1-4) rappresentante il trimestre seguito dalla lettera <i>T</i> e un anno a quattro cifre, per esempio 25 dicembre 2004 viene rappresentato nel seguente modo 4 Q 2004.
ww WK YYYY	Numero a due cifre rappresentante la settimana dell'anno, seguito dalle lettere <i>ST</i> e un anno a quattro cifre. La settimana dell'anno viene calcolata presupponendo che il primo giorno della settimana è Lunedì e che esiste almeno un giorno nella prima settimana.

Time

Il linguaggio CLEM supporta i formati di ora seguenti.

Formato	Esempi
HHMMSS	120112, 010101, 221212
HHMM	1223, 0745, 2207
MMSS	5558, 0100
HH:MM:SS	12:01:12, 01:01:01, 22:12:12
HH:MM	12:23, 07:45, 22:07
MM:SS	55:58, 01:00
(H)H:(M)M:(S)S	12:1:12, 1:1:1, 22:12:12
(H)H:(M)M	12:23, 7:45, 22:7
(M)M:(S)S	55:58, 1:0
HH.MM.SS	12.01.12, 01.01.01, 22.12.12
HH.MM	12.23, 07.45, 22.07
MM.SS	55.58, 01.00
(H)H.(M)M.(S)S	12.1.12, 1.1.1, 22.12.12
(H)H.(M)M	12.23, 7.45, 22.7
(M)M.(S)S	55.58, 1.0

Operatori CLEM

Sono disponibili i seguenti operatori:

Operazione	Commenti	Precedenza (vedere la sezione seguente)
or	Utilizzata tra due espressioni CLEM. Restituisce il valore vero se una delle due espressioni è vera oppure se entrambe sono vere.	10
and	Utilizzata tra due espressioni CLEM. Restituisce il valore vero se entrambe le espressioni sono vere.	9

Operazione	Commenti	Precedenza (vedere la sezione seguente)
=	Utilizzata tra qualsiasi coppia di elementi confrontabili. Restituisce vero se ELEMENTO1 è uguale a ELEMENTO2.	7
==	Identico a =.	7
/=	Utilizzata tra qualsiasi coppia di elementi confrontabili. Restituisce vero se ELEMENTO1 <i>non</i> è uguale a ELEMENTO2.	7
/==	Identico a /=.	7
>	Utilizzata tra qualsiasi coppia di elementi confrontabili. Restituisce vero se ELEMENTO1 è maggiore di ELEMENTO2.	6
>=	Utilizzata tra qualsiasi coppia di elementi confrontabili. Restituisce vero se ELEMENTO1 è maggiore o uguale a ELEMENTO2.	6
<	Utilizzata tra qualsiasi coppia di elementi confrontabili. Restituisce vero se ELEMENTO1 è minore di ELEMENTO2.	6
<=	Utilizzata tra qualsiasi coppia di elementi confrontabili. Restituisce vero se ELEMENTO1 è minore o uguale a ELEMENTO2.	6
&&=_0	Utilizzata tra due numeri interi. Equivalente dell'espressione Booleana $INT1 \ \&\& \ INT2 = 0$.	6
&&/=_0	Utilizzata tra due numeri interi. Equivalente dell'espressione Booleana $INT1 \ \&\& \ INT2 \neq 0$.	6
+	Somma due numeri: $NUM1 + NUM2$.	5
><	Concatena due stringhe, per esempio $STRING1 \ >< \ STRING2$.	5
-	Sottrae un numero dall'altro: $NUM1 - NUM2$. Può inoltre precedere un numero: $- \ NUM$.	5
*	Utilizzata per moltiplicare due numeri: $NUM1 * NUM2$.	4
&&	Utilizzata tra due numeri interi. Il risultato è l'AND bitwise di INT1 e INT2.	4
&&~~	Utilizzata tra due numeri interi. Il risultato è l'AND bitwise di INT1 e il complemento bitwise di INT2.	4
	Utilizzata tra due numeri interi. Il risultato è l'OR inclusivo bitwise di INT1 e INT2.	4
~~	Utilizzata prima di un numero intero. Genera il complemento bitwise di INT.	4

Operazione	Commenti	Precedenza (vedere la sezione seguente)
&	Utilizzata tra due numeri interi. Il risultato è l'OR esclusivo bitwise di INT1 e INT2.	4
INT1 << N	Utilizzata tra due numeri interi. Genera lo schema di bit di INT con spostamento a sinistra di N posizioni.	4
INT1 >> N	Utilizzata tra due numeri interi. Genera lo schema di bit di INT con spostamento a destra di N posizioni.	4
/	Utilizzata per dividere un numero per un altro: NUM1 / NUM2.	4
**	Utilizzata tra due numeri: BASE ** POTENZA. Restituisce BASE elevata alla potenza POTENZA.	3
rem	Utilizzata tra due numeri interi: INT1 rem INT2. Restituisce il resto, INT1 - (INT1 div INT2) * INT2.	2
div	Utilizzata tra due numeri interi: INT1 div INT2. Esegue una divisione di interi.	2

Precedenza degli operatori

Le precedenze influiscono sull'analisi delle espressioni complesse, in particolare delle espressioni non racchiuse tra parentesi e contenenti più operatori infissi. Ad esempio,

$3 + 4 * 5$

viene analizzato come $3 + (4 * 5)$ anziché $(3 + 4) * 5$ poiché le precedenze relative determinano che * deve essere analizzato prima di +. Nel linguaggio CLEM, a ogni operatore è associato un valore di precedenza. Minore è il valore, maggiore è l'importanza che gli viene attribuita nell'elenco di analisi. Ciò significa che verrà elaborato prima degli altri operatori con valori di precedenza maggiori.

Riferimenti alle funzioni

Per le operazioni con i dati in IBM® SPSS® Modeler sono disponibili le seguenti funzioni CLEM. È possibile inserire queste funzioni sotto forma di codice in numerose finestre di dialogo, tra cui i nodi Nuovo campo e Crea flag, oppure utilizzare il Generatore espressioni per creare espressioni CLEM valide senza che sia necessario memorizzare elenchi di funzioni o nomi di campi.

Tipo di funzione	Descrizione
Informazioni	Utilizzate per ottenere informazioni sui valori dei campi. Per esempio, la funzione is_string restituisce "vero" per tutti i record il cui tipo è costituito da una stringa.
Conversione	Utilizzate per creare nuovi campi o convertire tipi di archiviazione. Per esempio, la funzione to_timestamp converte il campo selezionato in un timestamp.

Tipo di funzione	Descrizione
Confronto	Utilizzate per confrontare i valori dei campi tra di loro oppure con una stringa determinata. Per esempio, <= viene utilizzata per stabilire se i valori di due campi siano uno uguale o inferiore all'altro.
Logiche	Utilizzate per eseguire operazioni logiche, quali le operazioni if, then, else.
Numeriche	Utilizzate per eseguire calcoli numerici, quale il logaritmo naturale dei valori dei campi.
Trigonometriche	Utilizzate per eseguire calcoli trigonometrici, quale l'arcocoseno di un angolo specificato.
Probabilità	Restituiscono probabilità in base a varie distribuzioni, quali la probabilità che un valore dalla distribuzione <i>t</i> per studenti sia inferiore a un determinato valore.
Bitwise	Utilizzate per manipolare i numeri interi come schemi di bit.
Casuali	Utilizzate per la selezione casuale di elementi o la creazione casuale di numeri.
Stringa	Utilizzate per eseguire numerose operazioni sulle stringhe, per esempio stripchar, che consente di rimuovere un carattere specificato.
SoundEx	Utilizzate per trovare stringhe quando non se ne conosce l'esatta ortografia, in base a ipotesi fonetiche sulla pronuncia di determinate lettere.
Data e orario	Utilizzate per eseguire svariate operazioni sui campi di data, ora e timestamp.
Sequenza	Utilizzate per avere informazioni sulla sequenza di record di un insieme di dati o per eseguire operazioni in base a tale sequenza.
Globali	Utilizzate per accedere a valori globali creati da un nodo Globali. Per esempio, @MEAN viene utilizzata per fare riferimento alla media di tutti i valori di un campo nell'intero insieme di dati.
Vuoti e nulli	Utilizzate per accedere, contrassegnare e spesso riempire i vuoti specificati dall'utente o i valori mancanti del sistema. Per esempio, la funzione @BLANK(FIELD) viene utilizzata per assegnare un flag true a record contenenti valori vuoti.
Campi speciali	Utilizzate per denotare i campi specifici da esaminare. Per esempio, la funzione @FIELD viene utilizzata quando si derivano più campi.

Convenzioni utilizzate nelle descrizioni delle funzioni

Le convenzioni seguenti vengono utilizzate in questa guida quando fanno riferimento agli elementi in una funzione.

Convenzione	Descrizione
<i>BOOL</i>	Valore booleano o flag, per esempio vero o falso.
<i>NUM, NUM1, NUM2</i>	Qualsiasi numero.
<i>REALE, REALE1, REALE2</i>	Qualsiasi numero come 1.234 o -77.01.
<i>INT, INT1, INT2</i>	Qualsiasi numero come 1 o -77.
<i>CAR</i>	Un codice di carattere come 'A'.
<i>STRING</i>	Una stringa come "referrerID".
<i>LIST</i>	Un elenco di elementi come ["abc" "def"].
<i>ELEMENTO</i>	Un campo come Customer o extract_concept.

Convenzione	Descrizione
<i>DATE</i>	Un campo data come <i>start_date</i> , i cui valori sono in un formato quale DD-MON-YYYY.
<i>ORA</i>	Un campo ora come <i>power_flux</i> , i cui valori sono in un formato quale HHMMSS.

Le funzioni in questa guida vengono elencate con la funzione in una colonna, il tipo di risultato (intero, stringa, ecc.) in un'altra e una descrizione (se disponibile) in una terza colonna. Per esempio, di seguito è riportata la descrizione della funzione *rem*.

Funzione	Risultato	Descrizione
<i>INT1 rem INT2</i>	<i>Number</i>	Restituisce il resto della divisione di <i>INT1</i> diviso per <i>INT2</i> . Per esempio, $INT1 - (INT1 \text{ div } INT2) * INT2$.

I dettagli sulle convenzioni di utilizzo, quali la modalità di elencare gli elementi oppure di specificare i caratteri in una funzione, vengono descritti altrove. [Per ulteriori informazioni, vedere l'argomento Tipi di dati CLEM a pag. 119.](#)

Funzioni di informazione

Le funzioni di informazione vengono utilizzate per ottenere informazioni sui valori di un determinato campo e in genere per derivare campi flag. Per esempio, è possibile utilizzare la funzione *@BLANK* per creare un campo flag che indica i record contenenti valori vuoti per il campo selezionato. Allo stesso modo, è possibile controllare il tipo di archiviazione per un campo utilizzando qualsiasi funzione relativa al tipo di archiviazione, per esempio *is_string*.

Funzione	Risultato	Descrizione
<i>@BLANK(FIELD)</i>	<i>Booleana</i>	Restituisce un valore vero per tutti i record contenenti valori vuoti in base alle regole di gestione dei vuoti impostate in un nodo Input o Tipo a monte mediante la scheda Tipi. Si noti che questa funzione non può essere chiamata da uno script. Per ulteriori informazioni, vedere l'argomento Espressioni CLEM negli script in il capitolo 3 in IBM SPSS Modeler 14.2 Guida per script e automazione.
<i>@NULL(ITEM)</i>	<i>Booleana</i>	Restituisce vero per tutti i record contenenti valori non definiti. I valori non definiti sono rappresentati da valori di sistema nulli, visualizzati in IBM® SPSS® Modeler come \$null\$. Si noti che questa funzione non può essere chiamata da uno script. Per ulteriori informazioni, vedere l'argomento Espressioni CLEM negli script in il capitolo 3 in IBM SPSS Modeler 14.2 Guida per script e automazione.
<i>is_date(ITEM)</i>	<i>Booleana</i>	Restituisce vero per tutti i record il cui tipo è rappresentato da una data.
<i>is_datetime(ITEM)</i>	<i>Booleana</i>	Restituisce vero per tutti i record il cui tipo è rappresentato da una data, un'ora o un timestamp.
<i>is_integer(ITEM)</i>	<i>Booleana</i>	Restituisce vero per tutti i record il cui tipo è rappresentato da un numero intero.
<i>is_number(ITEM)</i>	<i>Booleana</i>	Restituisce vero per tutti i record il cui tipo è rappresentato da un numero.

Funzione	Risultato	Descrizione
is_real(ITEM)	<i>Booleana</i>	Restituisce vero per tutti i record il cui tipo è rappresentato da un numero reale.
is_string(ITEM)	<i>Booleana</i>	Restituisce vero per tutti i record il cui tipo è rappresentato da una stringa.
is_time(ITEM)	<i>Booleana</i>	Restituisce vero per tutti i record il cui tipo è rappresentato da un'ora.
is_timestamp(ITEM)	<i>Booleana</i>	Restituisce vero per tutti i record il cui tipo è rappresentato da un timestamp.

Funzioni di conversione

Le funzioni di conversione consentono di creare nuovi campi e convertire il tipo di archiviazione dei file esistenti. Per esempio, per formare nuove stringhe è possibile unire o separare stringhe. Per unire due stringhe, utilizzare l'operatore >>. Per esempio, se il campo Site include il valore "BRAMLEY", allora "xx" >> Site restituisce "xxBRAMLEY". Il risultato di >> è sempre rappresentato da una stringa, anche se gli argomenti non sono stringhe. Di conseguenza, se il campo V1 è 3 e il campo V2 è 5, allora V1 >> V2 restituisce "35" (una stringa, non un numero).

Le funzioni di conversione, e qualsiasi altra funzione che richieda un tipo di input specifico, quale un valore di data o ora, dipendono dai formati correnti specificati nella finestra di dialogo Opzioni stream. Per esempio, se si desidera convertire un campo stringa con valori *Gen 2003*, *Feb 2003* e così via, selezionare il formato di data corrispondente MES AAAA come formato di data di default per lo stream. [Per ulteriori informazioni, vedere l'argomento Impostazione delle opzioni per gli stream in il capitolo 5 a pag. 51.](#)

Funzione	Risultato	Descrizione
ITEM1 >> ITEM2	<i>String</i>	Esegue la concatenazione dei valori di due campi e restituisce la stringa risultante come <i>ELEMENTO1ELEMENTO2</i> .
to_integer(ITEM)	<i>Intero</i>	Converte il tipo di archiviazione del campo specificato in un numero intero.
to_real(ITEM)	<i>Reale</i>	Converte il tipo di archiviazione del campo specificato in un numero reale.
to_number(ITEM)	<i>Number</i>	Converte il tipo di archiviazione del campo specificato in un numero.
to_string(ITEM)	<i>String</i>	Converte il tipo di archiviazione del campo specificato in una stringa.
to_time(ITEM)	<i>Time</i>	Converte il tipo di archiviazione del campo specificato in un'ora.
to_date(ITEM)	<i>Date</i>	Converte il tipo di archiviazione del campo specificato in una data.
to_timestamp(ITEM)	<i>Timestamp</i>	Converte il tipo di archiviazione del campo specificato in un timestamp.

Funzione	Risultato	Descrizione
to_datetime(ITEM)	Data e ora	Converte il tipo di archiviazione del campo specificato in un valore di data, ora o timestamp.
datetime_date(ITEM)	Date	Restituisce il valore di data per un <i>numero</i> , una <i>stringa</i> o un <i>timestamp</i> . Si noti che questa è la sola funzione che consente di convertire un numero (in secondi) nuovamente in una data. Se ITEM è una stringa, crea una data analizzando una stringa nel formato di data corrente. La funzione verrà eseguita correttamente solo se il formato di data specificato nella finestra di dialogo Proprietà stream è corretto. Se ITEM è un numero, viene interpretato come numero di secondi dalla data (o epoca) di riferimento. Le frazioni di giorno vengono troncate. Se ITEM è un timestamp, viene restituita la parte relativa alla data del timestamp. Se ITEM è una data, la data viene restituita invariata.

Funzioni di confronto

Le funzioni di confronto vengono utilizzate per confrontare i valori dei campi tra di loro oppure con una stringa specificata. Per esempio, è possibile controllare l'uguaglianza delle stringhe utilizzando =. Un esempio di verifica dell'uguaglianza delle stringhe è rappresentato da: `Class = "class 1"`.

Ai fini del confronto numerico, *maggiore* indica una maggiore vicinanza all'infinito positivo, mentre *minore* corrisponde a una maggiore prossimità all'infinito negativo. Tutti i numeri negativi sono quindi minori di qualsiasi numero positivo.

Funzione	Risultato	Descrizione
count_equal(ITEM1, LIST)	Intero	Restituisce il numero di valori da un elenco di campi che sono uguali a <i>ELEMENTO1</i> oppure nullo se <i>ELEMENTO1</i> è nullo. Per ulteriori informazioni, vedere l'argomento Riepilogo di più campi in il capitolo 7 a pag. 108.
count_greater_than(ITEM1, LIST)	Intero	Restituisce il numero di valori da un elenco di campi che sono maggiori di <i>ELEMENTO1</i> oppure nullo se <i>ELEMENTO1</i> è nullo.
count_less_than(ITEM1, LIST)	Intero	Restituisce il numero di valori da un elenco di campi che sono minori di <i>ELEMENTO1</i> oppure nullo se <i>ELEMENTO1</i> è nullo.
count_not_equal(ITEM1, LIST)	Intero	Restituisce il numero di valori da un elenco di campi che non sono uguali a <i>ELEMENTO1</i> oppure nullo se <i>ELEMENTO1</i> è nullo.
count_nulls(LIST)	Intero	Restituisce il numero di valori nulli da un elenco di campi.
count_non_nulls(LIST)	Intero	Restituisce il numero di valori non nulli da un elenco di campi.
date_before(ITEM1, ITEM2)	Booleana	Viene utilizzata per controllare l'ordine dei valori di data. Restituisce un valore vero se <i>ITEM1</i> precede <i>ITEM2</i> .
first_index(ITEM, LIST)	Intero	Restituisce l'indice del primo campo contenente <i>ELEMENTO</i> da un ELENCO di campi oppure 0 se il valore non viene trovato. Supportata solo per tipi stringa, numero intero e numero reale. Per ulteriori informazioni, vedere l'argomento Utilizzo di dati a risposta multipla in il capitolo 7 a pag. 109.

Funzione	Risultato	Descrizione
first_non_null(LIST)	<i>Qualsiasi</i>	Restituisce il primo valore non nullo nell'elenco di campi fornito. Sono supportati tutti i tipi di archiviazione.
first_non_null_index(LIST)	<i>Intero</i>	Restituisce l'indice del primo campo dell'ELENCO specificato contenente un valore non nullo oppure 0 se tutti i valori sono nulli. Sono supportati tutti i tipi di archiviazione.
ITEM1 = ITEM2	<i>Booleana</i>	Restituisce vero per i record in cui <i>ELEMENTO1</i> è uguale a <i>ELEMENTO2</i> .
ITEM1 /= ITEM2	<i>Booleana</i>	Restituisce vero se le due stringhe non sono identiche e 0 se sono identiche.
ITEM1 < ITEM2	<i>Booleana</i>	Restituisce vero per i record in cui <i>ELEMENTO1</i> è minore di <i>ELEMENTO2</i> .
ITEM1 <= ITEM2	<i>Booleana</i>	Restituisce vero per i record in cui <i>ELEMENTO1</i> è minore o uguale a <i>ELEMENTO2</i> .
ITEM1 > ITEM2	<i>Booleana</i>	Restituisce vero per i record in cui <i>ELEMENTO1</i> è maggiore di <i>ELEMENTO2</i> .
ITEM1 >= ITEM2	<i>Booleana</i>	Restituisce vero per i record in cui <i>ELEMENTO1</i> è maggiore o uguale a <i>ELEMENTO2</i> .
last_index(ITEM, LIST)	<i>Intero</i>	Restituisce l'indice dell'ultimo campo contenente ELEMENTO da un ELENCO di campi oppure 0 se il valore non viene trovato. Supportata solo per tipi stringa, numero intero e numero reale. Per ulteriori informazioni, vedere l'argomento Utilizzo di dati a risposta multipla in il capitolo 7 a pag. 109.
last_non_null(LIST)	<i>Qualsiasi</i>	Restituisce l'ultimo valore non nullo nell'elenco di campi fornito. Sono supportati tutti i tipi di archiviazione.
last_non_null_index(LIST)	<i>Intero</i>	Restituisce l'indice dell'ultimo campo dell'ELENCO specificato contenente un valore non nullo oppure 0 se tutti i valori sono nulli. Sono supportati tutti i tipi di archiviazione.
max(ITEM1, ITEM2)	<i>Qualsiasi</i>	Restituisce il maggiore dei due elementi, <i>ELEMENTO1</i> o <i>ELEMENTO2</i> .
max_index(LIST)	<i>Intero</i>	Restituisce l'indice del campo contenente il valore massimo da un elenco di campi numerici oppure 0 se tutti i valori sono nulli. Per esempio, se il terzo campo elencato contiene il valore massimo, viene restituito il valore di indice 3. Se più campi contengono il valore massimo, viene restituito quello elencato per primo (più a sinistra). Per ulteriori informazioni, vedere l'argomento Utilizzo di dati a risposta multipla in il capitolo 7 a pag. 109.
max_n(LIST)	<i>Number</i>	Restituisce il valore massimo da un elenco di campi numerici oppure nullo se tutti i valori di campo sono nulli. Per ulteriori informazioni, vedere l'argomento Riepilogo di più campi in il capitolo 7 a pag. 108.
member(ITEM, LIST)	<i>Booleana</i>	Restituisce vero se <i>ELEMENTO</i> è membro dell'elenco <i>ELENCO</i> specificato. In caso contrario, verrà restituito un valore falso. È inoltre possibile specificare un elenco di nomi di campo. Per ulteriori informazioni, vedere l'argomento Riepilogo di più campi in il capitolo 7 a pag. 108.
min(ITEM1, ITEM2)	<i>Qualsiasi</i>	Restituisce il minore dei due elementi, <i>ELEMENTO1</i> o <i>ELEMENTO2</i> .

Funzione	Risultato	Descrizione
min_index(LIST)	Intero	Restituisce l'indice del campo contenente il valore minimo da un elenco di campi numerici oppure 0 se tutti i valori sono nulli. Per esempio, se il terzo campo elencato contiene il valore minimo, viene restituito il valore di indice 3. Se più campi contengono il valore minimo, viene restituito quello elencato per primo (più a sinistra). Per ulteriori informazioni, vedere l'argomento Utilizzo di dati a risposta multipla in il capitolo 7 a pag. 109.
min_n(LIST)	Number	Restituisce il valore minimo da un elenco di campi numerici oppure nullo se tutti i valori di campo sono nulli.
time_before(TIME1, TIME2)	Booleana	Viene utilizzata per controllare l'ordine dei valori di ora. Restituisce un valore vero se <i>ORA1</i> precede <i>ORA2</i> .
value_at(INT, LIST)		Restituisce il valore di ogni campo elencato alla distanza INT oppure NULLO se la distanza non è compresa nell'intervallo di valori validi, cioè minore di 1 o maggiore del numero dei campi elencati. Sono supportati tutti i tipi di archiviazione.

Funzioni logiche

Le espressioni CLEM possono essere utilizzate per eseguire operazioni logiche.

Funzione	Risultato	Descrizione
COND1 and COND2	Booleana	Questa operazione rappresenta una congiunzione logica e restituisce un valore vero se <i>COND1</i> e <i>COND2</i> sono entrambe vere. Se <i>COND1</i> è falsa, la valutazione di <i>COND2</i> non verrà eseguita. Si può quindi disporre di congiunzioni in cui <i>COND1</i> verifica innanzitutto la validità di un'operazione in <i>COND2</i> . Per esempio, <code>length(Label) >=6</code> e <code>Label(6) = 'x'</code> .
COND1 or COND2	Booleana	Questa operazione rappresenta una disgiunzione logica inclusiva e restituisce un valore vero se <i>COND1</i> o <i>COND2</i> risulta vera oppure se entrambe vengono rilevate come vere. Se <i>COND1</i> è vera, la valutazione di <i>COND2</i> non verrà eseguita.
not(COND)	Booleana	Questa operazione rappresenta una negazione logica e restituisce un valore vero se <i>COND</i> è falsa. In caso contrario, questa operazione restituirà un valore 0.
if COND then EXPR1 else EXPR2 endif	Qualsiasi	Questa operazione rappresenta una valutazione condizionale. Se <i>COND</i> è vera, questa operazione restituirà il risultato di <i>EXPR1</i> . Altrimenti, verrà restituito il risultato della valutazione di <i>EXPR_N</i> .
if COND1 then EXPR1 elseif COND2 then EXPR2 else EXPR_N endif	Qualsiasi	Questa operazione rappresenta una valutazione condizionale con più diramazioni. Se <i>COND1</i> risulta vera, questa operazione restituirà il risultato di <i>EXPR1</i> . In caso contrario, se <i>COND2</i> è vera, restituirà il risultato della valutazione di <i>EXPR2</i> . Altrimenti, verrà restituito il risultato della valutazione di <i>EXPR_N</i> .

Funzioni numeriche

CLEM include alcune funzioni numeriche di uso comune.

Funzione	Risultato	Descrizione
-NUM	Number	Viene utilizzata per negare <i>NUM</i> . Restituisce il numero corrispondente con segno opposto.
NUM1 + NUM2	Number	Restituisce la somma di <i>NUM1</i> e <i>NUM2</i> .
codice -NUM2	Number	Restituisce il risultato della sottrazione del valore di <i>NUM2</i> da <i>NUM1</i> .
NUM1 * NUM2	Number	Restituisce il valore di <i>NUM1</i> moltiplicato per <i>NUM2</i> .
NUM1 / NUM2	Number	Restituisce il resto della divisione di <i>NUM1</i> diviso per <i>NUM2</i> .
INT1 div INT2	Number	Viene utilizzata per eseguire la divisione di numeri interi. Restituisce il valore di <i>INT1</i> diviso per <i>INT2</i> .
INT1 rem INT2	Number	Restituisce il resto della divisione di <i>INT1</i> diviso per <i>INT2</i> . Per esempio, $INT1 - (INT1 \text{ div } INT2) * INT2$.
INT1 mod INT2	Number	Questa funzione è obsoleta. Utilizzare la funzione <i>rem</i> .
BASE ** POWER	Number	Restituisce il valore <i>BASE</i> elevato alla potenza <i>POTENZA</i> . Entrambi i valori possono essere rappresentati da qualsiasi numero, ma <i>BASE</i> non deve essere uguale a zero se <i>POTENZA</i> corrisponde a un valore zero di tipo diverso dal numero intero 0. Se <i>POTENZA</i> rappresenta un numero intero, il calcolo viene eseguito mediante moltiplicazioni successive delle potenze di <i>BASE</i> . Di conseguenza, se <i>BASE</i> è un numero intero, il risultato sarà rappresentato da un numero intero. Se <i>POTENZA</i> corrisponde al numero intero 0, il risultato sarà sempre rappresentato da un 1 dello stesso tipo di <i>BASE</i> . In caso contrario, ovvero se <i>POWER</i> non corrisponde a un numero intero, il risultato verrà calcolato come $\exp(POWER * \log(BASE))$.
abs(NUM)	Number	Restituisce il valore assoluto di <i>NUM</i> , che corrisponde sempre a un numero dello stesso tipo.
exp(NUM)	Reale	Restituisce <i>e</i> elevato alla potenza <i>NUM</i> , dove <i>e</i> rappresenta la base dei logaritmi naturali.
fracof(NUM)	Reale	Restituisce la parte frazionaria di <i>NUM</i> , definita come $NUM - \text{intof}(NUM)$.
intof(NUM)	Intero	Tronca l'argomento in modo da ottenere un numero intero. Restituisce il numero intero dello stesso segno di <i>NUM</i> e con l'ampiezza maggiore per cui $\text{abs}(INT) \leq \text{abs}(NUM)$.
log(NUM)	Reale	Restituisce il logaritmo naturale (base <i>e</i>) di <i>NUM</i> , che non deve essere uno zero di alcun tipo.
log10(NUM)	Reale	Restituisce il logaritmo in base 10 di <i>NUM</i> , che non deve essere uno zero di alcun tipo. Questa funzione è definita come $\log(NUM) / \log(10)$.
negate(NUM)	Number	Viene utilizzata per negare <i>NUM</i> . Restituisce il numero corrispondente con segno opposto.
round(NUM)	Intero	Viene utilizzata per arrotondare <i>NUM</i> a un numero intero come $\text{intof}(NUM+0.5)$ se <i>NUM</i> è positivo oppure come $\text{intof}(NUM-0.5)$ se <i>NUM</i> è negativo.
sign(NUM)	Number	Viene utilizzata per determinare il segno di <i>NUM</i> . Se <i>NUM</i> è un numero intero, questa operazione restituirà -1, 0 o 1. Se <i>NUM</i> è un numero reale, restituirà -1.0, 0.0, or 1.0, a seconda che <i>NUM</i> sia negativo, uguale a zero o positivo.

Funzione	Risultato	Descrizione
sqrt(NUM)	Reale	Restituisce la radice quadrata di NUM. NUM deve essere positivo.
sum_n(LIST)	Number	Restituisce la somma di valori da un elenco di campi numerici oppure nullo se tutti i valori di campo sono nulli. Per ulteriori informazioni, vedere l'argomento Riepilogo di più campi in il capitolo 7 a pag. 108.
mean_n(LIST)	Number	Restituisce il valore medio da un elenco di campi numerici oppure nullo se tutti i valori di campo sono nulli.
sdev_n(LIST)	Number	Restituisce la deviazione standard da un elenco di campi numerici oppure nullo se tutti i valori di campo sono nulli.

Funzioni trigonometriche

Tutte le funzioni illustrate in questa sezione accettano un angolo come argomento oppure ne restituiscono uno come risultato. In entrambi i casi, le unità dell'angolo, rappresentate da radianti o gradi, vengono controllate mediante l'impostazione dell'opzione corrispondente dello stream.

Funzione	Risultato	Descrizione
arccos(NUM)	Reale	Calcola l'arcocoseno dell'angolo specificato.
arccosh(NUM)	Reale	Calcola l'arcocoseno iperbolico dell'angolo specificato.
arcsin(NUM)	Reale	Calcola l'arcoseno dell'angolo specificato.
arcsinh(NUM)	Reale	Calcola l'arcoseno iperbolico dell'angolo specificato.
arctan(NUM)	Reale	Calcola l'arcotangente dell'angolo specificato.
arctan2(NUM_Y, NUM_X)	Reale	Calcola l'arcotangente di NUM_Y / NUM_X e utilizza i segni dei due numeri per derivare le informazioni sui quadranti. Il risultato è un numero reale nell'intervallo $-\pi < \text{ANGLE} \leq \pi$ (radians) – $-180 < \text{ANGLE} \leq 180$ (degrees)
arctanh(NUM)	Reale	Calcola l'arcotangente iperbolico dell'angolo specificato.
cos(NUM)	Reale	Calcola il coseno dell'angolo specificato.
cosh(NUM)	Reale	Calcola il coseno iperbolico dell'angolo specificato.
pi	Reale	Questa costante è la migliore approssimazione reale di pi.
sin(NUM)	Reale	Calcola il seno dell'angolo specificato.
sinh(NUM)	Reale	Calcola il seno iperbolico dell'angolo specificato.
tan(NUM)	Reale	Calcola la tangente dell'angolo specificato.
tanh(NUM)	Reale	Calcola la tangente iperbolica dell'angolo specificato.

Funzioni di probabilità

Le funzioni di probabilità restituiscono probabilità in base a varie distribuzioni, quali la probabilità che un valore dalla distribuzione t di Student sia minore di un determinato valore.

Funzione	Risultato	Descrizione
<code>cdf_chisq(NUM, DF)</code>	Reale	Restituisce la probabilità che un valore dalla distribuzione chi-quadrato con i gradi di libertà specificati sia inferiore al numero specificato.
<code>cdf_f(NUM, DF1, DF2)</code>	Reale	Restituisce la probabilità che un valore dalla distribuzione F , con gradi di libertà $DF1$ e $DF2$, sia inferiore al numero specificato.
<code>cdf_normal(NUM, MEAN, STDDEV)</code>	Reale	Restituisce la probabilità che un valore dalla distribuzione normale con la deviazione standard e media specificate sia inferiore al numero specificato.
<code>cdf_t(NUM, DF)</code>	Reale	Restituisce la probabilità che un valore dalla distribuzione t per studenti con i gradi di libertà specificati sia inferiore al numero specificato.

Operazioni bitwise sui numeri interi

Queste funzioni consentono la manipolazione dei numeri interi come schemi di bit rappresentanti valori in complemento a due, in cui alla posizione di bit N corrisponde il peso 2^{*N} . Ai bit viene applicata una numerazione crescente a partire da 0. Queste operazioni vengono eseguite come se il bit del segno di un numero intero venisse esteso a sinistra in modo indefinito. Di conseguenza, in qualsiasi posizione precedente al bit più significativo, un numero intero positivo dispone di 0 bit e un numero intero negativo di 1 bit.

Nota: non è possibile chiamare le funzioni bitwise da script. [Per ulteriori informazioni, vedere l'argomento Espressioni CLEM negli script in il capitolo 3 in IBM SPSS Modeler 14.2 Guida per script e automazione.](#)

Funzione	Risultato	Descrizione
<code>~~ INT1</code>	Intero	Genera il complemento bitwise del numero intero $INT1$. Nel risultato è quindi presente un 1 per ciascuna posizione di bit in cui si riscontra uno 0 in $INT1$. È sempre vero che $~~ INT = -(INT + 1)$. Si noti che questa funzione non può essere chiamata da uno script. Per ulteriori informazioni, vedere l'argomento Espressioni CLEM negli script in il capitolo 3 in IBM SPSS Modeler 14.2 Guida per script e automazione.
<code>INT1 INT2</code>	Intero	Il risultato di questa operazione è l'OR inclusivo bitwise di $INT1$ e $INT2$. Nel risultato è quindi presente un 1 per ogni posizione di bit in cui si riscontra un 1 in $INT1$ o $INT2$ oppure in entrambi.
<code>INT1 /& INT2</code>	Intero	Il risultato di questa operazione è l'OR esclusivo bitwise di $INT1$ e $INT2$. Nel risultato è quindi presente un 1 per ogni posizione di bit in cui si riscontra un 1 in $INT1$ o in $INT2$, ma non in entrambi.

Funzione	Risultato	Descrizione
<code>INT1 && INT2</code>	<i>Intero</i>	Genera l'AND bitwise dei numeri interi <i>INT1</i> e <i>INT2</i> . Nel risultato è quindi presente un 1 per ciascuna posizione di bit in cui si riscontra un 1 sia in <i>INT1</i> che in <i>INT2</i> .
<code>INT1 && ~ INT2</code>	<i>Intero</i>	Genera l'AND bitwise di <i>INT1</i> e il complemento bitwise di <i>INT2</i> . Nel risultato è quindi presente un 1 per ogni posizione di bit in cui si riscontra un 1 in <i>INT1</i> e uno 0 in <i>INT2</i> . Equivale a <code>INT1 && (~INT2)</code> e si rivela utile per cancellare i bit di <i>INT1</i> impostati in <i>INT2</i> .
<code>INT << N</code>	<i>Intero</i>	Genera lo schema di bit di <i>INT1</i> con spostamento a sinistra di <i>N</i> posizioni. L'utilizzo di un valore negativo per <i>N</i> determina uno spostamento a destra.
<code>INT >> N</code>	<i>Intero</i>	Genera lo schema di bit di <i>INT1</i> con spostamento a destra di <i>N</i> posizioni. L'utilizzo di un valore negativo per <i>N</i> determina uno spostamento a sinistra.
<code>INT1 &&= _0 INT2</code>	<i>Booleana</i>	Equivale all'espressione booleana <code>INT1 && INT2 == 0</code> , ma offre una maggiore efficienza.
<code>INT1 &&/= _0 INT2</code>	<i>Booleana</i>	Equivale all'espressione booleana <code>INT1 && INT2 == 0</code> , ma offre una maggiore efficienza.
<code>integer_bitcount(INT)</code>	<i>Intero</i>	Conta il numero di bit 1 o 0 nella rappresentazione in complemento a due di <i>INT</i> . Se <i>INT</i> non è negativo, <i>N</i> corrisponderà al numero di bit 1. Se <i>INT</i> è negativo, rappresenterà il numero di bit 0. A causa dell'estensione del segno, in un numero intero non negativo è presente un numero infinito di bit 0 e un numero intero negativo contiene un numero infinito di bit 1. È sempre <code>integer_bitcount(INT) = integer_bitcount(-(INT+1))</code> .
<code>integer_leastbit(INT)</code>	<i>Intero</i>	Restituisce la posizione di bit <i>N</i> dell'insieme di bit meno significativo nel numero intero <i>INT</i> . <i>N</i> rappresenta la più elevata potenza di 2 per cui <i>INT</i> può essere diviso esattamente.
<code>integer_length(INT)</code>	<i>Intero</i>	Restituisce la lunghezza in bit di <i>INT</i> come intero in complemento a due. Ovvero, <i>N</i> è l'intero più piccolo tale che <code>INT < (1 << N)</code> if <code>INT >= 0</code> <code>INT >= (-1 << N)</code> if <code>INT < 0</code> . Se <i>INT</i> non è negativo, per la rappresentazione di <i>INT</i> come numero intero senza segno sarà richiesto un campo di almeno <i>N</i> bit. In alternativa, per rappresentare <i>INT</i> come numero intero con segno, sarà richiesto un minimo di <i>N+1</i> bit, indipendentemente dal segno.
<code>testbit(INT, N)</code>	<i>Booleana</i>	Verifica il bit in posizione <i>N</i> nel numero intero <i>INT</i> e restituisce lo stato del bit <i>N</i> sotto forma di valore booleano, ovvero vero per 1 e falso per 0.

Funzioni casuali

Le funzioni illustrate di seguito vengono utilizzate per la selezione casuale di elementi o la generazione casuale di numeri.

Funzione	Risultato	Descrizione
oneof(LIST)	<i>Qualsiasi</i>	Restituisce un elemento di <i>ELENCO</i> scelto in modo casuale. Gli elementi dell'elenco devono essere specificati come [ITEM1,ITEM2,...,ITEM_N]. Si noti che è inoltre possibile specificare un elenco di nomi di campo. Per ulteriori informazioni, vedere l'argomento Riepilogo di più campi in il capitolo 7 a pag. 108.
random(NUM)	<i>Number</i>	Restituisce un numero casuale distribuito in modo uniforme dello stesso tipo (<i>INT</i> o <i>REALE</i>), da 1 a <i>NUM</i> . Se si utilizza un numero intero, verranno restituiti solo numeri interi. Se si utilizza un numero reale (decimale), verranno restituiti numeri reali la cui precisione decimale è determinata dalle opzioni dello stream. Il numero casuale più grande restituito dalla funzione può essere uguale a <i>NUM</i> .
random0(NUM)	<i>Number</i>	Ha le stesse proprietà di <i>random(NUM)</i> , ma a partire da 0. Il numero casuale più grande restituito dalla funzione non sarà mai uguale a <i>X</i> .

Funzioni stringa

Nel linguaggio CLEM sono supportate le seguenti operazioni con le stringhe:

- Confronto di stringhe
- Creazione di stringhe
- Accesso ai caratteri

In un'espressione CLEM, una stringa è rappresentata da qualsiasi sequenza di caratteri racchiusa tra virgolette doppie corrispondenti ("string quotes"). I caratteri (*CHAR*) possono essere rappresentati da qualsiasi carattere alfanumerico singolo. Sono dichiarati in espressioni CLEM che utilizzano virgolette singole a barra rovesciata nel formato `<carattere>`, per esempio ``z``, ``A`` o ``2``. I caratteri che rappresentano indici fuori intervallo o negativi in una stringa determineranno un comportamento non definito.

Nota: i confronti tra stringhe che utilizzano e non utilizzano il push back SQL potrebbero generare risultati diversi in presenza di spazi finali.

Funzione	Risultato	Descrizione
allbutfirst(N, STRING)	<i>String</i>	Restituisce una stringa risultante dalla rimozione dei primi <i>N</i> caratteri da <i>STRINGA</i> .
allbutlast(N, STRING)	<i>String</i>	Restituisce una stringa risultante dalla rimozione degli ultimi caratteri da <i>STRINGA</i> .
alphabefore(STRING1, STRING2)	<i>Booleana</i>	Viene utilizzata per controllare l'ordine alfabetico delle stringhe. Restituisce vero se <i>STRINGA1</i> precede <i>STRINGA2</i> .

Funzione	Risultato	Descrizione
endstring(LENGTH, STRING)	<i>String</i>	Estrae gli ultimi <i>N</i> caratteri dalla stringa specificata. Se la lunghezza della stringa è inferiore o uguale alla lunghezza specificata, rimane invariata.
hasendstring(STRING, SUBSTRING)	<i>Intero</i>	La funzione è analoga a <code>isendstring(SUBSTRING, STRING)</code> .
hasmidstring(STRING, SUBSTRING)	<i>Intero</i>	La funzione è analoga a <code>ismidstring(SUBSTRING, STRING)</code> (sottostringa incorporata).
hasstartstring(STRING, SUBSTRING)	<i>Intero</i>	La funzione è analoga a <code>isstartstring(SUBSTRING, STRING)</code> .
hassubstring(STRING, N, SUBSTRING)	<i>Intero</i>	La funzione è analoga a <code>issubstring(SUBSTRING, N, STRING)</code> , dove il valore di default di <i>N</i> è 1.
count_substring(STRING, SUBSTRING)	<i>Intero</i>	Restituisce il numero di volte in cui la sottostringa specificata è presente all'interno della stringa. Ad esempio, <code>count_substring("foooo.txt", "oo")</code> restituisce 3.
hassubstring(STRING, SUBSTRING)	<i>Intero</i>	La funzione è analoga a <code>issubstring(SUBSTRING, 1, STRING)</code> , dove il valore di default di <i>N</i> è 1.
isalphacode(CHAR)	<i>Booleana</i>	Restituisce un valore vero se <i>CAR</i> rappresenta un carattere della stringa specificata (spesso un nome di campo) con codice alfabetico. In caso contrario, questa funzione restituirà un valore 0. Per esempio, <code>isalphacode(produce_num(1))</code> .
isendstring(SUBSTRING, STRING)	<i>Intero</i>	Se la stringa <i>STRINGA</i> termina con la sottostringa <i>SOTTOSTRINGA</i> , questa funzione restituirà il numero di indice di <i>SOTTOSTRINGA</i> in <i>STRINGA</i> . In caso contrario, questa funzione restituirà un valore 0.
islowercode(CHAR)	<i>Booleana</i>	Restituisce il valore vero se <i>CAR</i> rappresenta una lettera minuscola per la stringa specificata (spesso un nome di campo). In caso contrario, restituirà il valore 0. Per esempio, sia <code>islowercode(`)</code> che <code>islowercode(country_name(2))</code> sono espressioni valide.
ismidstring(SUBSTRING, STRING)	<i>Intero</i>	Se <i>SOTTOSTRINGA</i> è una sottostringa di <i>STRINGA</i> ma non ha inizio con il primo carattere di <i>STRINGA</i> né termina con l'ultimo, questa funzione restituirà l'indice in corrispondenza del quale inizia la sottostringa. In caso contrario, questa funzione restituirà un valore 0.
isnumbercode(CHAR)	<i>Booleana</i>	Restituisce il valore vero se <i>CAR</i> per la stringa specificata (spesso un nome di campo) rappresenta un carattere con codice numerico. In caso contrario, questa funzione restituirà un valore 0. Per esempio, <code>isnumbercode(product_id(2))</code> .

Funzione	Risultato	Descrizione
isstartstring(SUBSTRING, STRING)	<i>Intero</i>	Se la stringa <i>STRINGA</i> ha inizio con la sottostringa <i>SOTTOSTRINGA</i> , questa funzione restituirà l'indice 1. In caso contrario, restituirà il valore 0.
issubstring(SUBSTRING, N, STRING)	<i>Intero</i>	Cerca una sottostringa uguale a <i>SOTTOSTRINGA</i> all'interno della stringa <i>STRINGA</i> , a partire dal carattere <i>N</i> . Se la ricerca ha esito positivo, questa funzione restituisce il valore di indice intero in corrispondenza del quale ha inizio la sottostringa corrispondente. In caso contrario, restituirà il valore 0. Se non viene specificato, <i>N</i> verrà impostato per default su 1.
issubstring(SUBSTRING, STRING)	<i>Intero</i>	Cerca una sottostringa uguale a <i>SOTTOSTRINGA</i> all'interno della stringa <i>STRINGA</i> , a partire dal carattere <i>N</i> . Se la ricerca ha esito positivo, questa funzione restituisce il valore di indice intero in corrispondenza del quale ha inizio la sottostringa corrispondente. In caso contrario, restituirà il valore 0. Se non viene specificato, <i>N</i> verrà impostato per default su 1.
issubstring_count(SUBSTRING, N, STRING):	<i>Intero</i>	Restituisce l'indice dell' <i>N</i> esima occorrenza della <i>SOTTOSTRINGA</i> all'interno della <i>STRINGA</i> specificata. Se le occorrenze della <i>SOTTOSTRINGA</i> sono inferiori a <i>N</i> , viene restituito 0.
issubstring_lim(SUBSTRING, N, STARTLIM, ENDLIM, STRING)	<i>Intero</i>	Questa funzione è equivalente a <i>issubstring</i> , ma la corrispondenza deve iniziare prima dell'indice <i>LIMINIZIO</i> o a partire da esso e terminare con l'indice <i>LIMFINE</i> o prima di esso. I vincoli <i>LIMINIZIO</i> o <i>LIMFINE</i> possono essere disattivati fornendo un valore falso per questi argomenti, per esempio <i>issubstring_lim(SUBSTRING, N, false, false, STRING)</i> equivale a <i>issubstring</i> .
isuppercode(CHAR)	<i>Booleana</i>	Restituisce un valore vero se <i>CAR</i> rappresenta una lettera maiuscola. In caso contrario, restituirà il valore 0. Per esempio, sia <i>isuppercode('')</i> che <i>isuppercode(country_name(2))</i> sono espressioni valide.
last(CHAR)	<i>String</i>	Restituisce l'ultimo carattere <i>CAR</i> di <i>STRINGA</i> (che deve contenere almeno un carattere).
length(STRING)	<i>Intero</i>	Restituisce la lunghezza della stringa <i>STRINGA</i> , ovvero il numero di caratteri in essa contenuti.

Funzione	Risultato	Descrizione
locchar(CHAR, N, STRING)	Intero	Utilizzata per identificare la posizione dei caratteri nei campi simbolici. La funzione cerca il carattere <i>CAR</i> all'interno della stringa <i>STRINGA</i> , iniziando la ricerca dal carattere <i>Nesimo</i> di <i>STRINGA</i> . Questa funzione restituisce un valore che indica la posizione (a partire da <i>N</i>) in cui si trova il carattere. Se il carattere non viene trovato, questa funzione restituirà un valore 0. Se la funzione include una distanza (<i>N</i>) non valida (per esempio una distanza superiore alla lunghezza della stringa), la funzione restituisce \$null\$. Per esempio, locchar(`n`, 2, web_page) ricerca all'interno del campo <i>web_page</i> il carattere `n` a partire dal secondo carattere del valore del campo. <i>Nota:</i> racchiudere il carattere specificato tra virgolette singole a barra rovesciata.
locchar_back(CHAR, N, STRING)	Intero	Funzione analoga a locchar, tranne per il fatto che la ricerca viene eseguita all'indietro, a partire dall' <i>Nesimo</i> carattere. Per esempio, locchar_back(`n`, 9, web_page) esegue la ricerca all'interno del campo <i>web_page</i> a partire dal nono carattere e procedendo all'indietro verso l'inizio della stringa. Se contiene un margine non valido, per esempio superiore alla lunghezza della stringa, questa funzione restituirà \$null\$. Idealmente, sarà necessario utilizzare locchar_back insieme alla funzione length(<field>) per utilizzare dinamicamente la lunghezza del valore corrente del campo. Per esempio, locchar_back(`n`, (length(web_page)), web_page).
lowertoupper(CHAR) lowertoupper (STRING)	CAR o Stringa	L'input può essere rappresentato da una stringa o da un carattere, utilizzato in questa funzione per restituire un nuovo elemento dello stesso tipo, con i caratteri minuscoli convertiti nei relativi maiuscoli. Per esempio, lowertoupper(`a`), lowertoupper("My string") e lowertoupper(field_name(2)) sono espressioni valide.
matches	Booleana	Restituisce vero se una stringa corrisponde allo schema specificato. Lo schema deve essere una stringa letterale e non può essere un nome di campo contenente uno schema. È possibile includere nello schema un punto interrogativo (?) corrispondente esattamente a un carattere; un asterisco (*) corrisponde a zero o più caratteri. Per indicare un punto interrogativo o un asterisco letterale, ovvero non utilizzato come carattere jolly, è possibile utilizzare la barra rovesciata (\).

Funzione	Risultato	Descrizione
replace(SUBSTRING, NEWSUBSTRING, STRING)	String	All'interno della <i>STRINGA</i> specificata, sostituisce tutte le istanze della <i>SOTTOSTRINGA</i> con la <i>NUOVASOTTOSTRINGA</i> .
replicate(COUNT, STRING)	String	Restituisce una stringa costituita dalla stringa originale copiata il numero di volte specificato.
stripchar(CHAR,STRING)	String	Consente di rimuovere i caratteri specificati da una stringa o un campo. È possibile utilizzare questa funzione, per esempio, per rimuovere dai dati i simboli aggiuntivi, come i simboli di valuta, in modo da ottenere un nome o un numero semplice. Per esempio, la sintassi stripchar('\$', 'Cost') restituisce un nuovo campo in cui il simbolo del dollaro è stato rimosso da tutti i valori. <i>Nota:</i> racchiudere il carattere specificato tra virgolette singole a barra rovesciata.
skipchar(CHAR, N, STRING)	Intero	Cerca nella stringa <i>STRINGA</i> qualsiasi carattere diverso da <i>CAR</i> , a partire dall' <i>Nesimo</i> carattere. Questa funzione restituisce una sottostringa di numeri interi che indica il punto in cui è stato trovato un carattere oppure 0, se ogni carattere successivo a <i>Nesimo</i> corrisponde a <i>CAR</i> . Se contiene un margine non valido, per esempio superiore alla lunghezza della stringa, questa funzione restituirà \$null\$. locchar è spesso utilizzato insieme alle funzioni skipchar per determinare il valore di <i>N</i> (punto dal quale iniziare la ricerca nella stringa). Per esempio, skipchar('s', (locchar('s', 1, "MyString")), "MyString").
skipchar_back(CHAR, N, STRING)	Intero	Questa funzione è analoga a skipchar, tranne che per il fatto che la ricerca viene eseguita a ritroso , a partire dall' <i>Nesimo</i> carattere.
startstring(LENGTH, STRING)	String	Estrae i primi <i>N</i> caratteri dalla stringa specificata. Se la lunghezza della stringa è inferiore o uguale alla lunghezza specificata, rimane invariata.
strmember(CHAR, STRING)	Intero	Equivale a locchar(CHAR, 1, STRING). Restituisce una sottostringa di numeri interi che indica il punto in cui viene rilevata la prima occorrenza di <i>CHAR</i> oppure 0. Se contiene una distanza non valida, per esempio superiore alla lunghezza della stringa, questa funzione restituirà \$null\$.
subscrs(N, STRING)	CAR	Restituisce l' <i>Nesimo</i> carattere <i>CAR</i> della stringa di input <i>STRINGA</i> . Per questa funzione è inoltre disponibile la forma abbreviata STRING(N). Per esempio, lowertoupper("name"(1)) è un'espressione valida.

Funzione	Risultato	Descrizione
substring(N, LEN, STRING)	String	Restituisce una stringa <i>SOTTOSTRINGA</i> , rappresentata dal numero di caratteri di <i>STRINGA</i> indicati da <i>LUNGH</i> , a partire dal carattere in corrispondenza dell'indice <i>N</i> .
substring_between(N1, N2, STRING)	String	Restituisce la sottostringa di <i>STRINGA</i> che ha inizio in corrispondenza dell'indice <i>N1</i> e termina con l'indice <i>N2</i> .
trim(STRING)	String	Rimuove i caratteri spazio bianco iniziale e finale dalla stringa specificata.
trim_start(STRING)	String	Rimuove i caratteri spazio bianco iniziale dalla stringa specificata.
trimend(STRING)	String	Rimuove i caratteri spazio bianco finale dalla stringa specificata.
unicode_char(NUM)	CAR	Restituisce il carattere con valore Unicode <i>NUM</i> .
unicode_value(CHAR)	NUM	Restituisce il valore Unicode di <i>CAR</i>
uppertolower(CHAR) uppertolower (STRING)	CAR o Stringa	L'input può essere rappresentato da una stringa o da un carattere, utilizzato in questa funzione per restituire un nuovo elemento dello stesso tipo, con i caratteri maiuscoli convertiti nei relativi minuscoli. <i>Nota:</i> Ricordarsi di specificare le stringhe con le virgolette doppie e i caratteri con le virgolette singole a barra rovesciata. I nomi semplici di campo non devono essere racchiusi tra virgolette.

Funzioni SoundEx

SoundEx è un metodo utilizzato per trovare stringhe quando si conosce un suono ma non l'esatta ortografia. Sviluppato nel 1918, cerca parole con suoni simili in base a ipotesi fonetiche sulla pronuncia di determinate lettere. Può essere per esempio utilizzato per cercare nomi in un database, in cui ortografia e pronuncia di nomi simili possono variare. L'algoritmo SoundEx di base è documentato in numerose fonti e nonostante i limiti conosciuti (per esempio, la combinazione iniziale ph e la lettera f non corrispondono nonostante il loro suono sia simile) è supportato in qualche forma dalla maggioranza dei database.

Funzione	Risultato	Descrizione
soundex(STRING)	Intero	Restituisce il codice SoundEx di quattro caratteri per la <i>STRINGA</i> specificata.
soundex_difference(STRING1, STRING2)	Intero	Restituisce un numero intero compreso tra 0 e 4 che indica il numero di caratteri che sono identici nella codifica SoundEx per le due stringhe, dove 0 indica nessuna somiglianza e 4 indica una forte somiglianza o stringhe identiche.

Funzioni di data e ora

Nel linguaggio CLEM è disponibile un gruppo di funzioni per la gestione dei campi con archiviazione di tipo datetime delle variabili che rappresentano le date e le ore. I formati di data e ora utilizzati sono specifici di ogni stream e vengono impostati nella finestra di dialogo delle proprietà dello stream. Le funzioni di data e ora analizzano le stringhe di data e ora in base al formato correntemente selezionato.

Quando si specifica un anno in una data che utilizza solo due cifre, ovvero in cui non è indicato il secolo, verrà utilizzato il secolo di default specificato nella finestra di dialogo delle proprietà dello stream.

Nota: Non è possibile chiamare le funzioni data e ora da script. [Per ulteriori informazioni, vedere l'argomento Espressioni CLEM negli script in il capitolo 3 in IBM SPSS Modeler 14.2 Guida per script e automazione.](#)

Funzione	Risultato	Descrizione
@TODAY	String	Se si seleziona Cambia giorni/min nella finestra di dialogo delle proprietà dello stream, questa funzione restituirà la data corrente sotto forma di stringa nel formato di data corrente. Se si utilizza un formato di data a due cifre e non si seleziona Cambia giorni/min, questa funzione restituirà \$null\$ sul server corrente. Si noti che questa funzione non può essere chiamata da uno script. Per ulteriori informazioni, vedere l'argomento Espressioni CLEM negli script in il capitolo 3 in IBM SPSS Modeler 14.2 Guida per script e automazione.
to_time(ITEM)	Time	Converte il tipo di archiviazione del campo specificato in un'ora.
to_date(ITEM)	Date	Converte il tipo di archiviazione del campo specificato in una data.
to_timestamp(ITEM)	Timestamp	Converte il tipo di archiviazione del campo specificato in un timestamp.
to_datetime(ITEM)	Data e ora	Converte il tipo di archiviazione del campo specificato in un valore di data, ora o timestamp.
datetime_date(ITEM)	Date	Restituisce il valore di data per un <i>numero</i> , una <i>stringa</i> o un <i>timestamp</i> . Si noti che questa è la sola funzione che consente di convertire un numero (in secondi) nuovamente in una data. Se ITEM è una stringa, crea una data analizzando una stringa nel formato di data corrente. La funzione verrà eseguita correttamente solo se il formato di data specificato nella finestra di dialogo Proprietà stream è corretto. Se ITEM è un numero, viene interpretato come numero di secondi dalla data (o epoca) di riferimento. Le frazioni di giorno vengono troncate. Se ITEM è un timestamp, viene restituita la parte relativa alla data del timestamp. Se ITEM è una data, la data viene restituita invariata.
date_before(DATE1, DATE2)	Booleana	Restituisce un valore vero se <i>DATE1</i> rappresenta una data o un timestamp precedente a quella rappresentata da <i>DATE2</i> . In caso contrario, questa funzione restituirà un valore 0.

Funzione	Risultato	Descrizione
date_days_difference(DATE1, DATE2)	<i>Intero</i>	Restituisce il tempo in giorni dalla data o dal timestamp rappresentato da <i>DATA1</i> a quello rappresentato da <i>DATA2</i> , sotto forma di numero intero. Se <i>DATA2</i> è antecedente a <i>DATA1</i> , questa funzione restituisce un numero negativo.
date_in_days(DATE)	<i>Intero</i>	Restituisce il tempo in giorni trascorso dalla data di riferimento alla data o al timestamp rappresentato da <i>DATA</i> , sotto forma di numero intero. Se il valore di <i>DATA</i> è precedente alla data di riferimento, questa funzione restituisce un numero negativo. Per ottenere un calcolo corretto è necessario includere una data valida. Per esempio, 29 febbraio 2001 non è una data valida. Infatti, dal momento che il 2001 non è un anno bisestile, questa data non esiste.
date_in_months(DATE)	<i>Reale</i>	Restituisce il tempo in mesi trascorso dalla data di riferimento alla data o al timestamp rappresentato da <i>DATA</i> , sotto forma di numero reale. Tale numero corrisponde a una cifra approssimativa basata su un mese formato da 30.4375 giorni. Se il valore di <i>DATA</i> è precedente alla data di riferimento, questa funzione restituisce un numero negativo. Per ottenere un calcolo corretto è necessario includere una data valida. Per esempio, 29 febbraio 2001 non è una data valida. Infatti, dal momento che il 2001 non è un anno bisestile, questa data non esiste.
date_in_weeks(DATE)	<i>Reale</i>	Restituisce il tempo in settimane trascorso dalla data di riferimento alla data o al timestamp rappresentato da <i>DATA</i> , sotto forma di numero reale. Tale numero è basato su una settimana formata da sette giorni. Se il valore di <i>DATA</i> è precedente alla data di riferimento, questa funzione restituisce un numero negativo. Per ottenere un calcolo corretto è necessario includere una data valida. Per esempio, 29 febbraio 2001 non è una data valida. Infatti, dal momento che il 2001 non è un anno bisestile, questa data non esiste.
date_in_years(DATE)	<i>Reale</i>	Restituisce il tempo in anni trascorso dalla data di riferimento alla data o al timestamp rappresentato da <i>DATA</i> , sotto forma di numero reale. Tale numero corrisponde a una cifra approssimativa basata su un anno formato da 365.25 giorni. Se il valore di <i>DATA</i> è precedente alla data di riferimento, questa funzione restituisce un numero negativo. Per ottenere un calcolo corretto è necessario includere una data valida. Per esempio, 29 febbraio 2001 non è una data valida. Infatti, dal momento che il 2001 non è un anno bisestile, questa data non esiste.
date_months_difference(DATE1, DATE2)	<i>Reale</i>	Restituisce il tempo in mesi dalla data o dal timestamp rappresentato da <i>DATA1</i> a quello rappresentato da <i>DATA2</i> , sotto forma di numero reale. Tale numero corrisponde a una cifra approssimativa basata su un mese formato da 30.4375 giorni. Se <i>DATA2</i> è antecedente a <i>DATA1</i> , questa funzione restituisce un numero negativo.
datetime_date(YEAR, MONTH, DAY)	<i>Date</i>	Crea il valore di data per gli argomenti <i>ANNO</i> , <i>MESE</i> e <i>GIORNO</i> specificati. Gli argomenti devono essere numeri interi.

Funzione	Risultato	Descrizione
datetime_day(DATE)	<i>Intero</i>	Restituisce il giorno del mese di una <i>DATA</i> o un timestamp. Il risultato è un numero intero nell'intervallo da 1 a 31.
datetime_day_name(DAY)	<i>String</i>	Restituisce il nome completo dell'argomento <i>GIORNO</i> specificato. L'argomento deve essere un numero intero nell'intervallo da 1 (domenica) a 7 (sabato).
datetime_hour(TIME)	<i>Intero</i>	Restituisce l'ora dal valore <i>ORA</i> o da un timestamp. Il risultato è un numero intero nell'intervallo da 0 a 23.
datetime_in_seconds(TIME)	<i>Reale</i>	Restituisce la parte in secondi memorizzata in <i>ORA</i> .
datetime_in_seconds(DATE), datetime_in_seconds(DATE-TIME)	<i>Reale</i>	Restituisce il numero accumulato, convertito in secondi, dalla differenza tra il valore <i>DATA</i> o <i>DATAORA</i> corrente e la data di riferimento (1900-01-01).
datetime_minute(TIME)	<i>Intero</i>	Restituisce i minuti dal valore <i>ORA</i> o da un timestamp. Il risultato è un numero intero nell'intervallo da 0 a 59.
datetime_month(DATE)	<i>Intero</i>	Restituisce il mese da una <i>DATA</i> o da un timestamp. Il risultato è un numero intero nell'intervallo da 1 a 12.
datetime_month_name(MONTH)	<i>String</i>	Restituisce il nome completo dell'argomento <i>MESE</i> specificato. L'argomento deve essere un numero intero nell'intervallo da 1 a 12.
datetime_now	<i>Timestamp</i>	Restituisce l'ora corrente come timestamp.
datetime_second(TIME)	<i>Intero</i>	Restituisce i secondi dal valore <i>ORA</i> o da un timestamp. Il risultato è un numero intero nell'intervallo da 0 a 59.
datetime_day_short_name(DAY)	<i>String</i>	Restituisce il nome abbreviato dell'argomento <i>GIORNO</i> specificato. L'argomento deve essere un numero intero nell'intervallo da 1 (domenica) a 7 (sabato).
datetime_month_short_name(MONTH)	<i>String</i>	Restituisce il nome abbreviato dell'argomento <i>MESE</i> specificato. L'argomento deve essere un numero intero nell'intervallo da 1 a 12.
datetime_time(HOUR, MINUTE, SECOND)	<i>Time</i>	Restituisce il valore di ora per gli argomenti <i>ORA</i> , <i>MINUTO</i> , e <i>SECONDO</i> specificati. Gli argomenti devono essere numeri interi.
datetime_time(ITEM)	<i>Time</i>	Restituisce il valore di ora dell' <i>ELEMENTO</i> specificato.
datetime_timestamp(YEAR, MONTH, DAY, HOUR, MINUTE, SECOND)	<i>Timestamp</i>	Restituisce il valore di timestamp per l' <i>ANNO</i> , il <i>MESE</i> , il <i>GIORNO</i> , l' <i>ORA</i> , i <i>MINUTE</i> e i <i>SECONDI</i> specificati.
datetime_timestamp(DATE, TIME)	<i>Timestamp</i>	Restituisce il valore di timestamp per la <i>DATA</i> e l' <i>ORA</i> specificate.
datetime_timestamp (NUMBER)	<i>Timestamp</i>	Restituisce il valore di timestamp del numero di secondi specificato.
datetime_weekday(DATE)	<i>Intero</i>	Restituisce il giorno della settimana dalla <i>DATA</i> o dal timestamp specificato.
datetime_year(DATE)	<i>Intero</i>	Restituisce l'anno da una <i>DATA</i> o da un timestamp. Il risultato è un numero intero come 2002.

Funzione	Risultato	Descrizione
date_weeks_difference (DATE1, DATE2)	Reale	Restituisce il tempo in settimane dalla data o dal timestamp rappresentato da <i>DATA1</i> a quello rappresentato da <i>DATA2</i> , sotto forma di numero reale. Tale numero è basato su una settimana formata da sette giorni. Se <i>DATA2</i> è antecedente a <i>DATA1</i> , questa funzione restituisce un numero negativo.
date_years_difference (DATE1, DATE2)	Reale	Restituisce il tempo in anni dalla data o dal timestamp rappresentato da <i>DATA1</i> a quello rappresentato da <i>DATA2</i> , sotto forma di numero reale. Tale numero corrisponde a una cifra approssimativa basata su un anno formato da 365.25 giorni. Se <i>DATA2</i> è antecedente a <i>DATA1</i> , questa funzione restituisce un numero negativo.
time_before(TIME1, TIME2)	Booleana	Restituisce un valore vero se <i>ORA1</i> rappresenta un'ora o un timestamp precedente a quello rappresentata da <i>ORA2</i> . In caso contrario, questa funzione restituirà un valore 0.
time_hours_difference (TIME1, TIME2)	Reale	Restituisce il tempo trascorso in ore tra le ore o timestamp rappresentati da <i>ORA1</i> e <i>ORA2</i> , sotto forma di numero reale. Se si seleziona Cambia giorni/min nella finestra di dialogo delle proprietà dello stream, un valore superiore di <i>ORA1</i> verrà interpretato come riferito al giorno precedente. Se non si seleziona l'opzione di trasformazione, un valore superiore di <i>ORA1</i> determina la restituzione di un valore negativo.
time_in_hours(TIME)	Reale	Restituisce il tempo in ore rappresentato dall'ora <i>TIME</i> , sotto forma di numero reale. Con il formato di ora HHMM, per esempio l'espressione <code>time_in_hours('0130')</code> restituisce 1.5. <i>TIME</i> può rappresentare un'ora o un timestamp.
time_in_mins(TIME)	Reale	Restituisce il tempo in minuti rappresentato da <i>ORA</i> , sotto forma di numero reale. <i>TIME</i> può rappresentare un'ora o un timestamp.
time_in_secs(TIME)	Intero	Restituisce il tempo in secondi rappresentato da <i>ORA</i> , sotto forma di numero intero. <i>TIME</i> può rappresentare un'ora o un timestamp.
time_mins_difference(TIME1, TIME2)	Reale	Restituisce il tempo trascorso in minuti tra le ore o timestamp rappresentati da <i>ORA1</i> e <i>ORA2</i> , sotto forma di numero reale. Se si seleziona Cambia giorni/min nella finestra di dialogo delle proprietà dello stream, un valore superiore a <i>ORA1</i> verrà interpretato come riferito al giorno precedente (o all'ora precedente, se per il formato corrente sono specificati solo ore e minuti). Se non si seleziona l'opzione di trasformazione, un valore superiore a <i>ORA1</i> determina la restituzione di un valore negativo.
time_secs_difference(TIME1, TIME2)	Intero	Restituisce la differenza in secondi tra le ore o timestamp rappresentati da <i>ORA1</i> e <i>ORA2</i> , sotto forma di numero intero. Se si seleziona Cambia giorni/min nella finestra di dialogo delle proprietà dello stream, un valore superiore a <i>ORA1</i> verrà interpretato come riferito al giorno precedente (o all'ora precedente, se per il formato corrente sono specificati solo ore e minuti). Se non si seleziona l'opzione di trasformazione, un valore superiore di <i>ORA1</i> determina la restituzione di un valore negativo.

Conversione dei valori di data e ora

Si noti che le funzioni di conversione, e qualsiasi altra funzione che richieda un tipo di input specifico quale un valore di data o ora, dipendono dai formati correnti specificati nella finestra di dialogo Opzioni stream. Per esempio, un campo denominato *DATA*, archiviato come stringa con i valori *Gen 2003*, *Feb 2003* e così via, potrebbe essere convertito in campo di archiviazione di date nel modo seguente:

```
to_date(DATA)
```

Perché la conversione funzioni è necessario selezionare il formato di data MES AAAA corrispondente come default per lo stream. [Per ulteriori informazioni, vedere l'argomento Impostazione delle opzioni per gli stream in il capitolo 5 a pag. 51.](#)

Per un esempio di conversione di valori di stringa in date mediante un nodo Riempimento, vedere lo stream *broadband_create_models.str*, installato nella cartella *\Demos* all'interno della sottocartella *streams*. [Per ulteriori informazioni, vedere l'argomento Previsione mediante il nodo Serie storica in il capitolo 15 in Guida alle applicazioni di IBM SPSS Modeler 14.2.](#)

Date archiviate come numeri. Si noti che *DATA* nell'esempio precedente è il nome di un campo, mentre *to_date* è una funzione CLEM. È possibile convertire le date archiviate come numeri utilizzando la funzione *datetime_date*, in cui il numero viene interpretato come numero di secondi trascorsi dalla data (o epoca) di riferimento.

```
datetime_date(DATA)
```

Mediante la conversione di una data in un numero di secondi e viceversa, è possibile effettuare calcoli, quali la data corrente più o meno un numero determinato di giorni. Per esempio:

```
datetime_date((date_in_days(DATA)-7)*60*60*24)
```

Funzioni sequenza

Per alcune operazioni, la sequenza degli eventi si rivela importante. L'applicazione supporta le seguenti sequenze di record:

- Sequenze e serie storiche
- Funzioni sequenza
- Indicizzazione di record
- Calcolo della media, somma e confronto di valori
- Monitoraggio delle modifiche e differenziazione
- @SINCE
- Valori di distanza
- Funzionalità aggiuntive relative alle sequenze

In numerose applicazioni, ogni record in uno stream può essere considerato come un caso singolo, indipendente da tutti gli altri. In queste circostanze, l'ordine dei record è in genere irrilevante.

Per alcune classi di problemi, tuttavia, la sequenza di record è estremamente importante. Si tratta in genere di situazioni relative a serie storiche, in cui la sequenza di record rappresenta una sequenza ordinata di occorrenze ed eventi. Ogni record rappresenta un'istantanea di un momento specifico nel tempo. È tuttavia possibile che la maggior parte delle informazioni utili non sia contenuta nei valori del momento, bensì risieda nella variazione e nel comportamento di tali valori nel corso del tempo.

Il parametro rilevante può anche non essere rappresentato dal tempo. Gli stessi principi possono essere applicati anche a record che rappresentano, per esempio, analisi eseguite a distanze diverse lungo una linea.

Le funzioni speciali e di sequenza sono immediatamente riconoscibili grazie alle seguenti caratteristiche:

- Sono tutte precedute dal prefisso @.
- I nomi sono riportati in lettere maiuscole.

Le funzioni sequenza possono fare riferimento al record correntemente elaborato da un nodo, ai record che hanno già attraversato un nodo e anche, in un caso, ai record non ancora passati attraverso un nodo. Le funzioni sequenza possono essere combinate liberamente con altri componenti delle espressioni CLEM, benché alcune presentino restrizioni relative agli argomenti utilizzabili.

Esempi

Può rivelarsi utile conoscere da quanto tempo si è verificato un evento specifico o una condizione si è rivelata vera. A tale scopo, è possibile utilizzare la funzione @SINCE, per esempio:

```
@SINCE(Reddito > Risultati)
```

Questa funzione restituisce la distanza dell'ultimo record in cui questa condizione è stata rilevata come vera, ovvero il numero di record precedenti a quello considerato in cui la condizione era vera. Se la condizione non è mai stata rilevata come vera, @SINCE restituirà @INDEX + 1.

Talvolta può rivelarsi utile fare riferimento a un valore del record corrente nell'espressione utilizzata con @SINCE. A questo scopo, è possibile utilizzare la funzione @THIS, con cui si specifica che il nome di un campo deve essere sempre applicato al record corrente. Per trovare la distanza dell'ultimo record contenente un campo Concentration con valore due volte superiore rispetto a quello del record corrente, si può utilizzare quanto segue:

```
@SINCE(Concentrazione > 2 * @THIS(Concentrazione))
```

In alcuni casi, la condizione specificata per @SINCE è vera per il record corrente per definizione, per esempio:

```
@SINCE(ID == @THIS(ID))
```

Per questo motivo, @SINCE non valuta la condizione per il record corrente. Per valutare la condizione per il record corrente così come per i precedenti, utilizzare l'analoga funzione @SINCE0. Se la condizione è vera nel record corrente, @SINCE0 restituirà 0.

Nota: non è possibile chiamare le funzioni @ da script. Per ulteriori informazioni, vedere l'argomento [Espressioni CLEM negli script in il capitolo 3 in IBM SPSS Modeler 14.2 Guida per script e automazione.](#)

Funzione	Risultato	Descrizione
MEAN(FIELD)	Reale	Restituisce la media dei valori per l'argomento CAMPO o CAMPI specificato.
@MEAN(FIELD, EXPR)	Reale	Restituisce la media dei valori per CAMPO negli ultimi record ESPR ricevuti dal nodo corrente, incluso il record corrente. CAMPO deve essere il nome di un campo numerico. ESPR può essere qualsiasi valutazione di espressione per un numero intero maggiore di 0. Se ESPR non viene specificato o se il valore specificato è superiore al numero di record ricevuti, viene restituita la media dei valori inclusi in tutti i record ricevuti. Si noti che questa funzione non può essere chiamata da uno script. Per ulteriori informazioni, vedere l'argomento Espressioni CLEM negli script in il capitolo 3 in IBM SPSS Modeler 14.2 Guida per script e automazione.
@MEAN(FIELD, EXPR, INT)	Reale	Restituisce la media dei valori per CAMPO negli ultimi record ESPR ricevuti dal nodo corrente, incluso il record corrente. CAMPO deve essere il nome di un campo numerico. ESPR può essere qualsiasi valutazione di espressione per un numero intero maggiore di 0. Se ESPR non viene specificato o se il valore specificato è superiore al numero di record ricevuti, viene restituita la media dei valori inclusi in tutti i record ricevuti. INT indica il numero massimo di valori da cercare all'indietro. Se si specifica questo terzo argomento, si otterrà un'operazione più precisa.
@DIFF1(FIELD)	Reale	Restituisce il primo differenziale di CAMPO1. La forma con un singolo argomento restituisce semplicemente la differenza tra il valore corrente e il valore precedente del campo. Restituisce 0 se per il campo specificato non esiste alcun valore precedente.
@DIFF1(FIELD1, FIELD2)	Reale	La forma con due argomenti fornisce il primo differenziale di CAMPO1 rispetto a CAMPO2. Restituisce 0 se per il campo specificato non esiste alcun valore precedente.
@DIFF2(FIELD)	Reale	Restituisce il secondo differenziale di CAMPO1. La forma con un singolo argomento restituisce semplicemente la differenza tra il valore corrente e il valore precedente del campo. Restituisce 0 se per il campo specificato non esiste alcun valore precedente.
@DIFF2(FIELD1, FIELD2)	Reale	La forma con due argomenti fornisce il primo differenziale di CAMPO1 rispetto a CAMPO2. Restituisce 0 se per il campo specificato non esiste alcun valore precedente.
@INDEX	Intero	Restituisce l'indice del record corrente. Gli indici vengono assegnati ai record mano a mano che arrivano al nodo corrente. Al primo record viene assegnato l'indice 1, che viene incrementato di 1 per ciascun record successivo.

Funzione	Risultato	Descrizione
@LAST_NON_BLANK(FIELD)	Qualsiasi	Restituisce l'ultimo valore per <i>CAMPO</i> che non era vuoto, come definito in un nodo Input o Tipo a monte. Se i record letti non contengono valori non vuoti per <i>CAMPO</i> , viene restituito \$null\$. Si noti che è possibile definire separatamente per ogni campo valori vuoti, denominati anche valori mancanti definiti dall'utente.
@MAX(FIELD)	Number	Restituisce il valore massimo per il campo specificato con l'argomento <i>CAMPO</i> .
@MAX(FIELD, EXPR)	Number	Restituisce il valore massimo per <i>CAMPO</i> negli ultimi record <i>ESPR</i> ricevuti, incluso il record corrente. <i>CAMPO</i> deve essere il nome di un campo numerico. <i>ESPR</i> può essere qualsiasi valutazione di espressione per un numero intero maggiore di 0.
@MAX(FIELD, EXPR, INT)	Number	Restituisce il valore massimo per <i>CAMPO</i> negli ultimi record <i>ESPR</i> ricevuti, incluso il record corrente. <i>CAMPO</i> deve essere il nome di un campo numerico. <i>ESPR</i> può essere qualsiasi valutazione di espressione per un numero intero maggiore di 0. Se <i>ESPR</i> non viene specificato o se il valore specificato è superiore al numero di record ricevuti, viene restituito il valore massimo dei valori inclusi in tutti i record ricevuti. <i>INT</i> indica il numero massimo di valori da cercare all'indietro. Se si specifica questo terzo argomento, si otterrà un'operazione più precisa.
@MIN(FIELD)	Number	Restituisce il valore minimo per il campo specificato con l'argomento <i>CAMPO</i> .
@MIN(FIELD, EXPR)	Number	Restituisce il valore minimo per <i>CAMPO</i> negli ultimi record <i>ESPR</i> ricevuti, incluso il record corrente. <i>CAMPO</i> deve essere il nome di un campo numerico. <i>ESPR</i> può essere qualsiasi valutazione di espressione per un numero intero maggiore di 0.
@MIN(FIELD, EXPR, INT)	Number	Restituisce il valore minimo per <i>CAMPO</i> negli ultimi record <i>ESPR</i> ricevuti, incluso il record corrente. <i>CAMPO</i> deve essere il nome di un campo numerico. <i>ESPR</i> può essere qualsiasi valutazione di espressione per un numero intero maggiore di 0. Se <i>ESPR</i> non viene specificato o se il valore specificato è superiore al numero di record ricevuti, viene restituito il valore minimo dei valori inclusi in tutti i record ricevuti. <i>INT</i> indica il numero massimo di valori da cercare all'indietro. Se si specifica questo terzo argomento, si otterrà un'operazione più precisa.

Funzione	Risultato	Descrizione
@OFFSET(FIELD, EXPR)	Qualsiasi	Restituisce il valore di <i>CAMPO</i> nel record con una distanza dal record corrente pari al valore di <i>ESPR</i> . Una distanza positiva fa riferimento a un record già passato, mentre una distanza negativa specifica una ricerca in avanti di un record non ancora pervenuto. Per esempio, @OFFSET(Status, 1) restituisce il valore del campo Status nel record precedente mentre @OFFSET(Status, -4) viene eseguita la ricerca in avanti su quattro record nella sequenza, ovvero su record che non sono ancora passati attraverso questo nodo) per ottenere il valore. <i>Si noti che una distanza negativa (ricerca in avanti) deve essere specificata come costante.</i> Per le sole distanze positive, <i>ESPR</i> può inoltre rappresentare un'espressione CLEM arbitraria, che può essere valutata per fornire la distanza in relazione al record corrente. In tal caso, l'utilizzo di una versione a tre argomenti di questa funzione dovrebbe migliorare la performance (vedere la funzione successiva). Se questa espressione restituisce un valore diverso da un numero intero non negativo, verrà generato un errore. Non è quindi consentito il calcolo di distanze per la ricerca in avanti. <i>Nota:</i> Una funzione @OFFSET autoreferenziale non può utilizzare una "ricerca in avanti" letterale. Per esempio, in un nodo Riempimento non è possibile sostituire il valore di field1 utilizzando un'espressione come @OFFSET(field1,-2).
@OFFSET(FIELD, EXPR, INT)	Qualsiasi	Esegue la stessa operazione della funzione @OFFSET con l'aggiunta di un terzo argomento, <i>INT</i> , che specifica il numero massimo di valori da cercare all'indietro. In casi in cui la distanza viene calcolata da un'espressione, il terzo argomento dovrebbe migliorare le prestazioni. Per esempio in un'espressione come @OFFSET(Foo, Month, 12) il sistema conserva solo gli ultimi dodici valori di Foo, altrimenti deve archiviare ogni valore. Per i casi in cui il valore di distanza sia una costante, incluse le distanze di "ricerca in avanti" negative, che devono essere costanti, il terzo argomento è inutile ed è necessario usare la versione a due argomenti di questa funzione. Vedere anche la nota relativa alle funzioni autoreferenziali nella versione a due argomenti descritte sopra.
@SDEV(FIELD)	Reale	Restituisce la deviazione standard dei valori per il <i>CAMPO</i> o i <i>CAMPI</i> specificati.
@SDEV(FIELD, EXPR)	Reale	Restituisce la deviazione standard dei valori per <i>CAMPO</i> negli ultimi record <i>ESPR</i> ricevuti dal nodo corrente, incluso il record corrente. <i>CAMPO</i> deve essere il nome di un campo numerico. <i>ESPR</i> può essere qualsiasi valutazione di espressione per un numero intero maggiore di 0. Se <i>ESPR</i> non viene specificato o se il valore specificato è superiore al numero di record ricevuti, viene restituita la deviazione standard in tutti i record ricevuti.

Funzione	Risultato	Descrizione
@SDEV(FIELD, EXPR, INT)	Reale	Restituisce la deviazione standard dei valori per <i>CAMPO</i> negli ultimi record <i>ESPR</i> ricevuti dal nodo corrente, incluso il record corrente. <i>CAMPO</i> deve essere il nome di un campo numerico. <i>ESPR</i> può essere qualsiasi valutazione di espressione per un numero intero maggiore di 0. Se <i>ESPR</i> non viene specificato o se il valore specificato è superiore al numero di record ricevuti, viene restituita la deviazione standard in tutti i record ricevuti. <i>INT</i> indica il numero massimo di valori da cercare all'indietro. Se si specifica questo terzo argomento, si otterrà un'operazione più precisa.
@SINCE(EXPR)	Qualsiasi	Restituisce il numero di record passati da quando l'espressione CLEM arbitraria rappresentata da <i>ESPR</i> è stata rilevata come vera.
@SINCE(EXPR, INT)	Qualsiasi	L'aggiunta del secondo argomento, <i>INT</i> , consente di specificare il numero massimo di record da cercare all'indietro. Se <i>ESPR</i> non è mai stata vera <i>INT</i> corrisponderà a @INDEX+1.
@SINCE0(EXPR)	Qualsiasi	A differenza di @SINCE, considera il record corrente. @SINCE0 restituisce 0 se <i>ESPR</i> risulta vera per il record corrente.
@SINCE0(EXPR, INT)	Qualsiasi	L'aggiunta del secondo argomento, <i>INT</i> , consente di specificare il numero massimo di record da cercare all'indietro.
@SUM(FIELD)	Number	Restituisce la somma dei valori per l'argomento <i>CAMPO</i> o <i>CAMPI</i> specificato.
@SUM(FIELD, EXPR)	Number	Restituisce la somma dei valori per <i>CAMPO</i> negli ultimi record <i>ESPR</i> ricevuti dal nodo corrente, incluso il record corrente. <i>CAMPO</i> deve essere il nome di un campo numerico. <i>ESPR</i> può essere qualsiasi valutazione di espressione per un numero intero maggiore di 0. Se <i>ESPR</i> non viene specificato o se il valore specificato è superiore al numero di record ricevuti, viene restituita la somma dei valori inclusi in tutti i record ricevuti.
@SUM(FIELD, EXPR, INT)	Number	Restituisce la somma dei valori per <i>CAMPO</i> negli ultimi record <i>ESPR</i> ricevuti dal nodo corrente, incluso il record corrente. <i>CAMPO</i> deve essere il nome di un campo numerico. <i>ESPR</i> può essere qualsiasi valutazione di espressione per un numero intero maggiore di 0. Se <i>ESPR</i> non viene specificato o se il valore specificato è superiore al numero di record ricevuti, viene restituita la somma dei valori inclusi in tutti i record ricevuti. <i>INT</i> indica il numero massimo di valori da cercare all'indietro. Se si specifica questo terzo argomento, si otterrà un'operazione più precisa.
@THIS(FIELD)	Qualsiasi	Restituisce il valore del campo denominato <i>CAMPO</i> nel record corrente. Viene utilizzata solo nelle espressioni @SINCE.

Funzioni globali

Le funzioni @MEAN,@SUM, @MIN, @MAX e @SDEV possono essere utilizzate su tutti i record letti fino al record corrente incluso. In alcuni casi, tuttavia, è utile poter confrontare i valori nel record corrente con i valori rilevati nell'intero insieme di dati. Se si utilizza un nodo Globali per generare valori nell'intero insieme di dati, è possibile accedere a tali valori in un'espressione CLEM mediante le funzioni globali.

Ad esempio,

@GLOBAL_MAX(Età)

restituisce il valore più elevato di Age nell'insieme di dati, mentre l'espressione

(Valore - @GLOBAL_MEAN(Valore)) / @GLOBAL_SDEV(Valore)

esprime la differenza tra il valore Value di questo record e la media globale come numero di deviazioni standard. È possibile utilizzare i valori globali solo dopo che sono stati calcolati in un nodo Globali. Tutti i valori globali correnti possono essere annullati facendo clic sul pulsante Cancella valori globali nella scheda Valori globali della finestra di dialogo delle proprietà dello stream.

Nota: non è possibile chiamare le funzioni @ da script. [Per ulteriori informazioni, vedere l'argomento Espressioni CLEM negli script in il capitolo 3 in IBM SPSS Modeler 14.2 Guida per script e automazione.](#)

Funzione	Risultato	Descrizione
@GLOBAL_MAX(FIELD)	Number	Restituisce il valore massimo per CAMPO nell'intero insieme di dati, generato in precedenza da un nodo Globali. CAMPO deve essere il nome di un campo numerico. Se il valore globale corrispondente non è stato impostato, si verifica un errore. Si noti che questa funzione non può essere chiamata da uno script. Per ulteriori informazioni, vedere l'argomento Espressioni CLEM negli script in il capitolo 3 in IBM SPSS Modeler 14.2 Guida per script e automazione.
@GLOBAL_MIN(FIELD)	Number	Restituisce il valore minimo per CAMPO nell'intero insieme di dati, generato in precedenza da un nodo Globali. CAMPO deve essere il nome di un campo numerico. Se il valore globale corrispondente non è stato impostato, si verifica un errore.
@GLOBAL_SDEV(FIELD)	Number	Restituisce la deviazione standard dei valori per CAMPO nell'intero insieme di dati, generato in precedenza da un nodo Globali. CAMPO deve essere il nome di un campo numerico. Se il valore globale corrispondente non è stato impostato, si verifica un errore.

Funzione	Risultato	Descrizione
@GLOBAL_MEAN(FIELD)	Number	Restituisce la media dei valori per <i>CAMPO</i> nell'intero insieme di dati, generato in precedenza da un nodo Globali. <i>CAMPO</i> deve essere il nome di un campo numerico. Se il valore globale corrispondente non è stato impostato, si verifica un errore.
@GLOBAL_SUM(FIELD)	Number	Restituisce la somma dei valori per <i>CAMPO</i> nell'intero insieme di dati, generato in precedenza da un nodo Globali. <i>CAMPO</i> deve essere il nome di un campo numerico. Se il valore globale corrispondente non è stato impostato, si verifica un errore.

Funzioni per la gestione di valori vuoti e nulli

Mediante il linguaggio CLEM è possibile stabilire che valori specifici in un campo devono essere considerati come valori vuoti o mancanti. Le funzioni illustrate di seguito gestiscono i valori vuoti.

Nota: non è possibile chiamare le funzioni @ da script. [Per ulteriori informazioni, vedere l'argomento Espressioni CLEM negli script in il capitolo 3 in IBM SPSS Modeler 14.2 Guida per script e automazione.](#)

Funzione	Risultato	Descrizione
@BLANK(FIELD)	Booleana	Restituisce un valore vero per tutti i record contenenti valori vuoti in base alle regole di gestione dei vuoti impostate in un nodo Input o Tipo a monte mediante la scheda Tipi. Si noti che questa funzione non può essere chiamata da uno script. Per ulteriori informazioni, vedere l'argomento Espressioni CLEM negli script in il capitolo 3 in IBM SPSS Modeler 14.2 Guida per script e automazione.
@LAST_NON_BLANK(FIELD)	Qualsiasi	Restituisce l'ultimo valore per <i>CAMPO</i> che non era vuoto, come definito in un nodo Input o Tipo a monte. Se i record letti non contengono valori non vuoti per <i>CAMPO</i> , viene restituito \$null\$. Si noti che è possibile definire separatamente per ogni campo valori vuoti, denominati anche valori mancanti definiti dall'utente.
@NULL(FIELD)	Booleana	Restituisce vero se il valore di <i>CAMPO</i> è il valore mancante di sistema \$null\$. Restituisce falso per tutti gli altri valori, inclusi i valori vuoti definiti dall'utente. Per ricercare entrambi, utilizzare @BLANK(FIELD) e @NULL(FIELD).
undef	Qualsiasi	Viene utilizzata in genere in CLEM per immettere un valore \$null\$, per esempio per sostituire i valori vuoti con valori nulli nel nodo Riempimento.

I campi vuoti possono essere "riempiti" con il nodo Riempimento. Sia nei nodi Riempimento che Nuovo campo, solo in modalità multipla, la funzione speciale CLEM@FIELD fa riferimento ai campi correnti esaminati.

Campi speciali

Le funzioni speciali vengono utilizzate per indicare i campi specifici esaminati oppure per generare un elenco di campi come input. In caso di derivazione di più campi, per esempio, è opportuno utilizzare @FIELD per indicare che l'operazione di derivazione deve essere eseguita sui campi selezionati. Se si utilizza l'espressione $\log(@FIELD)$, viene derivato un nuovo campo di logaritmo per ogni campo selezionato.

Nota: non è possibile chiamare le funzioni @ da script. [Per ulteriori informazioni, vedere l'argomento Espressioni CLEM negli script in il capitolo 3 in IBM SPSS Modeler 14.2 Guida per script e automazione.](#)

Funzione	Risultato	Descrizione
@FIELD	<i>Qualsiasi</i>	Esegue un'operazione su tutti i campi specificati nel contesto dell'espressione. Si noti che questa funzione non può essere chiamata da uno script. Per ulteriori informazioni, vedere l'argomento Espressioni CLEM negli script in il capitolo 3 in IBM SPSS Modeler 14.2 Guida per script e automazione.
@TARGET	<i>Qualsiasi</i>	In caso di utilizzo di un'espressione CLEM in una funzione di analisi definita dall'utente, @TARGET rappresenta il campo obiettivo o il "valore corretto" per la coppia obiettivo/previsto analizzata. Questa funzione viene in genere utilizzata in un nodo Analisi.
@PREDICTED	<i>Qualsiasi</i>	In caso di utilizzo di un'espressione CLEM in una funzione di analisi definita dall'utente, @PREDICTED rappresenta il campo previsto per la coppia obiettivo/previsto analizzata. Questa funzione viene in genere utilizzata in un nodo Analisi.
@PARTITION_FIELD	<i>Qualsiasi</i>	Sostituisce il nome del campo di partizione corrente.
@TRAINING_PARTITION	<i>Qualsiasi</i>	Restituisce il valore della partizione dell'addestramento corrente. Per esempio, per selezionare i record di addestramento tramite un nodo Seleziona, usare l'espressione CLEM: @PARTITION_FIELD = @TRAINING_PARTITION. In tal modo, il nodo Seleziona funzionerà indipendentemente dai valori utilizzati per rappresentare ogni partizione nei dati.
@TESTING_PARTITION	<i>Qualsiasi</i>	Restituisce il valore della partizione di test corrente.
@VALIDATION_PARTITION	<i>Qualsiasi</i>	Restituisce il valore della partizione di validazione corrente.
@FIELDS_BETWEEN(start, end)	<i>Qualsiasi</i>	Restituisce l'elenco di nomi di campo compresi tra i campi iniziale e finale (inclusi) specificati in base all'ordine naturale (cioè di inserimento) dei campi nei dati. Per ulteriori informazioni, vedere l'argomento Riepilogo di più campi in il capitolo 7 a pag. 108.

Funzione	Risultato	Descrizione
@FIELDS_MATCHING(pattern)	<i>Qualsiasi</i>	Restituisce un elenco di nomi di campo corrispondenti a uno schema specificato. È possibile includere nello schema un punto interrogativo (?) corrispondente esattamente a un carattere; un asterisco (*) corrisponde a zero o più caratteri. Per indicare un punto interrogativo o un asterisco letterale, ovvero non utilizzato come carattere jolly, è possibile utilizzare la barra rovesciata (\). Per ulteriori informazioni, vedere l'argomento Riepilogo di più campi in il capitolo 7 a pag. 108.
@MULTI_RESPONSE_SET	<i>Qualsiasi</i>	Restituisce l'elenco dei campi nell'insieme di risposte multiple denominato. Per ulteriori informazioni, vedere l'argomento Utilizzo di dati a risposta multipla in il capitolo 7 a pag. 109.

Utilizzo di IBM SPSS Modeler con un repository

Informazioni su IBM SPSS Collaboration and Deployment Services Repository

IBM® SPSS® Modeler può essere utilizzato insieme a un repository IBM SPSS Collaboration and Deployment Services, il che consente di gestire il ciclo di vita dei modelli di data mining e degli oggetti predittivi correlati, oltre a consentire l'utilizzo di questi oggetti dalle applicazioni, dagli strumenti e dalle soluzioni aziendali. Gli oggetti SPSS Modeler che possono essere condivisi in questo modo includono gli stream, i nodi, gli output degli stream, gli scenari, i progetti e i modelli. Gli oggetti vengono memorizzati nel repository centralizzato, da dove è possibile condividerli con altre applicazioni e registrarli utilizzando funzionalità avanzate per il controllo delle versioni, dei metadati e della ricerca.

Nota: per l'accesso a un repository IBM® SPSS® Collaboration and Deployment Services è necessaria una licenza distinta. Per ulteriori informazioni, vedere <http://www.ibm.com/software/analytics/spss/products/deployment/cds/>

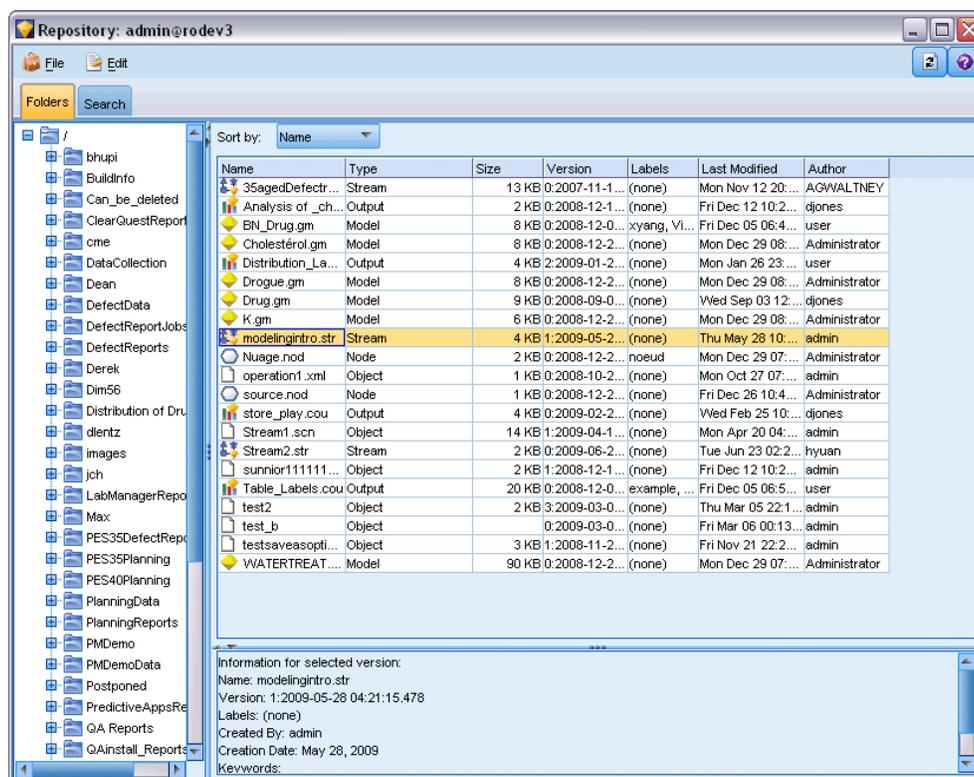
Prima di poter utilizzare SPSS Modeler con il repository, è necessario installare un adattatore sull'host del repository. Senza l'adattatore, quando si tenta di accedere agli oggetti del repository da alcuni nodi o modelli di SPSS Modeler è possibile che venga visualizzato il seguente messaggio:

Potrebbe essere necessario aggiornare il repository in modo che supporti nuovi tipi di nodi, modelli e output.

Per istruzioni sull'installazione dell'adattatore, vedere la guida *SPSS Modeler Deployment Installation*, disponibile in SPSS Modeler Deployment DVD. Per informazioni su come accedere agli oggetti del repository SPSS Modeler da IBM® SPSS® Collaboration and Deployment Services Deployment Manager, consultare il documento *SPSS Modeler Deployment Guide*.

Le sezioni che seguono contengono informazioni su come accedere al repository da SPSS Modeler.

Figura 9-1
 Oggetti in IBM SPSS Collaboration and Deployment Services Repository



Supporto completo del controllo delle versioni e della ricerca

Il repository offre funzionalità complete di controllo delle versioni e di ricerca. Per esempio, si supponga di creare uno stream e di memorizzarlo nel repository, in modo da poterlo condividere con i ricercatori di altre divisioni. Se in seguito si aggiorna lo stream in SPSS Modeler, è possibile aggiungere la versione aggiornata al repository senza sovrascrivere la versione precedente. Tutte le versioni rimangono accessibili ed è possibile eseguire ricerche per nome, etichetta, campi utilizzati o altri attributi. Per esempio, è possibile cercare tutte le versioni del modello che utilizzano le entrate nette come input o tutti i modelli creati da un utente specifico. Per eseguire queste operazioni in un file system tradizionale, sarebbe necessario salvare ogni versione con un nome file diverso e il programma non sarebbe a conoscenza delle relazioni tra le diverse versioni.

Single Sign-On

La funzione Single Sign-On consente agli utenti di collegarsi a IBM® SPSS® Collaboration and Deployment Services Repository senza dover immettere ogni volta nome utente e password. I dettagli di accesso alla rete locale dell'utente forniscono l'autenticazione necessaria per accedere

a IBM SPSS Collaboration and Deployment Services. Questa funzione dipende dalle seguenti condizioni:

- IBM SPSS Collaboration and Deployment Services deve essere configurato in modo da utilizzare un provider di Single Sign-On
- L'utente deve avere eseguito l'accesso a un host compatibile con il provider

Per ulteriori informazioni, vedere l'argomento [Connessione a IBM SPSS Collaboration and Deployment Services Repository](#) a pag. 159.

Archiviazione e deployment di oggetti di IBM SPSS Collaboration and Deployment Services Repository

Gli **stream** creati in IBM® SPSS® Modeler possono essere archiviati nel IBM® SPSS® Collaboration and Deployment Services Repository così come sono, come file con estensione *.str*. In questo modo, più utenti in tutta l'azienda possono accedere a un unico stream. [Per ulteriori informazioni, vedere l'argomento Archiviazione di oggetti in IBM SPSS Collaboration and Deployment Services Repository](#) a pag. 163.

È possibile anche eseguire il deployment di uno stream nel IBM SPSS Collaboration and Deployment Services Repository. Uno **stream implementato** viene archiviato come file con estensione *.str*, ma con metadati aggiuntivi per consentirne l'utilizzo da parte dell'applicazione thin client IBM® SPSS® Modeler Advantage. [Per ulteriori informazioni, vedere l'argomento Apertura di uno stream in IBM SPSS Modeler Advantage](#) in il capitolo 10 a pag. 197.

Tuttavia, per trarre il massimo vantaggio dalle funzioni aziendali di IBM® SPSS® Collaboration and Deployment Services, è necessario eseguire il deployment dello stream come **scenario**. Si tratta di un file (con estensione *.scn*) che contiene uno stream SPSS Modeler con specifici nodi, modelli e proprietà aggiuntive che ne consentono l'uso all'interno del IBM SPSS Collaboration and Deployment Services Repository per scopi quali il calcolo automatico del punteggio e l'aggiornamento dei modelli. Per esempio, un modello di Autoapprendimento (SLRM) può essere aggiornato automaticamente a intervalli regolari man mano che sono disponibili nuovi dati. In alternativa, è possibile eseguire il deployment di un insieme di stream per l'analisi Campione/Sfidante, in cui si confrontano gli scenari per determinare lo scenario che contiene il modello predittivo più efficace.

Per ulteriori informazioni sull'archiviazione degli stream come scenari, vedere [Opzioni di deployment degli stream](#) a pag. 186.

Requisiti per gli stream implementati come scenari

- Per garantire un accesso coerente ai dati aziendali, è necessario accedere agli stream di cui è stato eseguito il deployment tramite il componente IBM® SPSS® Collaboration and Deployment Services Enterprise View di IBM SPSS Collaboration and Deployment Services. Ciò significa che in SPSS Modeler deve esserci almeno un nodo di input Enterprise View all'interno di ogni ramo designato dello stream, sia di calcolo del punteggio che di creazione del modello. [Per ulteriori informazioni, vedere l'argomento Nodo Enterprise View](#) in il capitolo 2 in *IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output*.

- Per utilizzare il nodo Enterprise View, IBM SPSS Collaboration and Deployment Services deve essere installato, configurato e accessibile dal proprio sito, con una Enterprise View, le Visualizzazioni applicazione e le DPD già definite. Per ulteriori informazioni, rivolgersi all'amministratore di sistema oppure visitare il sito Web aziendale all'indirizzo <http://www.ibm.com/software/analytics/spss/products/deployment/cds/>.
- Una Definizione dei provider di dati (DPD) viene definita rispetto a una determinata sorgente dati ODBC. Per utilizzare una DPD da SPSS Modeler, è necessario che sull'host server di SPSS Modeler sia stata definita una sorgente dati ODBC con lo stesso nome e che si connette allo stesso archivio dati di quella a cui fa riferimento la Definizione dei provider di dati.
- È inoltre necessario che IBM® SPSS® Collaboration and Deployment Services Enterprise View Driver sia installato su ogni computer utilizzato per modificare o eseguire lo stream. Per Windows, installare semplicemente il driver sul computer su cui è installato IBM® SPSS® Modeler o IBM® SPSS® Modeler Server. Non sono necessarie ulteriori configurazioni del driver. In UNIX, è necessario aggiungere allo script di avvio un riferimento allo script *pev.sh*. Per ulteriori informazioni, vedere l'argomento [Configurazione di un driver per il nodo Enterprise View in l'appendice B in IBM SPSS Modeler Server 14.2 Guida della performance e amministrazione](#). Per dettagli sull'installazione di IBM SPSS Collaboration and Deployment Services Enterprise View Driver, rivolgersi all'amministratore di sistema.

Altre opzioni di deployment

Mentre IBM SPSS Collaboration and Deployment Services offre le più estese e avanzate funzionalità di gestione del contenuto aziendale, sono disponibili altri meccanismi per il deployment o l'esportazione di stream, tra cui:

- Usare Creazione guidata applicazioni predittive 4.x per esportare stream per il deployment in tale versione delle applicazioni predittive. Per ulteriori informazioni, vedere l'argomento [Creazione guidata applicazioni predittive 4.x in il capitolo 10 a pag. 198](#).
- Esportare lo stream e il modello da utilizzare successivamente con IBM® SPSS® Modeler Solution Publisher Runtime. Per ulteriori informazioni, vedere l'argomento [IBM SPSS Modeler Solution Publisher in il capitolo 2 in IBM SPSS Modeler 14.2 Solution Publisher](#).
- Esportare uno o più modelli in PMML, un formato basato su XML per la codifica delle informazioni sui modelli. Per ulteriori informazioni, vedere l'argomento [Importazione ed esportazione di modelli come PMML in il capitolo 10 a pag. 206](#).

Connessione a IBM SPSS Collaboration and Deployment Services Repository

- ▶ Per connettersi al IBM® SPSS® Collaboration and Deployment Services Repository, nel menu principale di IBM® SPSS® Modeler fare clic su:
Strumenti > Repository > Opzioni...
- ▶ Specificare le opzioni di login richieste.

Le impostazioni sono specifiche di ogni sito o installazione. Per informazioni specifiche sulla porta e l'accesso, rivolgersi all'amministratore del sistema locale.

Nota: per l'accesso a un repository IBM® SPSS® Collaboration and Deployment Services è necessaria una licenza distinta. Per ulteriori informazioni, vedere <http://www.ibm.com/software/analytics/spss/products/deployment/cds/>

Figura 9-2
Accesso a IBM SPSS Collaboration and Deployment Services Repository



Repository. L'installazione di IBM SPSS Collaboration and Deployment Services Repository a cui si desidera accedere. In genere corrisponde al nome del server host sul quale è installato il repository. È possibile connettersi a un solo archivio alla volta.

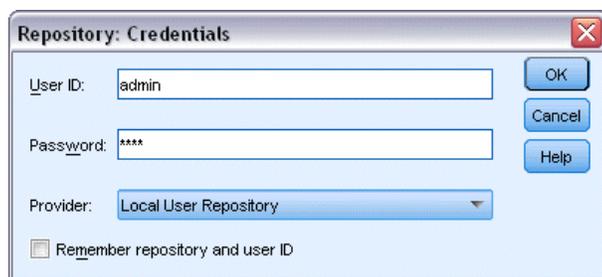
Porta: La porta utilizzata per l'hosting della connessione, in genere 8080 per default.

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

Abilita connessione sicura. 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 Collaboration and Deployment Services Repository. Se sono necessarie ulteriori informazioni, rivolgersi all'amministratore di sistema.

Immissione di credenziali per il IBM SPSS Collaboration and Deployment Services Repository

Figura 9-3
Immissione delle credenziali di IBM SPSS Collaboration and Deployment Services Repository



ID utente e Password. Specificare un nome utente e una password validi per il login. Se necessario, rivolgersi all'amministratore locale per ulteriori informazioni.

Provider. Scegliere un provider di protezione per l'autenticazione. Il IBM® SPSS® Collaboration and Deployment Services Repository può essere configurato in modo da utilizzare diversi provider di protezione. Se sono necessarie ulteriori informazioni, rivolgersi all'amministratore di sistema.

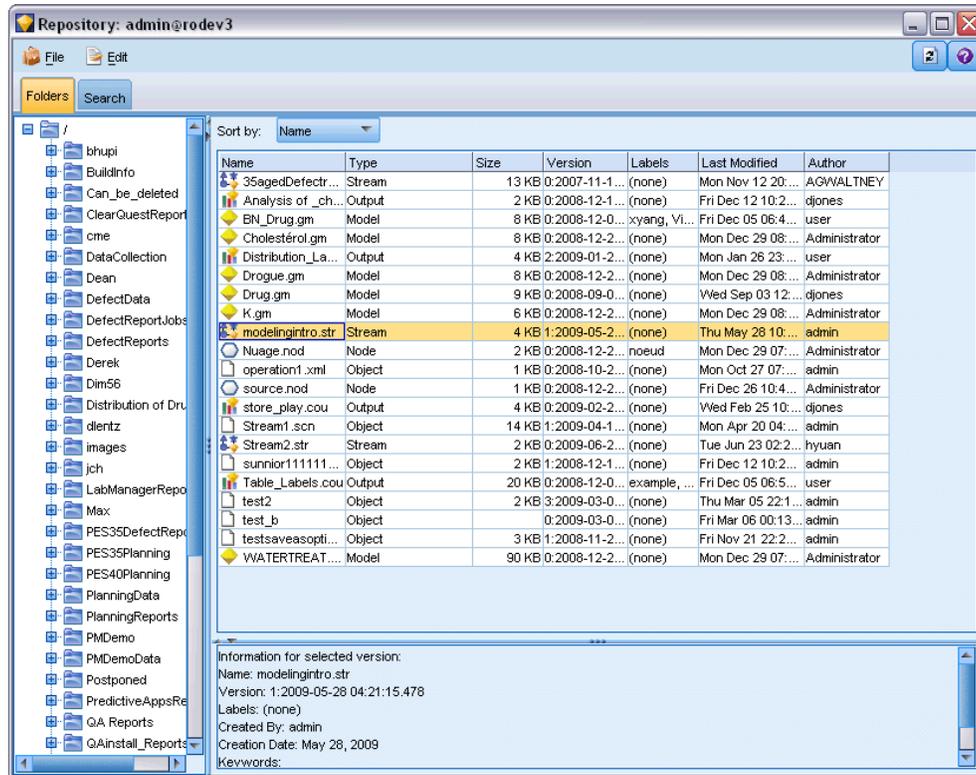
Memorizza archivio e ID utente. Salva le impostazioni correnti come di default, in modo tale che non sia necessario reinsertarle ogni volta che si desidera eseguire la connessione.

Esplorazione dei contenuti di IBM SPSS Collaboration and Deployment Services Repository

Il IBM® SPSS® Collaboration and Deployment Services Repository consente di esplorare i contenuti archiviati in modo simile a Esplora risorse di Windows; è possibile anche cercare le *versioni* di ogni oggetto archiviato.

- ▶ Per aprire la finestra del IBM SPSS Collaboration and Deployment Services Repository, nei menu di IBM® SPSS® Modeler fare clic su:
Strumenti > Repository > Esplora...
- ▶ Specificare le impostazioni relative alla connessione al repository, se necessario. [Per ulteriori informazioni, vedere l'argomento Connessione a IBM SPSS Collaboration and Deployment Services Repository a pag. 159.](#) Per informazioni specifiche sulla porta, la password e altri dettagli per la connessione, rivolgersi all'amministratore del sistema locale.

Figura 9-4
Esplorazione dei contenuti di IBM SPSS Collaboration and Deployment Services Repository

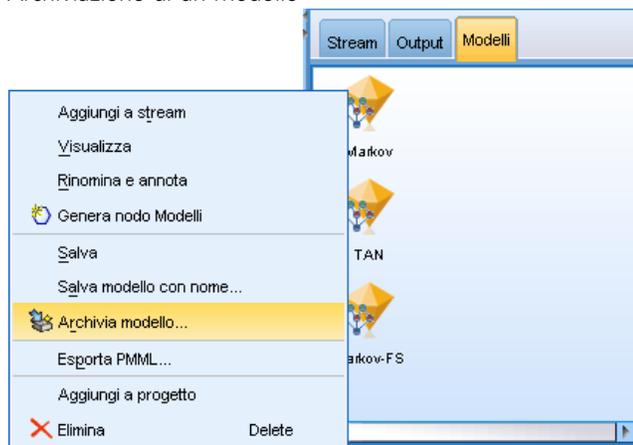


La finestra di esplorazione visualizza inizialmente una struttura ad albero della gerarchia delle cartelle. Fare clic sul nome di una cartella per visualizzarne il contenuto.

Nel riquadro destro vengono visualizzati gli oggetti corrispondenti al criterio di selezione o di ricerca corrente, nonché informazioni dettagliate sulla versione selezionata visualizzata nel riquadro in basso a destra. Gli attributi visualizzati sono validi per la versione più recente.

Archiviazione di oggetti in IBM SPSS Collaboration and Deployment Services Repository

Figura 9-5
Archiviazione di un modello



Stream, nodi, modelli, palette di modelli, progetti e oggetti di output possono essere archiviati nel repository, da dove altri utenti e applicazioni potranno accedervi.

Nota: per l'accesso a un repository IBM® SPSS® Collaboration and Deployment Services è necessaria una licenza distinta. Per ulteriori informazioni, vedere <http://www.ibm.com/software/analytics/spss/products/deployment/cds/>

Nel IBM® SPSS® Collaboration and Deployment Services Repository è anche possibile pubblicare output degli stream in un formato che ne consenta la visualizzazione ad altri utenti via Internet mediante IBM® SPSS® Collaboration and Deployment Services Deployment Portal. Per ulteriori informazioni, vedere l'argomento [Pubblica sul Web](#) in il capitolo 6 in *IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output*.

Impostazione delle proprietà degli oggetti

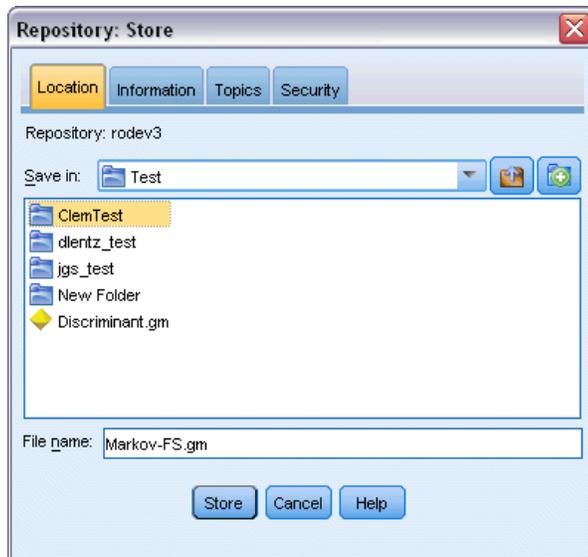
Quando si archivia un oggetto, il repository: Visualizza la finestra di dialogo Archivia che consente di specificare i valori di diverse proprietà dell'oggetto. È possibile:

- Scegliere il nome e la cartella del repository in cui archiviare l'oggetto
- Aggiungere informazioni sull'oggetto, per esempio l'etichetta versione e altre proprietà consultabili tramite ricerche
- Assegnare all'oggetto uno o più argomenti di classificazione
- Impostare le opzioni di sicurezza per l'oggetto

Le sezioni seguenti descrivono le proprietà che è possibile impostare.

Scelta della posizione di archiviazione degli oggetti

Figura 9-6
Scelta della posizione di archiviazione di un oggetto



Salva in. Mostra la cartella corrente, ovvero la posizione in cui verrà archiviato l'oggetto. Fare doppio clic sul nome di una cartella nell'elenco per impostare quella cartella come la cartella corrente. Utilizzare il pulsante Cartella superiore per navigare fino alla cartella principale. Utilizzare il pulsante Nuova cartella per creare una cartella al livello corrente.

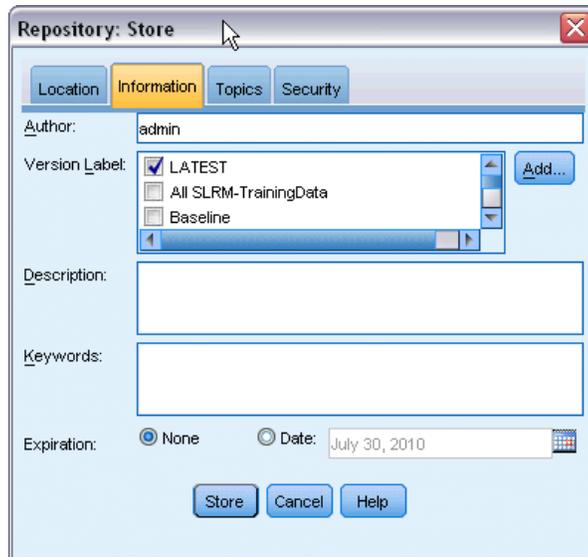
Nome file. Il nome con cui verrà archiviato l'oggetto.

Archivia. Archivia l'oggetto nella posizione corrente.

Aggiunta di informazioni sugli oggetti archiviati

Tutti i campi di questa scheda sono facoltativi.

Figura 9-7
Aggiunta di informazioni sull'oggetto



Autore. Il nome utente di chi ha creato l'oggetto nel repository. Per default, mostra il nome utente utilizzato per la connessione al repository, ma è possibile cambiare il nome qui.

Etichetta della versione. Selezionare un'etichetta dall'elenco per indicare la versione dell'oggetto, oppure fare clic su Aggiungi per creare una nuova etichetta. Evitare di usare il carattere “[” nell'etichetta. Se non si desidera assegnare un'etichetta a questa versione dell'oggetto, non selezionare alcuna casella. [Per ulteriori informazioni, vedere l'argomento Visualizzazione e modifica delle proprietà degli oggetti a pag. 181.](#)

Descrizione. Una descrizione dell'oggetto. Gli utenti possono effettuare ricerche di oggetti in base alla descrizione (vedere la nota).

Parole chiave. Una o più parole chiave che sono correlate all'oggetto e che possono essere utilizzate a scopo di ricerca (vedere la nota).

Scadenza. La data trascorsa la quale l'oggetto non è più visibile agli utenti generici, ma solo al proprietario e all'amministratore del repository. Per impostare una data di scadenza, selezionare l'opzione Data e inserire una data, oppure sceglierne una utilizzando il pulsante del calendario.

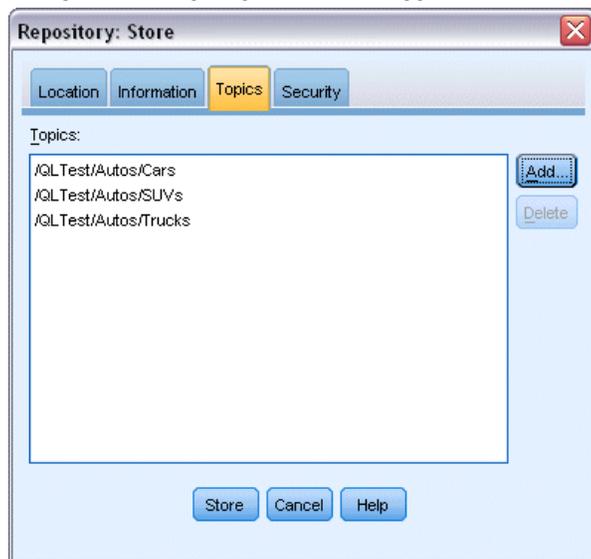
Archivia. Archivia l'oggetto nella posizione corrente.

Nota: le informazioni contenute nei campi Descrizione e Parole chiave sono considerate separatamente rispetto ai dati inseriti in IBM® SPSS® Modeler nella scheda Annotazioni dell'oggetto. Una ricerca per descrizione o parola chiave nel repository non restituisce le informazioni contenute nella scheda Annotazioni. [Per ulteriori informazioni, vedere l'argomento Ricerca di oggetti in IBM SPSS Collaboration and Deployment Services Repository a pag. 174.](#)

Assegnazione di argomenti a un oggetto archiviato

Gli argomenti sono un sistema di classificazione gerarchico per il contenuto archiviato nel repository. È possibile scegliere tra gli argomenti disponibili durante l'archiviazione degli oggetti e gli utenti possono inoltre cercare gli oggetti per argomento. L'elenco degli argomenti disponibili è impostato dagli utenti del repository che dispongono dei privilegi appropriati (per maggiori informazioni, vedere il manuale *Deployment Manager User's Guide*).

Figura 9-8
Assegnazione degli argomenti a un oggetto



Per assegnare un argomento all'oggetto:

- ▶ Fare clic sul pulsante Aggiungi.
- ▶ Fare clic sul nome di un argomento dall'elenco degli argomenti disponibili.
- ▶ Fare clic su OK.

Per rimuovere un'assegnazione di argomento:

- ▶ Selezionare l'argomento nell'elenco degli argomenti assegnati.
- ▶ Fare clic su Elimina.

Impostazione delle opzioni di protezione per gli oggetti archiviati

È possibile impostare o modificare una serie di opzioni di protezione per un oggetto archiviato. Per uno o più **principali** (ovvero, utenti o gruppi di utenti), è possibile:

- Assegnare i diritti di accesso all'oggetto
- Modificare i diritti di accesso all'oggetto
- Rimuovere i diritti di accesso all'oggetto

Figura 9-9
Impostazione delle opzioni di protezione per un oggetto



Principale. Il nome utente repository dell'utente o del gruppo che ha diritti di accesso su quell'oggetto.

Autorizzazioni. I diritti di accesso di cui questo utente o gruppo dispone per l'oggetto.

Aggiungi. Consente di aggiungere uno o più utenti o gruppi all'elenco di quelli che dispongono di diritti di accesso su questo oggetto. [Per ulteriori informazioni, vedere l'argomento Aggiunta di un utente all'elenco delle autorizzazioni a pag. 168.](#)

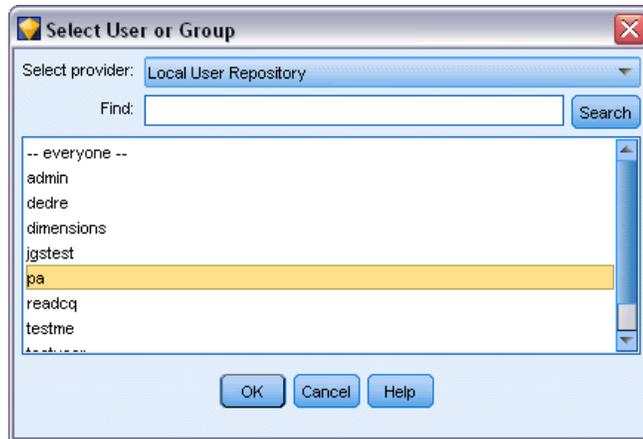
Modifica. Consente di modificare i diritti di accesso dell'utente o gruppo selezionato per questo oggetto. L'accesso in lettura è garantito per default. Questa opzione consente di garantire ulteriori diritti di accesso, nello specifico i diritti Proprietario, Scrittura, Eliminazione e Modifica autorizzazioni.

Elimina. Elimina l'utente o il gruppo selezionato dall'elenco delle autorizzazioni per questo oggetto.

Aggiunta di un utente all'elenco delle autorizzazioni

Figura 9-10

Aggiunta di un utente all'elenco delle autorizzazioni per un oggetto



Seleziona fornitore. Scegliere un provider di protezione per l'autenticazione. Il IBM® SPSS® Collaboration and Deployment Services Repository può essere configurato in modo da utilizzare diversi provider di protezione. Se sono necessarie ulteriori informazioni, rivolgersi all'amministratore di sistema.

Trova. Inserire il nome utente repository dell'utente o del gruppo che si desidera aggiungere, quindi fare clic su Cerca per visualizzare quel nome nell'elenco degli utenti. Per aggiungere più di un nome utente alla volta, lasciare vuoto questo campo e fare semplicemente clic su Cerca per visualizzare un elenco di tutti i nomi utente repository.

Elenco utenti. Selezionare uno o più nomi utente dall'elenco e fare clic su OK per aggiungerli all'elenco delle autorizzazioni.

Modifica dei diritti di accesso per un oggetto

Figura 9-11

Modifica dei diritti di accesso per un oggetto



Proprietario. Selezionare questa opzione per assegnare a questo utente o gruppo i diritti di accesso all'oggetto. Il proprietario ha il pieno controllo sull'oggetto, inclusi i diritti di accesso Eliminazione e Modifica.

Lettura. Per default, un utente o un gruppo che non è il proprietario dell'oggetto dispone solo dei diritti di accesso Lettura sull'oggetto. Selezionare le caselle di controllo appropriate per aggiungere i diritti di accesso Scrittura, Eliminazione e Modifica autorizzazioni per questo utente o gruppo.

Archiviazione di stream

È possibile archiviare uno stream come file *.str* nel repository, da dove altri utenti possono accedervi.

Nota: per informazioni sul deployment di uno stream per sfruttare appieno altre funzioni del repository, vedere Deployment di stream a pag. 184.

Per archiviare lo stream corrente:

- ▶ Nel menu principale, fare clic su:
File > Archivia > Memorizza come stream...
- ▶ Specificare le impostazioni relative alla connessione al repository, se necessario. [Per ulteriori informazioni, vedere l'argomento Connessione a IBM SPSS Collaboration and Deployment Services Repository a pag. 159.](#) Per informazioni specifiche sulla porta, la password e altri dettagli per la connessione, rivolgersi all'amministratore del sistema locale.
- ▶ Nel repository: nella finestra di dialogo Archivia, scegliere la cartella in cui archiviare l'oggetto, specificare le eventuali altre informazioni da registrare e fare clic sul pulsante Archivia. [Per ulteriori informazioni, vedere l'argomento Impostazione delle proprietà degli oggetti a pag. 163.](#)

Archiviazione di progetti

È possibile archiviare un progetto IBM® SPSS® Modeler completo come file *.cpj* nel repository, in modo che altri utenti possano accedervi.

Poiché un file di progetto è un contenitore per altri oggetti SPSS Modeler, è necessario dire a SPSS Modeler di archiviare gli oggetti del progetto nel IBM® SPSS® Collaboration and Deployment Services Repository. Questo parametro può essere impostato nella finestra di dialogo Proprietà progetto. [Per ulteriori informazioni, vedere l'argomento Impostazione delle proprietà del progetto in il capitolo 11 a pag. 216.](#)

Ogni volta che si aggiunge un nuovo oggetto a un progetto, SPSS Modeler chiede automaticamente di archivarlo se il progetto è configurato per l'archiviazione degli oggetti nel repository.

Al termine della sessione di SPSS Modeler è necessario archiviare una nuova versione del file di progetto che comprenda i nuovi oggetti aggiunti. Il file di progetto contiene e richiama automaticamente le ultime versioni degli oggetti. Se durante una sessione di SPSS Modeler non sono stati aggiunti nuovi oggetti, non è necessario archiviare nuovamente il file di progetto. Occorre però archiviare le nuove versioni degli oggetti del progetto (stream, output e così via) che sono stati modificati.

Per archiviare un progetto

- ▶ Selezionare il progetto nella scheda CRISP-DM o Classi nel riquadro dei manager in SPSS Modeler e dal menu principale fare clic su:
File > Progetto > Archivia progetto...
- ▶ Specificare le impostazioni relative alla connessione al repository, se necessario. [Per ulteriori informazioni, vedere l'argomento Connessione a IBM SPSS Collaboration and Deployment Services Repository a pag. 159.](#) Per informazioni specifiche sulla porta, la password e altri dettagli per la connessione, rivolgersi all'amministratore del sistema locale.
- ▶ Nel repository: nella finestra di dialogo Archivia, scegliere la cartella in cui archiviare l'oggetto, specificare le eventuali altre informazioni da registrare e fare clic sul pulsante Archivia. [Per ulteriori informazioni, vedere l'argomento Impostazione delle proprietà degli oggetti a pag. 163.](#)

Archiviazione di nodi

È possibile archiviare una definizione di nodo singolo dallo stream corrente come file *.nod* nel repository, da dove altri utenti possono accedervi.

Per archiviare un nodo:

- ▶ Fare clic sul nodo con il tasto destro del mouse nell'area di disegno dello stream e fare clic su Archivia nodo.
- ▶ Specificare le impostazioni relative alla connessione al repository, se necessario. [Per ulteriori informazioni, vedere l'argomento Connessione a IBM SPSS Collaboration and Deployment Services Repository a pag. 159.](#) Per informazioni specifiche sulla porta, la password e altri dettagli per la connessione, rivolgersi all'amministratore del sistema locale.
- ▶ Nel repository: nella finestra di dialogo Archivia, scegliere la cartella in cui archiviare l'oggetto, specificare le eventuali altre informazioni da registrare e fare clic sul pulsante Archivia. [Per ulteriori informazioni, vedere l'argomento Impostazione delle proprietà degli oggetti a pag. 163.](#)

Archiviazione di oggetti di output

È possibile archiviare un oggetto di output dallo stream corrente come file *.cou* nel repository, da dove altri utenti possono accedervi.

Per archiviare un oggetto di output:

- ▶ Selezionare l'oggetto nella scheda Output del riquadro dei manager in IBM® SPSS® Modeler e dal menu principale fare clic su:
File > Output > Memorizza output...
- ▶ In alternativa, fare clic con il tasto destro del mouse su un oggetto nella scheda Output e fare clic su Archivia.
- ▶ Specificare le impostazioni relative alla connessione al repository, se necessario. [Per ulteriori informazioni, vedere l'argomento Connessione a IBM SPSS Collaboration and Deployment](#)

[Services Repository](#) a pag. 159. Per informazioni specifiche sulla porta, la password e altri dettagli per la connessione, rivolgersi all'amministratore del sistema locale.

- ▶ Nel repository: nella finestra di dialogo Archivia, scegliere la cartella in cui archiviare l'oggetto, specificare le eventuali altre informazioni da registrare e fare clic sul pulsante Archivia. [Per ulteriori informazioni, vedere l'argomento Impostazione delle proprietà degli oggetti](#) a pag. 163.

Archiviazione di modelli e palette di modelli

È possibile archiviare un modello singolo come file *.gm* nel repository, da dove altri utenti possono accedervi. È anche possibile archiviare il contenuto completo della palette Modelli come file *.gen* nel repository.

Archiviazione di un modello

- ▶ Fare clic sull'oggetto nella palette Modelli in IBM® SPSS® Modeler e nel menu principale fare clic su:
File > Modelli > Archivia modello...
- ▶ In alternativa, fare clic con il tasto destro del mouse su un oggetto nella palette Modelli e fare clic su Archivia modello.
- ▶ Continuare da "Completamento della procedura di archiviazione".

Archiviazione di una palette Modelli

- ▶ Fare clic con il tasto destro del mouse sullo sfondo della palette Modelli.
- ▶ Nel menu di scelta rapida, fare clic su Archivia palette.
- ▶ Continuare da "Completamento della procedura di archiviazione".

Completamento della procedura di archiviazione

- ▶ Specificare le impostazioni relative alla connessione al repository, se necessario. [Per ulteriori informazioni, vedere l'argomento Connessione a IBM SPSS Collaboration and Deployment Services Repository](#) a pag. 159. Per informazioni specifiche sulla porta, la password e altri dettagli per la connessione, rivolgersi all'amministratore del sistema locale.
- ▶ Nel repository: nella finestra di dialogo Archivia, scegliere la cartella in cui archiviare l'oggetto, specificare le eventuali altre informazioni da registrare e fare clic sul pulsante Archivia. [Per ulteriori informazioni, vedere l'argomento Impostazione delle proprietà degli oggetti](#) a pag. 163.

Recupero di oggetti da IBM SPSS Collaboration and Deployment Services Repository

È possibile recuperare gli stream, i modelli, le palette di modelli, i nodi, i progetti e gli oggetti di output archiviati nel IBM® SPSS® Collaboration and Deployment Services Repository.

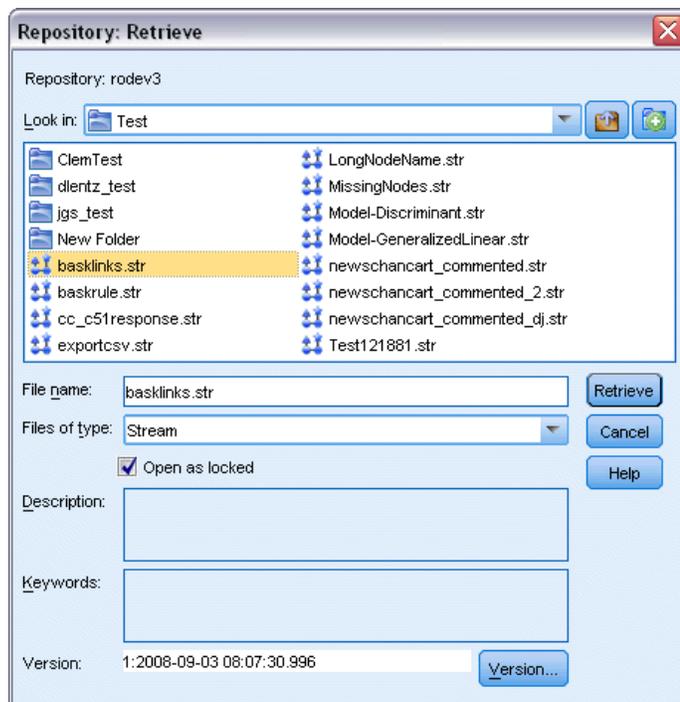
Nota: oltre a utilizzare le opzioni di menu descritte qui, è anche possibile recuperare gli stream, gli oggetti di output, i modelli e le palette di modelli facendo clic con il tasto destro del mouse sulla scheda appropriata del riquadro Manager nella parte superiore destra della finestra di IBM® SPSS® Modeler.

- ▶ Per recuperare uno stream, nel menu principale di SPSS Modeler fare clic su:
File > Recupera stream...
- ▶ Per recuperare un modello, una palette di modelli, un progetto o un oggetto di output, nel menu principale di SPSS Modeler fare clic su:
File > Modelli > Recupera modello...
- o*
File > Modelli > Recupera palette modelli...
- o*
File > Progetti > Recupera progetto...
- o*
File > Output > Recupera output...
- ▶ In alternativa, fare clic con il tasto destro del mouse nel riquadro dei manager o del progetto e fare clic su Recupera dal menu di scelta rapida.
- ▶ Per recuperare un nodo, nel menu principale di SPSS Modeler fare clic su:
Inserisci > Nodo (o Supernodo) dal repository...
- ▶ Specificare le impostazioni relative alla connessione al repository, se necessario. [Per ulteriori informazioni, vedere l'argomento Connessione a IBM SPSS Collaboration and Deployment Services Repository a pag. 159.](#) Per informazioni specifiche sulla porta, la password e altri dettagli per la connessione, rivolgersi all'amministratore del sistema locale.
- ▶ Nel repository: dalla finestra di dialogo Recupera, navigare fino all'oggetto, selezionarlo e fare clic sul pulsante Recupera.

Scelta di un oggetto da recuperare

Figura 9-12

Recupero di un oggetto dal IBM SPSS Collaboration and Deployment Services Repository



Cerca in. Mostra la gerarchia delle cartelle per la cartella corrente. Per navigare a una cartella diversa, selezionarne una da questo elenco per arrivare direttamente a quella cartella, oppure navigare utilizzando l'elenco di oggetti sotto a questo campo.

Pulsante Cartella superiore. Si sposta di un livello verso l'alto rispetto alla cartella corrente nella gerarchia.

Pulsante Nuova cartella. Crea una nuova cartella al livello corrente nella gerarchia.

Nome file. Il nome file del repository dell'oggetto selezionato. Per recuperare quell'oggetto, fare clic su Recupera.

File di tipo. Il tipo di oggetto che si è scelto di recuperare. Solo gli oggetti di questo tipo, insieme alle cartelle, vengono mostrati nell'elenco degli oggetti. Per visualizzare gli oggetti di un tipo diverso di recupero, selezionare il tipo di oggetto dall'elenco.

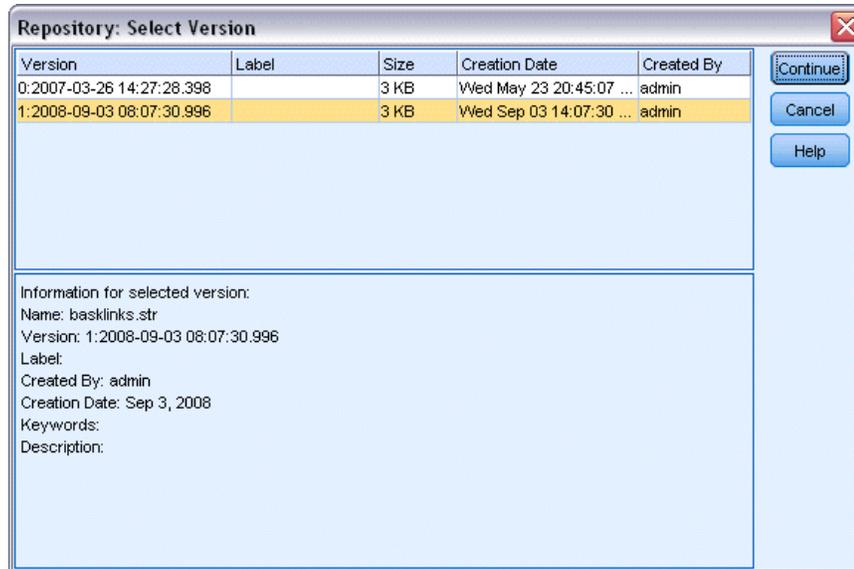
Apri come bloccato. Per default, quando un oggetto viene recuperato, è bloccato nel repository in modo che altri non possano aggiornarlo. Se non si desidera che l'oggetto venga bloccato al recupero, deselegionare questa casella.

Descrizione, Parole chiave. Se erano stati definiti ulteriori dettagli sull'oggetto al momento dell'archiviazione, tali dettagli verranno visualizzati qui. [Per ulteriori informazioni, vedere l'argomento Aggiunta di informazioni sugli oggetti archiviati a pag. 164.](#)

Versione. Per recuperare una versione dell'oggetto diversa dall'ultima, fare clic su questo pulsante. Vengono visualizzate informazioni dettagliate su tutte le versioni ed è possibile scegliere la versione desiderata.

Selezione della versione di un oggetto

Figura 9-13
Selezione della versione di un oggetto



Per selezionare una versione specifica di un oggetto del repository.

- ▶ (Facoltativo) Ordinare l'elenco per versione, etichetta, dimensione, data di creazione o autore della creazione, facendo doppio clic sull'intestazione della colonna appropriata.
- ▶ Selezionare la versione dell'oggetto con cui si desidera lavorare.
- ▶ Fare clic su Continua.

Ricerca di oggetti in IBM SPSS Collaboration and Deployment Services Repository

È possibile effettuare una ricerca degli oggetti in base al nome, alla cartella, al tipo, all'etichetta o in base ad altri criteri.

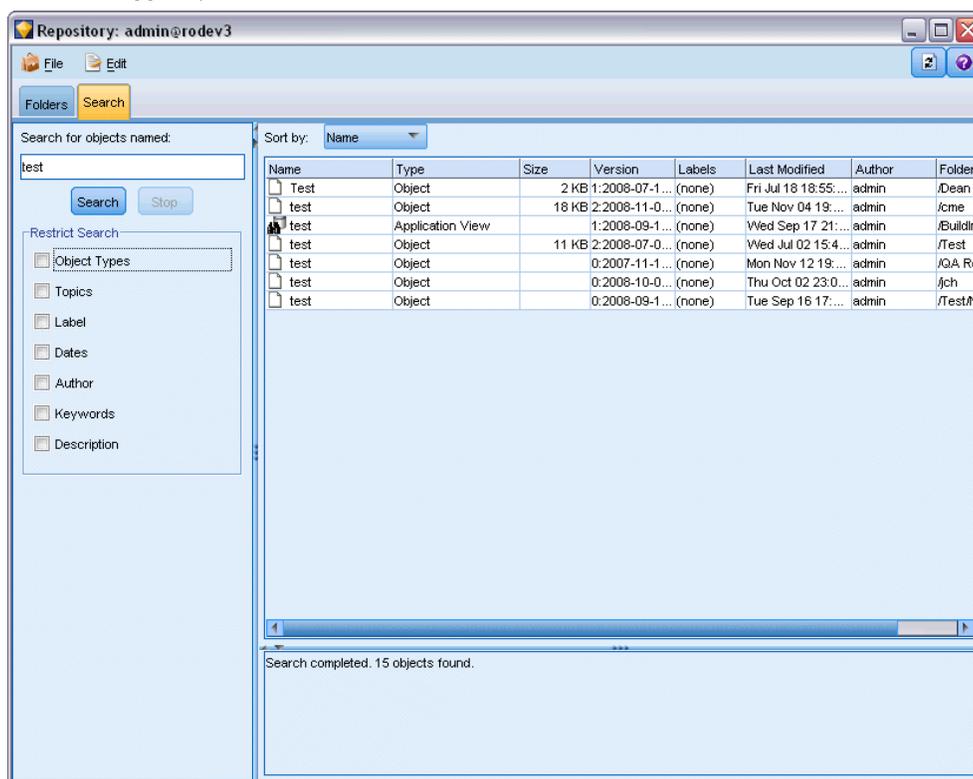
Ricerca per nome

Per cercare gli oggetti in base al nome:

- ▶ Nel menu principale di IBM® SPSS® Modeler, fare clic su:
Strumenti > IBM® SPSS® Collaboration and Deployment Services Repository > Esplora...

- ▶ Specificare le impostazioni relative alla connessione al repository, se necessario. [Per ulteriori informazioni, vedere l'argomento Connessione a IBM SPSS Collaboration and Deployment Services Repository a pag. 159.](#) Per informazioni specifiche sulla porta, la password e altri dettagli per la connessione, rivolgersi all'amministratore del sistema locale.
- ▶ Fare clic sulla scheda Cerca.
- ▶ Nel campo Cerca oggetti con nome, specificare il nome dell'oggetto che si desidera trovare.

Figura 9-14
Ricerca di oggetti per nome



Quando si cercano oggetti in base al nome è possibile utilizzare l'asterisco (*) come carattere jolly al posto di qualsiasi stringa di caratteri e il punto interrogativo (?) al posto di un singolo carattere. Per esempio, con la chiave di ricerca **cluster** saranno trovati tutti gli oggetti il cui nome contiene la stringa *cluster*. La stringa di ricerca *m0?_** restituirà *M01_cluster.str* e *M02_cluster.str*, ma non *M01a_cluster.str*. La funzione di ricerca non fa distinzione tra maiuscole e minuscole (*cluster* restituirà *Cluster* e anche *CLUSTER*).

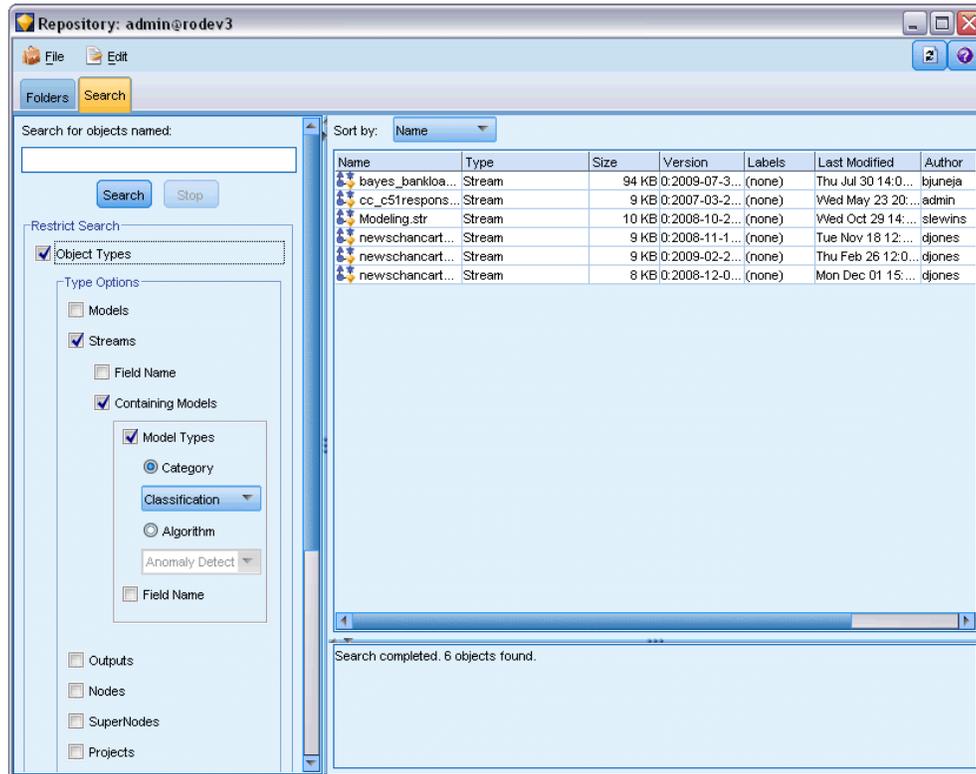
Nota: se gli oggetti sono molto numerosi, le ricerche possono richiedere più tempo.

Ricerca in base ad altri criteri

È possibile eseguire una ricerca in base a titolo, etichetta, date, autore, parole chiave, contenuto dell'indice o descrizione. Verranno trovati solo gli oggetti corrispondenti a *tutti* i criteri di ricerca specificati. Per esempio, si potrebbero individuare tutti gli stream contenenti uno o più

modelli di cluster che hanno anche un'etichetta specifica e che sono stati modificati dopo una determinata data.

Figura 9-15
Ricerca di stream contenenti un tipo specifico di modello



Tipi di oggetti. È possibile limitare la ricerca a modelli, stream, output, nodi, Supernodi, progetti, palette di modelli, scenari o altri tipi di oggetti.

- **Modelli.** È possibile cercare i modelli in base alla categoria (classificazione, approssimazione, raggruppamento cluster, ecc.) o in base a un algoritmo di modellazione specifico quale Kohonen.

È possibile inoltre eseguire la ricerca in base ai campi utilizzati, per esempio cercare tutti i modelli che utilizzano un campo denominato *reddito* come campo di input o di output (obiettivo).

- **Stream.** Per gli stream è possibile limitare la ricerca in base ai campi utilizzati o al tipo di modello (categoria o algoritmo) contenuti nello stream.

Argomenti. È possibile eseguire la ricerca sui modelli associati ad argomenti specifici da un elenco impostato dagli utenti del repository che dispongono dei privilegi appropriati (per maggiori informazioni, vedere il manuale *Deployment Manager User's Guide*). Per ottenere tale elenco, selezionare questa casella, quindi fare clic sul pulsante Aggiungi argomenti visualizzato, selezionare uno o più argomenti dall'elenco e fare clic su OK.

Etichetta. Limita la ricerca a specifiche etichette delle versioni degli oggetti.

Date. È possibile specificare una data di creazione o di modifica e cercare gli oggetti creati o modificati prima, dopo o nell'intervallo di date specificato.

Autore. Limita la ricerca agli oggetti creati da un utente specifico.

Parole chiave. Consente di eseguire la ricerca in base alle parole chiave specificate. In SPSS Modeler è possibile specificare le parole chiave nella scheda Annotazione di uno stream, modello o oggetto di output.

Descrizione. Consente di cercare termini specifici nel campo della descrizione. In SPSS Modeler, la descrizione di uno stream, un modello o un oggetto di output è specificata nella scheda Annotazione. Se sono presenti più stringhe di ricerca, è possibile separarle con un punto e virgola, per esempio reddito; tipo raccolto; valore domanda. Si noti che, all'interno di una stringa di ricerca, gli spazi sono importanti. Per esempio, tipo di raccolto con uno spazio e tipo di raccolto con due spazi non si equivalgono.

Modifica degli oggetti di IBM SPSS Collaboration and Deployment Services Repository

È possibile modificare gli oggetti esistenti nel IBM® SPSS® Collaboration and Deployment Services Repository direttamente da IBM® SPSS® Modeler. È possibile:

- Creare, rinominare o eliminare cartelle
- Bloccare o sbloccare oggetti
- Eliminare oggetti

Creazione, ridenominazione ed eliminazione delle cartelle

- ▶ Per eseguire operazioni sulle cartelle nel IBM® SPSS® Collaboration and Deployment Services Repository, dal menu principale di IBM® SPSS® Modeler fare clic su:
Strumenti > IBM® SPSS® Collaboration and Deployment Services Repository > Esplora...
- ▶ Specificare le impostazioni relative alla connessione al repository, se necessario. [Per ulteriori informazioni, vedere l'argomento Connessione a IBM SPSS Collaboration and Deployment Services Repository a pag. 159.](#) Per informazioni specifiche sulla porta, la password e altri dettagli per la connessione, rivolgersi all'amministratore del sistema locale.
- ▶ Assicurarsi che la scheda Cartelle sia attiva.
- ▶ Per creare una nuova cartella, fare clic con il tasto destro del mouse sulla cartella principale e fare clic su Nuova cartella.
- ▶ Per rinominare una cartella, fare clic sulla cartella con il tasto destro del mouse e fare clic su Rinomina cartella.
- ▶ Per eliminare una cartella, fare clic sulla cartella con il tasto destro del mouse e fare clic su Elimina cartella.

Blocco e sblocco degli oggetti di IBM SPSS Collaboration and Deployment Services Repository

È possibile bloccare un oggetto per impedire qualsiasi aggiornamento delle sue versioni esistenti o la creazione di nuove versioni. Un oggetto bloccato è contrassegnato dal simbolo del lucchetto sulla relativa icona:

Figura 9-16
Oggetto bloccato



Per bloccare un oggetto

- ▶ Nella finestra di navigazione del IBM® SPSS® Collaboration and Deployment Services Repository, fare clic sull'oggetto richiesto con il tasto destro del mouse.
- ▶ Fare clic su Blocca.

Per sbloccare un oggetto

- ▶ Nella finestra di navigazione del IBM SPSS Collaboration and Deployment Services Repository, fare clic sull'oggetto richiesto con il tasto destro del mouse.
- ▶ Fare clic su Sblocca.

Eliminazione degli oggetti da IBM SPSS Collaboration and Deployment Services Repository

Prima di eliminare un oggetto dal IBM® SPSS® Collaboration and Deployment Services Repository occorre decidere se si desidera eliminare tutte le versioni o solo una determinata versione dell'oggetto.

Per eliminare tutte le versioni di un oggetto

- ▶ Nella finestra di navigazione del IBM SPSS Collaboration and Deployment Services Repository, fare clic sull'oggetto richiesto con il tasto destro del mouse.
- ▶ Fare clic su Elimina oggetti.

Per eliminare la versione più recente di un oggetto

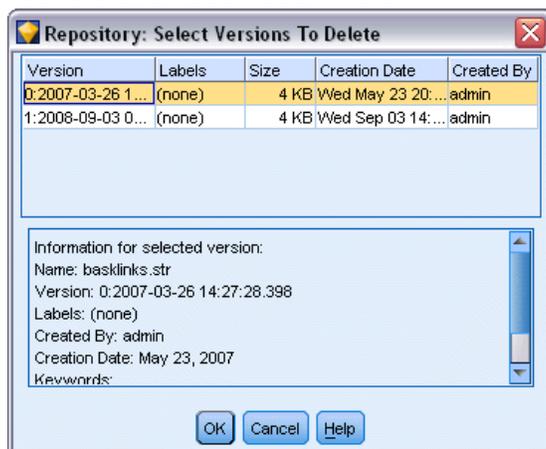
- ▶ Nella finestra di navigazione del IBM SPSS Collaboration and Deployment Services Repository, fare clic sull'oggetto richiesto con il tasto destro del mouse.
- ▶ Fare clic su Elimina.

Per eliminare una versione precedente di un oggetto

- ▶ Nella finestra di navigazione del IBM SPSS Collaboration and Deployment Services Repository, fare clic sull'oggetto richiesto con il tasto destro del mouse.

- ▶ Fare clic su Elimina versioni.
- ▶ Selezionare la versione o le versioni da eliminare e fare clic su OK.

Figura 9-17
Selezione delle versioni da eliminare



Gestione delle proprietà degli oggetti di IBM SPSS Collaboration and Deployment Services Repository

Da IBM® SPSS® Modeler è possibile controllare diverse proprietà degli oggetti. È possibile:

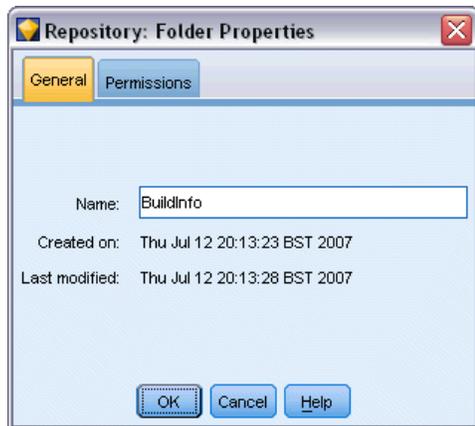
- Visualizzare le proprietà di una cartella
- Visualizzare e modificare le proprietà di un oggetto
- Creare, applicare ed eliminare le etichette di versione per un oggetto

Visualizzazione delle proprietà delle cartelle

Per visualizzare le proprietà di una cartella nella finestra del IBM® SPSS® Collaboration and Deployment Services Repository, fare clic con il tasto destro del mouse sulla cartella richiesta. Fare clic su Proprietà cartella.

scheda Generale

Figura 9-18
Proprietà cartella

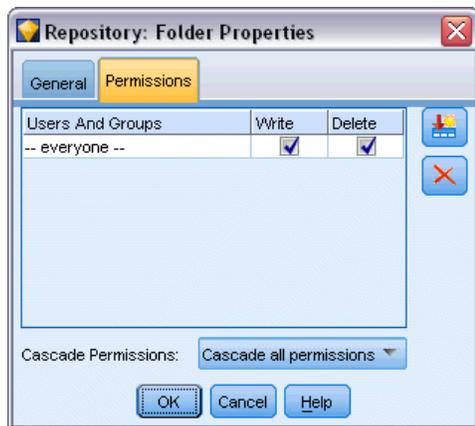


Visualizza il nome della cartella e le date di creazione e di modifica.

Scheda Autorizzazioni

Specifica le autorizzazioni di lettura e scrittura per la cartella. Sono elencati tutti gli utenti e i gruppi che dispongono dell'accesso alla cartella principale. Le autorizzazioni seguono una gerarchia. Per esempio, se non si dispone dell'autorizzazione alla lettura, non si potrà avere l'autorizzazione alla scrittura e se non si dispone dell'autorizzazione alla scrittura non si potrà avere l'autorizzazione all'eliminazione.

Figura 9-19
Proprietà cartella



Utenti e gruppi. Elenca gli utenti e i gruppi del repository che hanno almeno l'accesso in lettura a questa cartella. Selezionare le caselle di controllo Scrittura ed Eliminazione per aggiungere questi diritti di accesso per la cartella a un utente o un gruppo particolare. Fare clic sull'icona Aggiungi utenti o gruppi nella parte destra della scheda Autorizzazioni per assegnare l'accesso a utenti e gruppi aggiuntivi. L'elenco degli utenti e dei gruppi disponibili è gestito dall'amministratore.

Sovrapponi autorizzazioni. Scegliere un'opzione per controllare il modo in cui le modifiche apportate alla cartella corrente si applicano alle rispettive cartelle secondarie, se presenti.

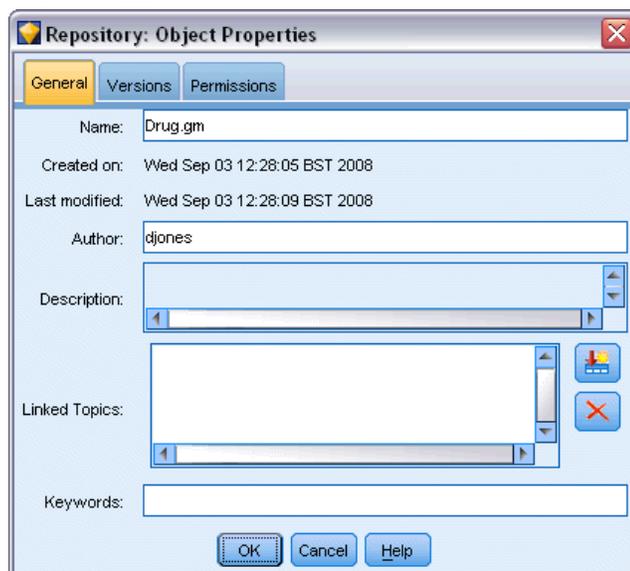
- **Sovrapponi tutte le autorizzazioni.** Sovrappone le autorizzazioni della cartella corrente a tutte le sottocartelle. In questo modo è possibile assegnare in modo rapido le autorizzazioni a più cartelle contemporaneamente. Impostare le autorizzazioni richieste per la cartella principale e quindi sovrapporre nel modo richiesto.
- **Sovrapponi solo modifiche.** Sovrappone solo le modifiche eseguite dall'ultima applicazione di modifiche. Per esempio, se è stato aggiunto un nuovo gruppo e si desidera assegnare a tale gruppo l'accesso a tutte le cartelle nella filiale Vendite, è possibile assegnare al gruppo l'accesso alla cartella Vendite radice e quindi sovrapporre la modifica a tutte le sottocartelle. Tutte le altre autorizzazioni per le cartelle esistenti rimarranno invariate.
- **Non sovrapporre.** Qualsiasi modifica viene applicata solo alla cartella corrente e non viene sovrapposta alle cartelle figlio.

Visualizzazione e modifica delle proprietà degli oggetti

La finestra di dialogo Proprietà oggetto del IBM® SPSS® Collaboration and Deployment Services Repository consente di visualizzare e modificare le proprietà. Alcune proprietà non possono essere modificate, ma è sempre possibile aggiornare un oggetto aggiungendo una nuova versione.

- ▶ Nella finestra del IBM SPSS Collaboration and Deployment Services Repository, fare clic sull'oggetto richiesto con il tasto destro del mouse.
- ▶ Fare clic su Proprietà oggetto.

Figura 9-20
Proprietà oggetto



Scheda Generale

Nome. Nome dell'oggetto visualizzato in IBM SPSS Collaboration and Deployment Services Repository.

Data di creazione. Data in cui l'oggetto (non la versione) è stato creato.

Ultima modifica. Data in cui è stata modificata la versione più recente.

Autore. Il nome di login dell'utente.

Descrizione. Per default, la descrizione corrisponde a quella specificata nella scheda Annotazione dell'oggetto in IBM® SPSS® Modeler.

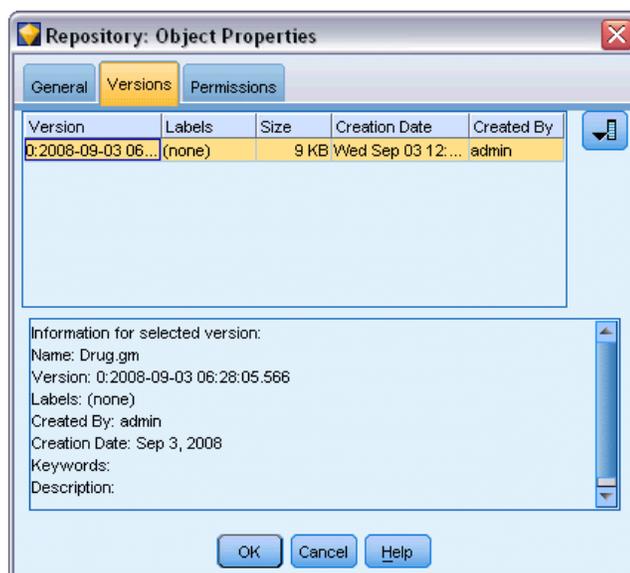
Argomenti collegati. Nel IBM SPSS Collaboration and Deployment Services Repository è possibile organizzare per argomenti i modelli e gli oggetti correlati. L'elenco degli argomenti disponibili è impostato dagli utenti del repository che dispongono dei privilegi appropriati (per maggiori informazioni, vedere il manuale *Deployment Manager User's Guide*).

Parole chiave. Le parole chiave sono specificate nella scheda Annotazione di uno stream, un modello o un oggetto di output. Se sono presenti più parole chiave, devono essere separate da spazi. Le parole chiave possono contenere al massimo 255 caratteri. (Se le parole chiave contengono spazi, separarle mediante le virgolette).

Scheda Versioni

Gli oggetti archiviati nel IBM SPSS Collaboration and Deployment Services Repository possono avere più versioni. Nella scheda Versioni sono visualizzate le informazioni relative a ogni versione.

Figura 9-21
Proprietà della versione



Per le versioni specifiche di un oggetto memorizzato è possibile definire o modificare le proprietà seguenti:

Versione. Identificatore univoco della versione, creato in base all'ora in cui la versione è stata archiviata.

Etichetta. Etichetta corrente della versione, se esistente. A differenza dell'identificatore di versione, le etichette possono essere spostate da una versione dell'oggetto a un'altra.

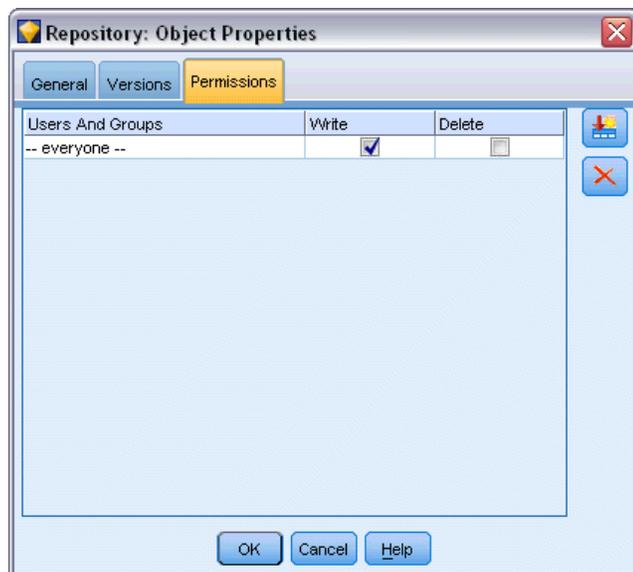
Per ogni versione vengono inoltre visualizzati la dimensione, la data di creazione e l'autore del file.

Modifica etichette. Fare clic sull'icona Modifica etichette nella parte superiore destra della scheda Versioni per definire, applicare o rimuovere le etichette per gli oggetti memorizzati. [Per ulteriori informazioni, vedere l'argomento Gestione delle etichette di versione degli oggetti a pag. 184.](#)

Scheda Autorizzazioni

La scheda Autorizzazioni consente di impostare le autorizzazioni di lettura e scrittura per l'oggetto. Sono elencati tutti gli utenti e i gruppi che dispongono dell'accesso all'oggetto corrente. Le autorizzazioni seguono una gerarchia. Per esempio, se non si dispone dell'autorizzazione alla lettura, non si potrà avere l'autorizzazione alla scrittura e se non si dispone dell'autorizzazione alla scrittura non si potrà avere l'autorizzazione all'eliminazione.

Figura 9-22
Diritti di accesso per l'oggetto



Utenti e gruppi. Elenca gli utenti e i gruppi del repository che hanno almeno l'accesso in lettura a questo oggetto. Selezionare le caselle di controllo Scrittura ed Eliminazione per aggiungere questi diritti di accesso per questo oggetto a un utente o un gruppo particolare. Fare clic sull'icona Aggiungi utenti o gruppi nella parte destra della scheda Autorizzazioni per assegnare l'accesso a utenti e gruppi aggiuntivi. L'elenco degli utenti e dei gruppi disponibili è gestito dall'amministratore.

Gestione delle etichette di versione degli oggetti

La finestra di dialogo Modifica etichette versione consente di:

- Applicare etichette all'oggetto selezionato
- Rimuovere etichette dall'oggetto selezionato
- Definire una nuova etichetta e applicarla all'oggetto

Per applicare etichette all'oggetto

- ▶ Selezionare una o più etichette nell'elenco Etichette disponibili.
- ▶ Fare clic sul pulsante freccia destra per spostare le etichette selezionate nell'elenco Etichette applicate.
- ▶ Fare clic su OK.

Per rimuovere etichette dall'oggetto

- ▶ Selezionare una o più etichette nell'elenco Etichette applicate.
- ▶ Fare clic sul pulsante freccia sinistra per spostare le etichette selezionate nell'elenco Etichette disponibili.
- ▶ Fare clic su OK.

Per definire una nuova etichetta e applicarla all'oggetto

- ▶ Digitare il nome dell'etichetta nel campo Nuova etichetta.
- ▶ Fare clic sul pulsante freccia destra per spostare la nuova etichetta nell'elenco Etichette applicate.
- ▶ Fare clic su OK.

Deployment di stream

Per consentire l'utilizzo di uno stream con l'applicazione thin client IBM® SPSS® Modeler Advantage, esso deve essere implementato come stream (file *.str*) nel repository.

Per sfruttare appieno le funzioni aziendali di IBM® SPSS® Collaboration and Deployment Services, è necessario eseguire il deployment di uno stream come scenario (file *.scn*) nel repository. [Per ulteriori informazioni, vedere l'argomento Archiviazione e deployment di oggetti di IBM SPSS Collaboration and Deployment Services Repository a pag. 158.](#)

Per eseguire il deployment dello stream corrente (metodo del menu File)

- ▶ Nel menu principale, fare clic su:
File > Archivia > Deployment
- ▶ Scegliere il tipo di deployment e compilare il resto della finestra di dialogo in funzione delle esigenze.

- ▶ Fare clic su Implementa come stream per implementare lo stream e utilizzarlo con IBM SPSS Modeler Advantage, oppure su Implementa come scenario per implementare lo stream e utilizzarlo con IBM SPSS Collaboration and Deployment Services.
- ▶ Fare clic sul pulsante Archivia. Per maggiori informazioni, fare clic sul pulsante Aiuto.
- ▶ Continuare da “Completamento del processo di deployment”.

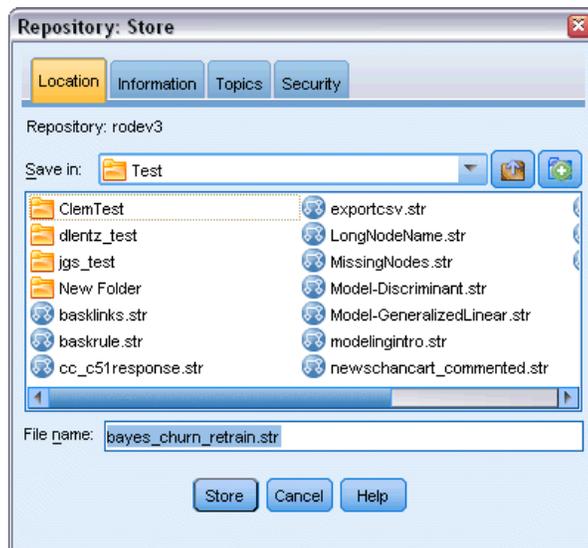
Per eseguire il deployment dello stream corrente (metodo del menu Strumenti)

- ▶ Nel menu principale, fare clic su:
Strumenti > Proprietà stream > Deployment
- ▶ Scegliere il tipo di deployment, completare la parte restante della scheda Deployment se necessario e fare clic sul pulsante Archivia. [Per ulteriori informazioni, vedere l'argomento Opzioni di deployment degli stream a pag. 186.](#)

Completamento del processo di deployment

- ▶ Specificare le impostazioni relative alla connessione al repository, se necessario. [Per ulteriori informazioni, vedere l'argomento Connessione a IBM SPSS Collaboration and Deployment Services Repository a pag. 159.](#) Per informazioni specifiche sulla porta, la password e altri dettagli per la connessione, rivolgersi all'amministratore del sistema locale.

Figura 9-23
Archiviazione di uno stream nel repository



- ▶ Nel repository: nella finestra di dialogo Archivia, scegliere la cartella in cui archiviare l'oggetto, specificare le eventuali altre informazioni da registrare e fare clic sul pulsante Archivia. [Per ulteriori informazioni, vedere l'argomento Impostazione delle proprietà degli oggetti a pag. 163.](#)

Opzioni di deployment degli stream

La scheda Deployment nella finestra di dialogo Opzioni stream consente di specificare le opzioni per il deployment dello stream, che può essere implementato come stream o come scenario.

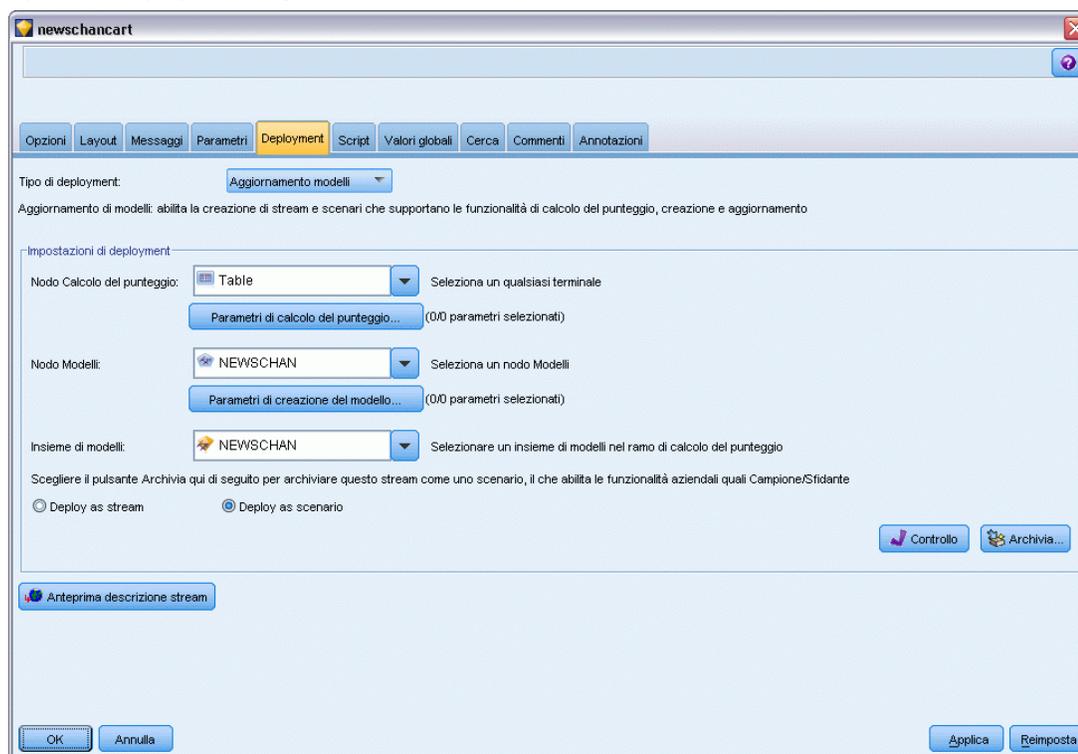
Quando si effettua il deployment come stream è possibile aprire e modificare lo stream nell'applicazione thin client IBM® SPSS® Modeler Advantage. Lo stream viene archiviato nel repository come file con estensione *.str*.

Quando si esegue il deployment di uno stream come scenario è possibile sfruttare appieno le funzionalità aggiuntive disponibili con IBM® SPSS® Collaboration and Deployment Services, per esempio l'accesso a più utenti, il calcolo automatico del punteggio, l'aggiornamento del modello e l'analisi Campione/Sfidante. Con il deployment come scenario, lo stream viene archiviato nel repository come file con estensione *.scn*. In questo modo, lo stream può essere utilizzato anche da Applicazioni predittive.

Da questa scheda è possibile eseguire un'anteprima della descrizione dello stream creata da IBM® SPSS® Modeler. [Per ulteriori informazioni, vedere l'argomento Descrizioni stream in il capitolo 5 a pag. 64.](#)

Nota: per assicurare un accesso regolare ai dati aziendali, uno stream sottoposto a deployment come scenario deve accedere ai propri dati di origine tramite IBM® SPSS® Collaboration and Deployment Services Enterprise View; in tal caso lo stream deve includere almeno un nodo di input Enterprise View all'interno di ogni ramo designato di calcolo del punteggio o di creazione del modello, a seconda dei casi. [Per ulteriori informazioni, vedere l'argomento Nodo Enterprise View in il capitolo 2 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)

Figura 9-24
Opzioni di deployment degli stream



Tipo di deployment. Scegliere come si desidera eseguire il deployment dello stream. Per poter essere sottoposti a deployment come scenario, tutti gli stream necessitano di un nodo di calcolo del punteggio designato; gli altri requisiti e opzioni dipendono dal tipo di deployment.

- <nessuno>. Il deployment dello stream nel repository non sarà eseguito. Tutte le opzioni sono disattivate tranne l'anteprima descrizione stream.
- Solo calcolo del punteggio. Il deployment dello stream nel repository avviene quando si fa clic sul pulsante Archivia. Il punteggio dei dati si può calcolare mediante il nodo designato nel campo Nodo Calcolo del punteggio.
- Aggiornamento modelli. Uguale a Solo calcolo del punteggio, ma consente anche di aggiornare il modello nel repository utilizzando gli oggetti specificati nei campi Nodo Modelli e Insieme di modelli.

Nodo Calcolo del punteggio. Selezionare un nodo grafico, di output o di esportazione per identificare il ramo dello stream da utilizzare per il calcolo del punteggio dei dati. Lo stream può contenere un numero qualsiasi di rami, modelli e nodi terminali validi, ma solo un ramo di calcolo del punteggio deve essere designato per il deployment. Questo è il requisito di base per il deployment di qualsiasi stream.

Parametri di calcolo del punteggio. Consente di specificare i parametri che possono essere modificati quando viene eseguito il ramo di calcolo del punteggio. [Per ulteriori informazioni, vedere l'argomento Parametri di calcolo del punteggio e di creazione del modello a pag. 189.](#)

Nodo Modelli. Per l'aggiornamento dei modelli, specifica il nodo Modelli utilizzato per rigenerare o aggiornare il modello nel repository. Deve essere un nodo Modelli dello stesso tipo di quello specificato per Insieme di modelli.

Parametri di creazione del modello. Consente di specificare i parametri che possono essere modificati quando viene eseguito il nodo Modelli. [Per ulteriori informazioni, vedere l'argomento Parametri di calcolo del punteggio e di creazione del modello a pag. 189.](#)

Insieme di modelli. Per l'aggiornamento dei modelli, specifica l'insieme di modelli da aggiornare o rigenerare ogni volta che viene aggiornato lo scenario nel repository (generalmente all'interno di un processo pianificato). Il modello deve essere situato sul ramo di calcolo del punteggio. Sebbene un ramo di calcolo del punteggio possa includere più modelli, è possibile specificarne solo uno. Si noti che quando lo scenario viene inizialmente creato, il modello potrebbe essere un semplice modello segnaposto che viene aggiornato o rigenerato non appena sono disponibili nuovi dati.

Implementa come stream. Fare clic su questa opzione se si desidera utilizzare lo stream con IBM SPSS Modeler Advantage.

Implementa come scenario. Fare clic su questa opzione se si desidera utilizzare lo stream con IBM SPSS Collaboration and Deployment Services (vedere la nota seguente).

Controllo. Fare clic su questo pulsante per controllare se lo stream è valido per il deployment. In caso di implementazione come scenario, lo stream deve contenere un nodo di input Enterprise View valido; se questo non è presente, viene visualizzato un messaggio di errore.

Archivia. Se lo stream è valido, esegue il deployment dello stream come scenario. In caso contrario, viene visualizzato un messaggio di errore. Fare clic sul pulsante Correggi, correggere l'errore e riprovare.

Anteprima descrizione stream. Consente di visualizzare il contenuto della descrizione dello stream creata da SPSS Modeler. [Per ulteriori informazioni, vedere l'argomento Descrizioni stream in il capitolo 5 a pag. 64.](#)

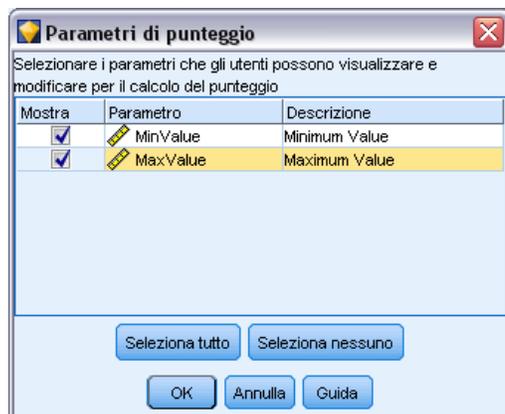
Nota: (solo per Implementa come scenario) nel ramo di modellazione è possibile utilizzare più nodi Enterprise View. In questo caso, l'uso di un'unica connessione dati per tutti i nodi Enterprise View all'interno del ramo è preferibile nella maggior parte dei casi e necessario per l'analisi Campione/Sfidante.

- Se il supporto di Campione/Sfidante non è richiesto, è possibile utilizzare diverse connessioni Enterprise View all'interno dello stesso ramo, a condizione che le connessioni varino solo in termini di Definizione provider di dati.
- Queste limitazioni sono applicabili solo all'interno di un dato ramo. Tra i rami di calcolo del punteggio e creazione del modello, è possibile utilizzare diverse connessioni Enterprise View senza tali limitazioni.

Parametri di calcolo del punteggio e di creazione del modello

Quando si effettua il deployment di uno stream a IBM® SPSS® Collaboration and Deployment Services, è possibile scegliere i parametri da visualizzare o modificare ogni volta che il modello viene aggiornato o di cui viene calcolato il punteggio. Per esempio, è possibile specificare valori minimi e massimi o un altro valore soggetto a modifica ogni volta che viene eseguito il processo.

Figura 9-25
Finestra di dialogo Parametri di calcolo del punteggio



- Per rendere visibile un parametro in modo che possa essere visualizzato o modificato dopo il deployment dello scenario, selezionarlo dall'elenco nella finestra di dialogo.

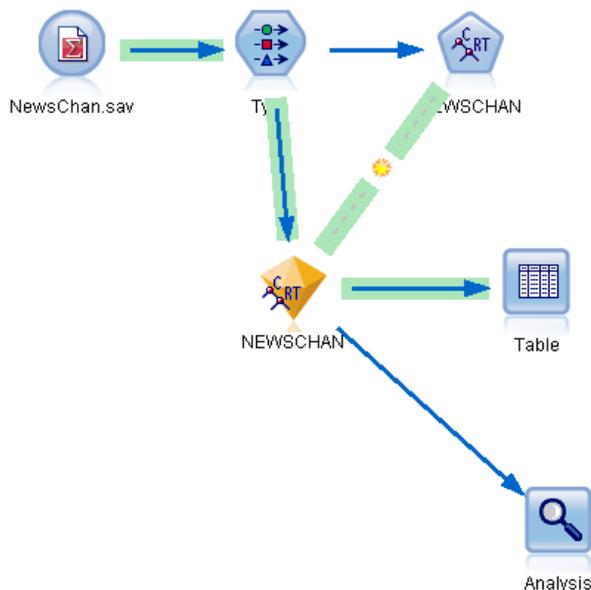
L'elenco di parametri disponibili è definito nella scheda Parametri della finestra di dialogo delle proprietà dello stream. [Per ulteriori informazioni, vedere l'argomento Impostazione di parametri per stream e sessione in il capitolo 5 a pag. 57.](#)

Il ramo di calcolo del punteggio

Se si sta eseguendo il deployment di uno stream, è necessario designare un ramo dello stream come **ramo di calcolo del punteggio** (ovvero, il ramo che contiene il nodo di calcolo del punteggio). Quando si designa un ramo come ramo di calcolo del punteggio, quel ramo viene evidenziato nell'area di disegno stream, così come il collegamento del modello all'insieme di modelli sul ramo di calcolo del punteggio. Questa rappresentazione grafica è particolarmente utile nel caso di stream complessi che presentano più rami, dove il ramo di calcolo del punteggio potrebbe non essere immediatamente ovvio.

Nota: è possibile designare come ramo di calcolo del punteggio un solo ramo dello stream.

Figura 9-26
Stream con ramo di calcolo del punteggio evidenziato



Se per lo stream era già stato definito un ramo di calcolo del punteggio, il nuovo ramo designato sostituisce quello esistente come ramo di calcolo del punteggio. È possibile impostare il colore dell'evidenziazione del ramo di calcolo del punteggio utilizzando un'opzione Colore personalizzato. [Per ulteriori informazioni, vedere l'argomento Impostazione delle opzioni di visualizzazione in il capitolo 12 a pag. 231.](#)

È possibile mostrare o nascondere l'evidenziazione del ramo di calcolo del punteggio utilizzando il pulsante della barra degli strumenti Mostra/nascondi markup dello stream.

Figura 9-27
Pulsante della barra degli strumenti Mostra/nascondi markup dello stream



Identificazione del ramo di calcolo del punteggio per il deployment

È possibile designare il ramo di calcolo del punteggio dal menu di scelta rapida di un nodo terminale, oppure dal menu Strumenti. Se si utilizza il menu di scelta rapida, il nodo di calcolo del punteggio viene impostato automaticamente nella scheda Deployment delle proprietà dello stream.

Per designare un ramo come ramo di calcolo del punteggio (menu di scelta rapida)

- Collegare l'insieme di modelli a un nodo terminale (un nodo di elaborazione o di output a valle rispetto all'insieme).

- ▶ Fare clic con il tasto destro del mouse sul nodo terminale.
- ▶ Nel menu, fare clic su Utilizza come ramo di calcolo del punteggio.

Per designare un ramo come ramo di calcolo del punteggio (menu Strumenti)

- ▶ Collegare l'insieme di modelli a un nodo terminale (un nodo di elaborazione o di output a valle rispetto all'insieme).
- ▶ Nel menu principale, fare clic su:
Strumenti > Proprietà stream > Deployment
- ▶ Nell'elenco Tipo di deployment, fare clic su Solo calcolo del punteggio o Aggiornamento modelli, in base alle necessità. [Per ulteriori informazioni, vedere l'argomento Opzioni di deployment degli stream a pag. 186.](#)
- ▶ Fare clic sul campo Nodo Calcolo del punteggio e selezionare un nodo terminale dall'elenco.
- ▶ Fare clic su OK.

Aggiornamento modelli

L'aggiornamento modelli è il processo di ricostruzione di un modello esistente in uno scenario utilizzando dati più recenti. Lo stream in sé e per sé non viene modificato nel repository. Per esempio, il tipo di algoritmo e le impostazioni specifiche dello stream rimangono invariate, ma il modello viene riaddestrato con nuovi dati e aggiornato se la nuova versione funziona meglio della precedente.

È possibile impostare per l'aggiornamento un solo insieme di modelli in uno scenario. Questa operazione è nota come **aggiornamento modelli**. Se si fa clic sull'opzione Aggiornamento modelli nella scheda Deployment delle proprietà dello stream (vedere Opzioni di deployment degli stream a pag. 186), l'insieme di modelli designato in quel momento diventa il modello di aggiornamento. È possibile inoltre designare un modello come modello di aggiornamento dal menu di scelta rapida di un insieme di modelli. Per fare questo, però, l'insieme deve già trovarsi sul ramo di calcolo del punteggio.

Disattivare lo stato di “modello di aggiornamento” di un insieme equivale a impostare il tipo di deployment dello scenario su Solo calcolo del punteggio, e in questo caso la scheda Deployment della finestra di dialogo delle proprietà dello stream verrebbe aggiornata di conseguenza. È possibile attivare e disattivare questo stato mediante l'opzione Utilizza come modello di aggiornamento nel menu di scelta rapida dell'insieme sul ramo di calcolo del punteggio corrente.

La rimozione del collegamento del modello di un insieme sul ramo di calcolo del punteggio rimuove anche lo stato di “modello di aggiornamento” dell'insieme. La rimozione del collegamento del modello si può annullare mediante il menu Modifica o la barra degli strumenti; in questo modo si ripristina inoltre lo status di “modello di aggiornamento” dell'insieme di modelli.

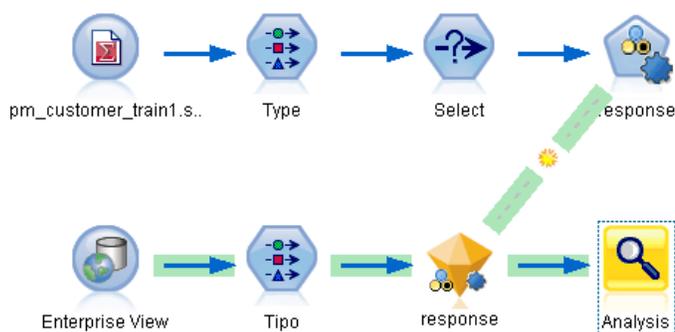
Modalità di selezione del modello di aggiornamento

Così come avviene per il ramo di calcolo del punteggio, anche il collegamento al modello di aggiornamento appare evidenziato nello stream. L'insieme di modelli scelto come modello di aggiornamento, e di conseguenza il collegamento evidenziato, dipende dal numero di insiemi presenti nello stream.

Modello singolo nello stream

Figura 9-28

Ramo di calcolo del punteggio con modello singolo nello stream



Se un insieme di modelli singolo con collegamento si trova sul ramo di calcolo del punteggio quando è identificato come tale, quell'insieme diventa il modello di aggiornamento dello scenario.

Modelli multipli nello stream

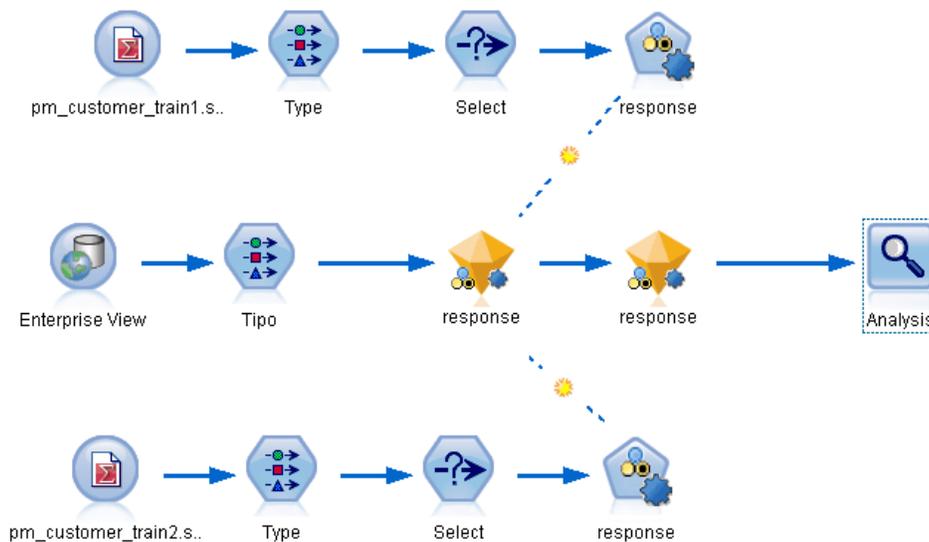
Se nello stream sono presenti più insiemi con collegamento, il modello di aggiornamento viene scelto secondo le modalità descritte di seguito.

Se un insieme di modelli è stato definito nella scheda Deployment della finestra di dialogo delle proprietà dello stream e si trova anche nello stream, quell'insieme diventa il modello di aggiornamento.

Se nella scheda Deployment non è stato definito alcun insieme, oppure c'è un insieme definito ma non si trova sul ramo di calcolo del punteggio, allora l'insieme più vicino al nodo terminale diventa il modello di aggiornamento.

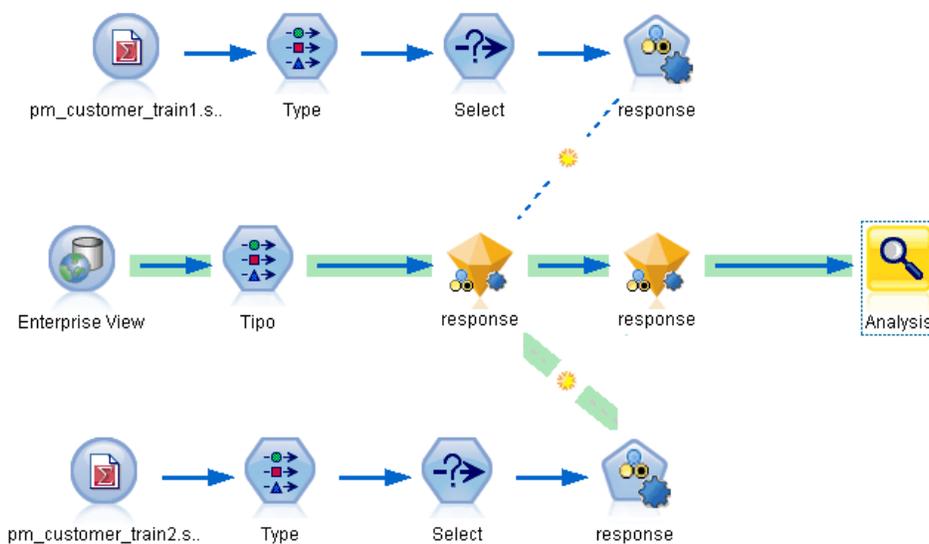
Per chiarire, si supponga di avere il seguente stream.

Figura 9-29
Ramo di calcolo del punteggio con più di un modello nello stream



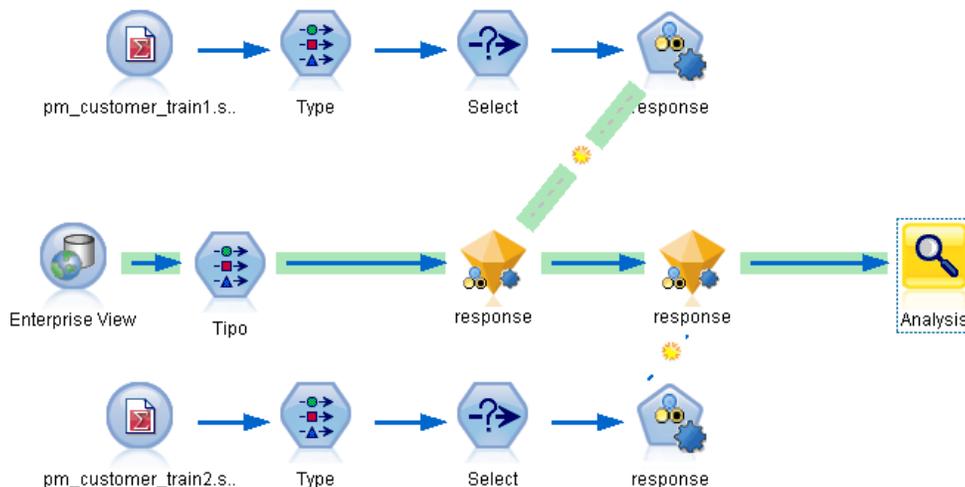
Si fa clic con il tasto destro del mouse sul nodo Analisi e si utilizza il relativo menu per impostare il ramo di calcolo del punteggio, che ora viene evidenziato. In questo modo, inoltre si designa il modello più vicino al nodo Analisi come modello di aggiornamento, come indicato dal collegamento di aggiornamento evidenziato.

Figura 9-30
Ramo di calcolo del punteggio evidenziato con più modelli e collegamenti di aggiornamento



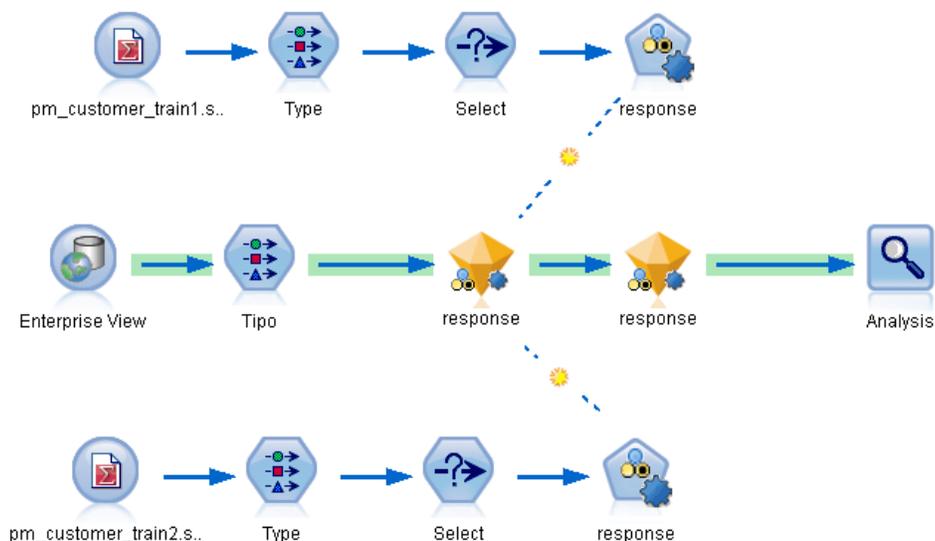
Tuttavia, si decide di utilizzare come modello di aggiornamento l'altro insieme di modelli nello stream, pertanto dal menu si imposta il relativo collegamento del modello in modo che venga utilizzato come collegamento di aggiornamento:

Figura 9-31
Ramo di calcolo del punteggio con collegamento di aggiornamento passato al primo insieme di modelli



Se successivamente si deseleggono entrambi i collegamenti del modello come collegamenti di aggiornamento, viene evidenziato solo il ramo di calcolo del punteggio, non i collegamenti. Il tipo di deployment è impostato su Solo calcolo del punteggio.

Figura 9-32
Ramo di calcolo del punteggio con più modelli e nessun collegamento di aggiornamento



Nota: è possibile scegliere di impostare uno dei collegamenti sullo stato Sostituisci, ma non l'altro. In questo caso, l'insieme di modelli scelto come modello di aggiornamento è quello che presenta un collegamento di aggiornamento e che si trova più vicino al nodo terminale quando si designa il ramo di calcolo del punteggio.

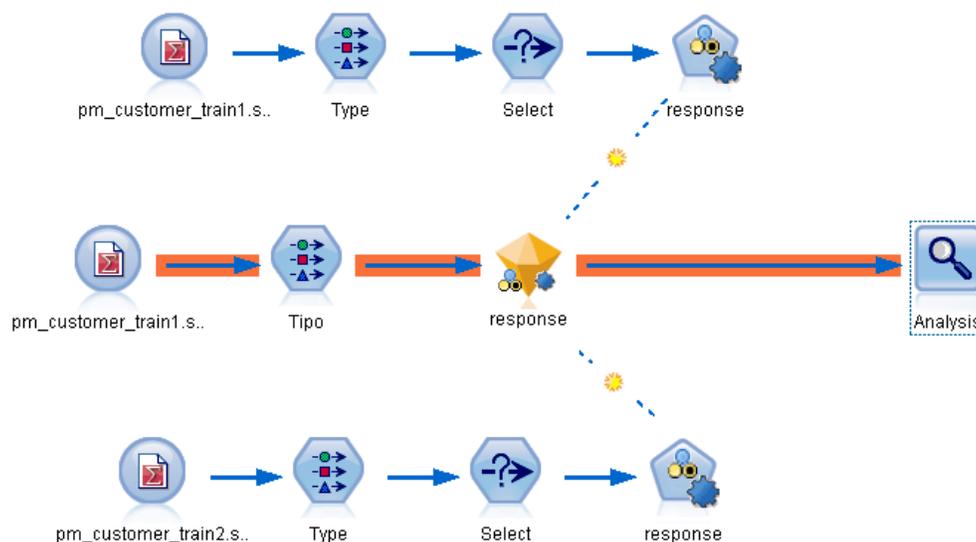
Nessun modello nello stream

Se non ci sono modelli nello stream, oppure se ci sono solo modelli senza collegamenti, il tipo di deployment è impostato su Solo calcolo del punteggio.

Controllo di un ramo di calcolo del punteggio per verificare la presenza di errori

Quando si designa il ramo di calcolo del punteggio, viene eseguito un controllo alla ricerca di eventuali errori, per esempio la mancanza di un nodo Enterprise View quando si esegue il deployment dello stream come scenario.

Figura 9-33
Ramo di calcolo del punteggio con errori



Se viene trovato un errore, il ramo di calcolo del punteggio viene evidenziato con il colore riservato agli errori e viene visualizzato un messaggio. È possibile impostare il colore degli errori utilizzando un'opzione Colore personalizzato. [Per ulteriori informazioni, vedere l'argomento Impostazione delle opzioni di visualizzazione in il capitolo 12 a pag. 231.](#)

Se viene trovato un errore, procedere come segue:

- Correggere l'errore in base al contenuto del messaggio di errore.

- ▶ Eseguire una delle seguenti operazioni:
 - Fare clic con il pulsante destro del mouse sul nodo terminale e fare clic su Controlla scenario nel menu di scelta rapida.
 - Nel menu principale, fare clic su:
Strumenti > Proprietà stream > Deployment
e fare clic sul pulsante Controllo.

- ▶ Se necessario, ripetere questo processo finché non viene più rilevato alcun errore.

Esportazione in applicazioni esterne

Informazioni sull'esportazione in applicazioni esterne

IBM® SPSS® Modeler consente di esportare l'intero processo di data mining in applicazioni esterne mediante diversi meccanismi, in modo che il lavoro svolto per la preparazione dei dati e la creazione dei modelli possa essere sfruttato anche fuori dall'ambiente SPSS Modeler.

La precedente sezione ha illustrato come eseguire il deployment degli stream in un repository IBM SPSS Collaboration and Deployment Services per sfruttarne la possibilità di accesso da parte di più utenti, di pianificazione dei processi e altre funzioni. In modo analogo è possibile utilizzare gli stream di SPSS Modeler con:

- IBM® SPSS® Modeler Advantage
- Applicazioni predittive 5.0
- Applicazioni predittive 4.x
- Applicazioni in grado di importare ed esportare file in formato PMML

Per ulteriori informazioni sull'utilizzo degli stream con IBM SPSS Modeler Advantage, vedere *Apertura di uno stream in IBM SPSS Modeler Advantage* a pag. 197.

Per esportare uno stream da utilizzare con Applicazioni predittive 5.0, seguire le istruzioni per il deployment come scenario. [Per ulteriori informazioni, vedere l'argomento *Deployment di stream in il capitolo 9* a pag. 184.](#)

Per informazioni sull'assemblaggio degli stream per il deployment con le versioni 4.x di Applicazioni predittive, vedere *Creazione guidata applicazioni predittive 4.x* a pag. 198.

Per informazioni sull'esportazione e importazione di modelli come file PMML per avere la possibilità di condividerli con qualsiasi altra applicazione in grado di supportare tale formato, vedere *Importazione ed esportazione di modelli come PMML* a pag. 206.

Nota: Applicazioni predittive è stato sostituito da IBM® SPSS® Decision Management. Il supporto per Applicazioni predittive sarà eliminato in una futura versione di SPSS Modeler.

Apertura di uno stream in IBM SPSS Modeler Advantage

Gli stream di IBM® SPSS® Modeler si possono utilizzare con l'applicazione thin client IBM® SPSS® Modeler Advantage. Benché sia possibile creare applicazioni personalizzate interamente all'interno di IBM SPSS Modeler Advantage, è possibile anche utilizzare uno stream già creato in SPSS Modeler come base per il flusso di lavoro di un'applicazione.

Per aprire uno stream IBM SPSS Modeler Advantage:

- ▶ Eseguire il deployment dello stream nel repository IBM® SPSS® Collaboration and Deployment Services facendo clic sull'opzione Implementa come stream. [Per ulteriori informazioni, vedere l'argomento Deployment di stream in il capitolo 9 a pag. 184.](#)
- ▶ Fare clic sul pulsante Apri in IBM SPSS Modeler Advantage nella barra degli strumenti o, nel menu principale, fare clic su:
File > Apri in IBM SPSS Modeler Advantage.
- ▶ Specificare le impostazioni relative alla connessione al repository, se necessario. [Per ulteriori informazioni, vedere l'argomento Connessione a IBM SPSS Collaboration and Deployment Services Repository in il capitolo 9 a pag. 159.](#) Per informazioni specifiche sulla porta, la password e altri dettagli per la connessione, rivolgersi all'amministratore del sistema locale.

Nota: il software IBM SPSS Modeler Advantage deve essere installato anche sul server del repository.

- ▶ Nel repository: nella finestra di dialogo Archivia, scegliere la cartella in cui archiviare l'oggetto, specificare le eventuali altre informazioni da registrare e fare clic sul pulsante Archivia. [Per ulteriori informazioni, vedere l'argomento Impostazione delle proprietà degli oggetti in il capitolo 9 a pag. 163.](#)

In questo modo viene avviato IBM SPSS Modeler Advantage con lo stream già aperto. Lo stream viene chiuso in SPSS Modeler.

Creazione guidata applicazioni predittive 4.x

Nota: Applicazioni predittive è stato sostituito da IBM® SPSS® Decision Management. Il supporto per Applicazioni predittive sarà eliminato in una futura versione di IBM® SPSS® Modeler.

Creazione guidata applicazioni predittive 4.x consente di assemblare gli stream di SPSS Modeler per il deployment tramite le applicazioni predittive di versione 4.x, compresa PredictiveMarketing 2.0 e versioni successive. Tutte le operazioni di manipolazione e modellazione dei dati eseguite in SPSS Modeler possono essere assemblate e salvate rapidamente come soluzione per il calcolo del punteggio. Dopo avere implementato l'applicazione, è possibile utilizzarla per incorporare i modelli di SPSS Modeler nella soluzione di campagna desiderata.

Per esempio, un'analista di dati desidera segmentare i propri clienti e sviluppare un modello di valore specializzato per l'argento, l'oro e il platino da utilizzare con PredictiveMarketing o un'altra applicazione. L'analista svilupperà innanzitutto il modello in SPSS Modeler e quindi lo esporterà per utilizzarlo in PredictiveMarketing, che utilizzerà il modello di SPSS Modeler in background per suggerire un'offerta con il valore più alto.

Per accedere alla procedura guidata

- ▶ Nel menu principale, fare clic su:
Strumenti > Creazione guidata applicazioni predittive 4.x

Operazioni preliminari all'utilizzo di Creazione guidata applicazioni predittive

Le informazioni seguenti forniscono una panoramica dell'integrazione tra IBM® SPSS® Modeler e altre applicazioni predittive. Prima di esportare uno stream mediante la procedura guidata, è consigliabile verificare l'integrazione e i prerequisiti per la pubblicazione.

Funzionamento dell'integrazione

In circostanze normali, è possibile eseguire le seguenti procedure di SPSS Modeler per espandere le funzionalità di data mining e deployment disponibili con le applicazioni predittive.

- ▶ Iniziare nell'applicazione predittiva in uso. Mediante le opzioni disponibili in Customer View Builder, esportare il modello di dati UCV (Unified Customer View) come file XML. Annotare la posizione del file XML poiché risulterà utile per le procedure da eseguire in SPSS Modeler.
- ▶ Impostare quindi i nodi di input in SPSS Modeler in modo che accedano a tutte le sorgenti dati (database, file piatti, ecc.) che contengono i campi del modello di dati UCV, elencati nel file XML esportato in precedenza. È possibile includere tutti i campi ai quali viene fatto riferimento nel modello di dati UCV oppure utilizzare unicamente la parte del modello UCV necessaria per i modelli che si desidera creare. In genere in SPSS Modeler si utilizzano numerosi nodi di input per accedere ai dati di modellazione.
- ▶ Utilizzare SPSS Modeler per eseguire le unioni di dati, le trasformazioni o le derivazioni necessarie per le operazioni di data mining.
- ▶ In un punto dello stream, assicurarsi di inserire un nodo Tipo e di denominarlo *UCV*. Questo nodo Tipo denominato *UCV* non solo viene utilizzato per la direzionalità durante la modellazione, ma anche per garantire che i dati corrispondano alle informazioni dei campi definite nel file XML. È consigliabile confrontare le impostazioni del nodo Tipo con le specifiche degli attributi del file XML creato in precedenza. [Per ulteriori informazioni, vedere l'argomento Passo 3: Selezione di un nodo UCV a pag. 202.](#)
- ▶ Considerare quindi il tipo di modello che si desidera creare in SPSS Modeler. Se si esegue il deployment di un modello di valore quale una rete neurale, è possibile esportare le previsioni binarie (per esempio churn vero/falso) come propensione, in modo tale da rendere la previsione confrontabile con le previsioni dei modelli generati dall'applicazione. [Per ulteriori informazioni, vedere l'argomento Esportazione di previsioni binarie come punteggi di propensione a pag. 200.](#)
- ▶ Dopo avere completato il modello e avere convertito le confidenze in propensioni, aggiungere un nodo terminale al ramo di deployment dello stream. Viene spesso utilizzato un nodo Tabella, ma un nodo terminale sarà comunque sufficiente. Assicurarsi che nel nodo terminale siano visibili unicamente i campi che si desidera siano visibili nell'applicazione esterna. In altre parole, potrebbe essere necessario filtrare i campi che non si desidera implementare prima del nodo terminale.
- ▶ Assicurarsi inoltre che i campi di previsione generati dal modello vengano istanziati prima dell'esportazione. A tale scopo, è possibile aggiungere un nodo Tipo tra il modello generato e il nodo "di esportazione" terminale.
- ▶ Come passaggio finale prima di utilizzare la procedura guidata, assicurarsi che lo stream sia pronto per il deployment mediante un'esecuzione di prova.

Ora lo stream è pronto per il deployment. Per accedere alla Creazione guidata applicazioni predittive, utilizzare il menu Strumenti di SPSS Modeler. Eseguire i passaggi della procedura guidata illustrati in questo documento per creare un pacchetto Deployment di Modeler (.cdp) contenente le informazioni e i metadati dello stream necessari per la pubblicazione nell'ambiente in tempo reale.

Esportazione di previsioni binarie come punteggi di propensione

Prima di procedere all'esportazione con Creazione guidata applicazioni predittive, verificare che il punteggio dell'output del modello di IBM® SPSS® Modeler sia coerente con le previsioni generate dall'applicazione predittiva. In molti casi può essere opportuno esportare le previsioni binarie come punteggi di propensione per poter confrontare con precisione l'attendibilità delle previsioni in numerosi modelli.

- Un **modello di valore** crea un singolo **punteggio di propensione** per ciascun record che classifica la probabilità di uno specifico risultato *sì* o *no* su una scala da 0.0 a 1.0. Per esempio, un modello del tasso di abbandono potrebbe generare un punteggio compreso tra 1,0 (abbandono probabile) e 0,0 (abbandono improbabile). Poiché i punteggi di propensione non sono probabilità, un punteggio di 0.5 non significa necessariamente “il 50% di probabilità di abbandono”, né il doppio di probabilità di abbandono rispetto a un soggetto con un punteggio di 0.25; significa, tuttavia, che il soggetto presenta *più probabilità* di abbandono rispetto a un altro con un punteggio di 0.4. I punteggi di propensione possono essere utilizzati per classificare e individuare, per esempio, il 10% dei clienti con le maggiori probabilità di abbandono.
- Un **modello di offerta** crea valori di previsione e confidenza che verranno utilizzati come attributi virtuali nel modello di dati UCV dopo che sono stati inviati all'applicazione. Questi valori possono essere intervalli numerici o continui.

Numerosi modelli di SPSS Modeler possono generare previsioni binarie, compresi i modelli rete neurale, albero decisionale e regressione logistica. Se si esegue il deployment di un modello di valore, prima di eseguire l'esportazione può essere necessario esportare i punteggi di propensione. I punteggi di propensione possono essere attivati nella scheda Analisi del nodo Modelli o nella scheda Impostazioni dell'insieme di modelli generato. [Per ulteriori informazioni, vedere l'argomento Opzioni della scheda Analizza di un nodo Modelli in il capitolo 3 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)

Passo 1: Panoramica

Quando si apre per la prima volta la Creazione guidata applicazioni predittive, viene visualizzata una schermata iniziale che indirizza l'utente al processo di creazione del pacchetto dei componenti stream necessari.

Figura 10-1
Creazione guidata applicazioni predittive

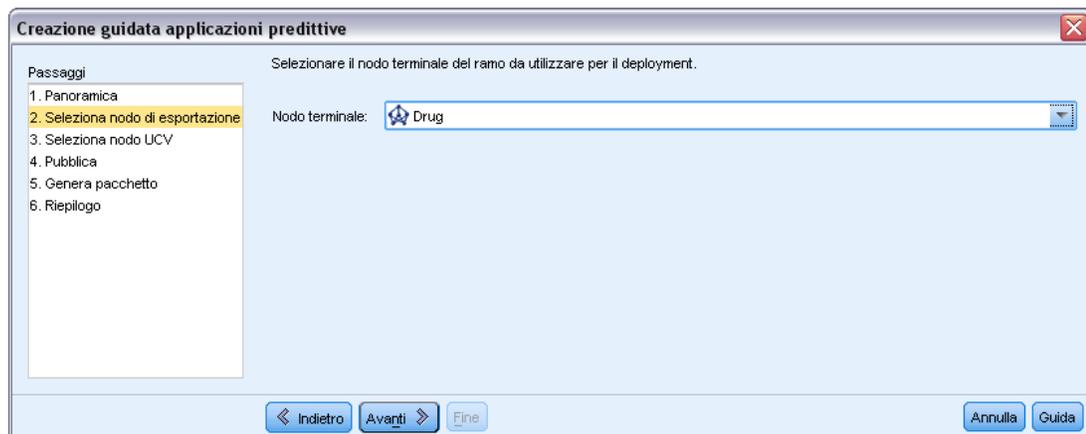


Le altre pagine della procedura guidano l'utente nel processo di creazione di un pacchetto per il deployment nell'ambiente in tempo reale. Prima di continuare, usare l'elenco di controllo dei prerequisiti per garantire che lo stream sia stato preparato per il deployment. [Per ulteriori informazioni, vedere l'argomento Operazioni preliminari all'utilizzo di Creazione guidata applicazioni predittive a pag. 199.](#)

Passo 2: Selezione di un nodo terminale

Nel passaggio 2, è possibile specificare un nodo terminale dello stream che rappresenta il ramo di punteggio o la parte dello stream dove viene eseguito il calcolo del punteggio.

Figura 10-2
Selezione di un nodo terminale



È importante saper distinguere tra nodi terminali con lo stesso nome, poiché l'elenco presente nella procedura guidata include solo il nome e il tipo di ogni nodo terminale dello stream. Per evitare problemi, assegnare nomi univoci ai nodi terminali dello stream.

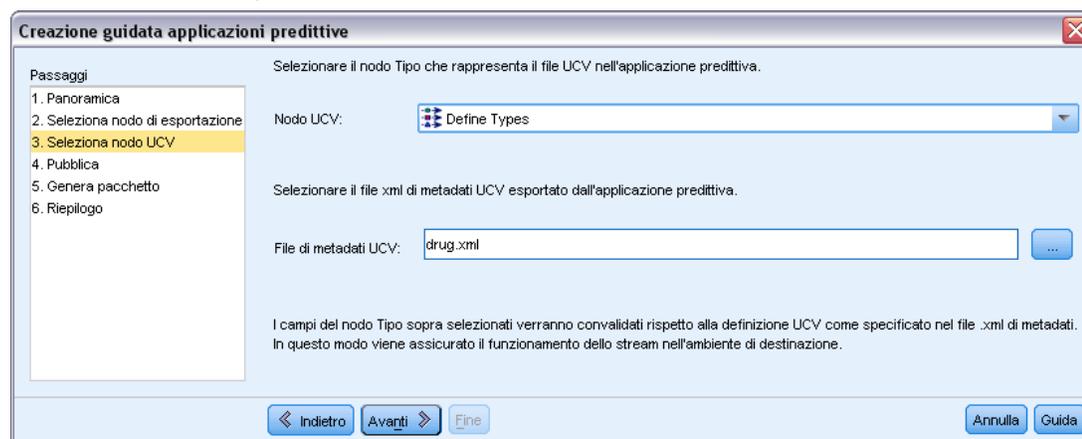
Assicurarsi inoltre che nel nodo terminale siano visibili unicamente i campi che si desidera rendere visibili nell'ambiente dell'applicazione. In altre parole, potrebbe essere necessario filtrare i campi che non si desidera implementare prima del nodo terminale.

Passo 3: Selezione di un nodo UCV

Nel passaggio 3 è possibile specificare due informazioni fondamentali per la pubblicazione, ovvero un nodo UCV e un file di metadati UCV.

Figura 10-3

Selezione di un nodo Tipo utilizzato come nodo UCV e file di metadati UCV (file XML)



Nodo UCV. Un nodo UCV è un nodo Tipo dello stream che consente di garantire la corrispondenza di tutti i dati con le definizioni presenti nel modello di dati UCV. Queste specifiche sono archiviate in un file XML, esportato in precedenza da Customer View Builder. Quando si fa clic su Avanti, le impostazioni del nodo Tipo vengono convalidate automaticamente in base alle specifiche del file XML indicato in questa schermata mediante l'opzione File di metadati UCV.

File di metadati UCV. Un file di metadati UCV è il file XML creato in precedenza mediante la funzionalità Customer View Builder. Il file XML selezionato contiene gli attributi dei dati richiesti per il deployment nell'applicazione.

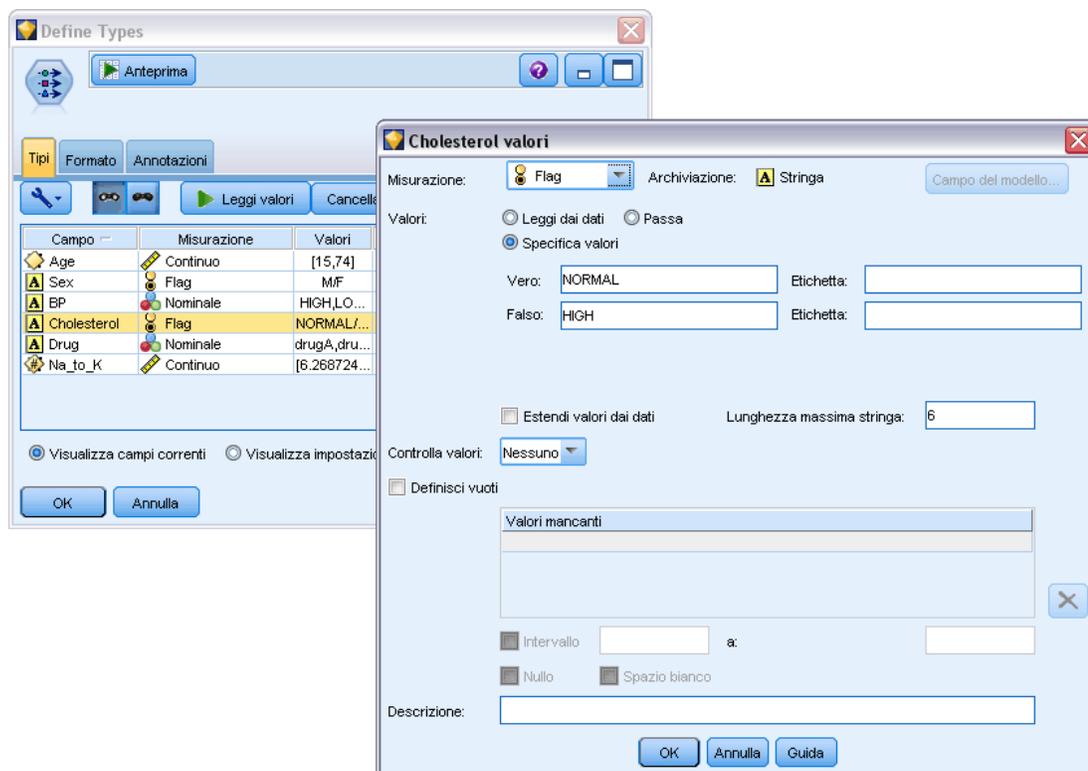
Errori di discordanza dei dati

Se la procedura guidata ha generato errori di discordanza dei dati, tornare allo stream di IBM® SPSS® Modeler e controllare le specifiche del nodo Tipo. Confrontare i dati dei campi del nodo Tipo con quelli contenuti nel file XML creato da Customer View Builder. Per aprire il file XML è possibile utilizzare un browser di testo, per esempio il Blocco note. *Verificare se le impostazioni del nodo Tipo corrispondono agli attributi UCV.* Per esempio, è possibile che nel file XML venga stabilito che il campo *Colesterolo* è richiesto da UCV e che tale campo contiene valori stringa:

```
<UcvAttribute Name="Colesterolo" Domain="String"/>
```

Nello stream di SPSS Modeler, controllare le impostazioni del nodo Tipo per verificare che vengano esportati da SPSS Modeler unicamente i campi utili per l'ambiente in tempo reale.

Figura 10-4
Controllo delle impostazioni del nodo Tipo e archiviazione nella finestra secondaria Valori



Mappatura della classe di archiviazione su un dominio

In alcuni casi, la mappatura delle impostazioni del nodo Tipo sulle specifiche UCV non risulta di comprensione immediata. In questa integrazione, il **dominio** come è definito nel file XML di UCV equivale alla **classe di archiviazione** di SPSS Modeler. La corrispondenza tra le impostazioni dei campi di ognuno di essi sarà realizzata durante l'esportazione.

Tabella 10-1
Mappatura di dominio e archiviazione

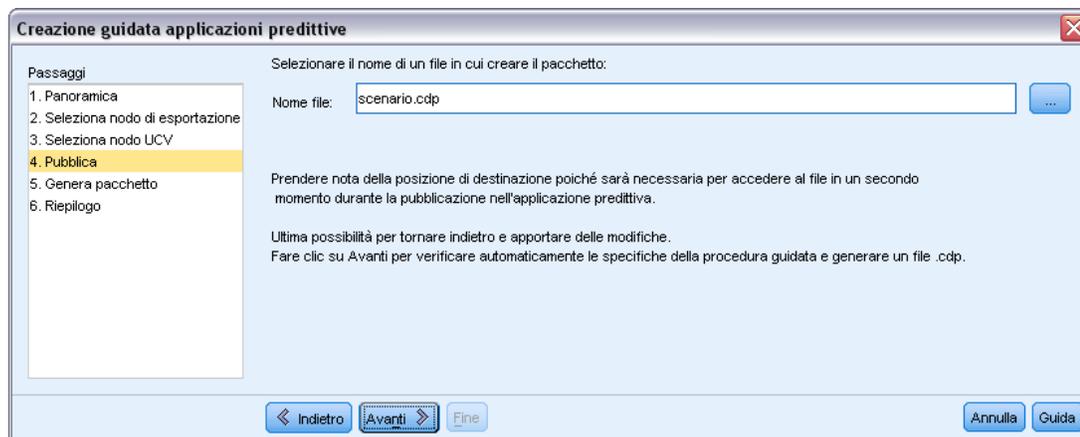
SPSS Modeler Storage	Domini delle applicazioni
Stringa	Stringa Carattere Bit
Reale	Reale singola precisione Double Decimali
Intero	Long Intero
Data	Data
Timestamp	Timestamp
Ora	Nessuna corrispondenza

È possibile modificare il tipo di archiviazione mediante le funzioni di conversione, per esempio `to_integer`, in un nodo Nuovo campo o Riempimento. [Per ulteriori informazioni, vedere l'argomento Conversione della classe di archiviazione mediante il nodo Riempimento in il capitolo 4 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)

Passo 4: Specifica di un pacchetto

Nel passaggio 4, è necessario specificare il nome e la posizione in cui verrà salvato il pacchetto dello stream. Questo passaggio consente di salvare una serie di oggetti come un **pacchetto Deployment di Modeler**, al quale è assegnata l'estensione `.cdp`.

Figura 10-5
Specifica del nome di un pacchetto

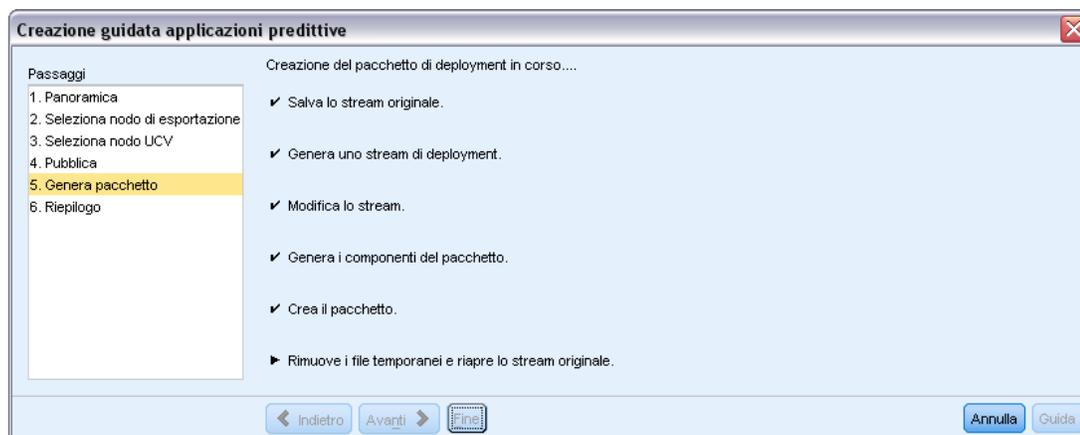


Fare clic su Avanti per controllare automaticamente i metadati dello stream e le specifiche. Se tutte le impostazioni sono corrette, verrà generato il file `.cdp`. Annotare la posizione di destinazione poiché risulterà necessaria per accedere al file in un momento successivo durante la pubblicazione nell'applicazione.

Passo 5: Creazione del pacchetto

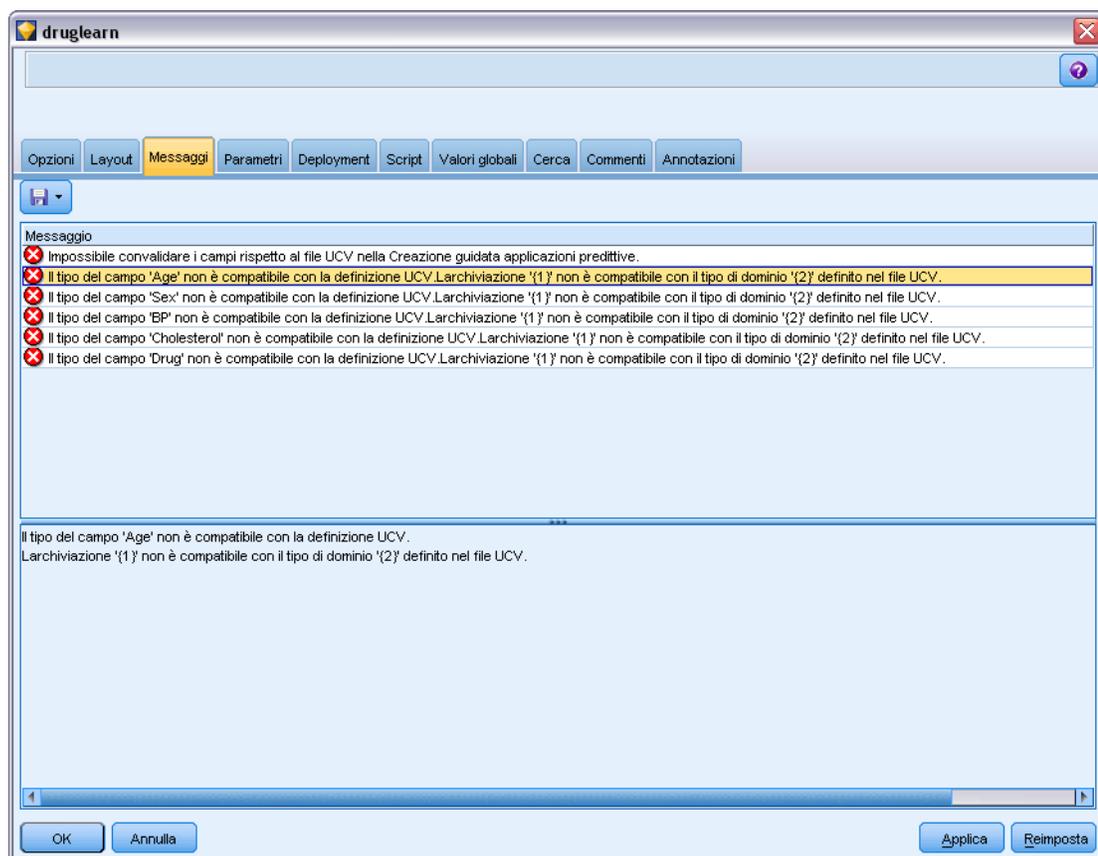
In questa fase vengono controllate automaticamente le informazioni dello stream e le specifiche inserite nelle schermate precedenti.

Figura 10-6
Creazione del pacchetto di deployment



I nomi e i tipi di campo dello stream di IBM® SPSS® Modeler vengono confrontati con quelli presenti nel file XML creato dal modello UCV. Se le informazioni dei campi non corrispondono, è possibile che la procedura guidata si interrompa e che vengano visualizzati messaggi di errore.

Figura 10-7
Messaggi di errore generati dalla discordanza dei metadati



Se si verifica un errore, tornare allo stream di SPSS Modeler ed eseguire le verifiche seguenti:

- Controllare che tutti i nomi di campo presenti nel nodo UCV (un nodo Tipo nello stream) siano presenti nel file XML che definisce il modello UCV e inoltre che i relativi tipi (domini) siano compatibili.
- Tenere presente che i nomi di campo sia di SPSS Modeler che di altre applicazioni rilevano la distinzione tra maiuscole e minuscole. L'ordine dei campi non è invece importante.

Nota: la dimensione del pacchetto *.cdp* creato deve essere inferiore a 5 KB. Se l'elenco dei campi creati è molto lungo, potrebbe essere necessario rimuoverne alcuni dalla descrizione in modo da rispettare tale limite.

Passo 6: Riepilogo

Dopo avere creato il pacchetto Deployment di IBM® SPSS® Modeler, il lavoro in SPSS Modeler è terminato.

Passaggio successivo

È quindi possibile importare il modello (un file di pacchetto *.cdp*) in Interaction Builder.

Importazione ed esportazione di modelli come PMML

PMML, acronimo di Predictive Model Markup Language, è un formato XML per descrivere modelli statistici e di data mining, inclusi input ai modelli, trasformazioni utilizzate per la preparazione dei dati per il data mining e parametri che definiscono i modelli stessi. In IBM® SPSS® Modeler è possibile importare ed esportare PMML, rendendo possibile la condivisione di modelli con altre applicazioni che supportano questo formato, quali IBM® SPSS® Statistics.

Nota: per l'esportazione di PMML è necessario disporre di una licenza di IBM® SPSS® Modeler Server.

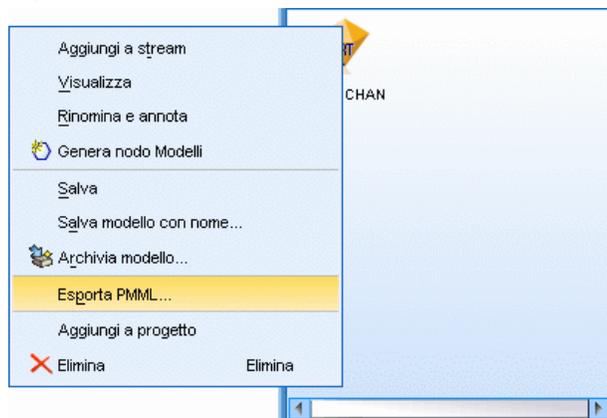
Per ulteriori informazioni sul formato PMML, vedere il sito Web del gruppo di data mining all'indirizzo <http://www.dmg.org>.

Per esportare un modello:

L'esportazione in formato PMML è supportata per la maggior parte dei tipi di modelli generati in SPSS Modeler. [Per ulteriori informazioni, vedere l'argomento Tipi di modello che supportano il formato PMML a pag. 209.](#)

- ▶ Fare clic con il pulsante destro del mouse su un insieme di modelli nella palette Modelli. In alternativa, fare doppio clic su un insieme di modelli nell'area di disegno e selezionare il menu File.
- ▶ Nel menu, fare clic su Esporta PMML.

Figura 10-8
Esportazione di un modello in formato PMML



- Nella finestra di dialogo Esporta (o Salva), specificare una directory di destinazione e un nome univoco per il modello.

Nota: è possibile modificare opzioni per l'esportazione in formato PMML nella finestra di dialogo Opzioni utente. Nel menu principale, fare clic su Strumenti > Opzioni > Opzioni utente

e selezionare la scheda PMML.

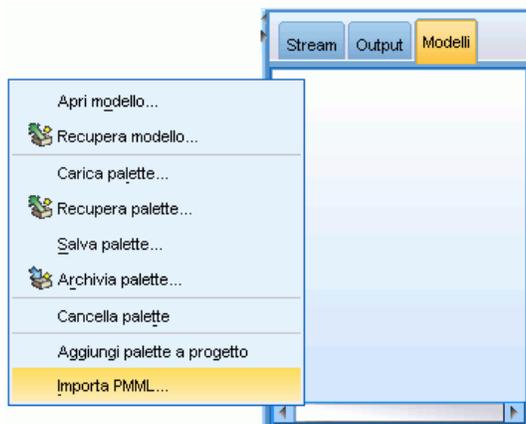
Per ulteriori informazioni, vedere l'argomento [Impostazione delle opzioni di esportazione PMML](#) in il capitolo 12 a pag. 236.

Per importare un modello salvato come PMML

I modelli esportati come PMML da SPSS Modeler o un'altra applicazione possono essere importati nella palette Modelli. [Per ulteriori informazioni, vedere l'argomento Tipi di modello che supportano il formato PMML a pag. 209.](#)

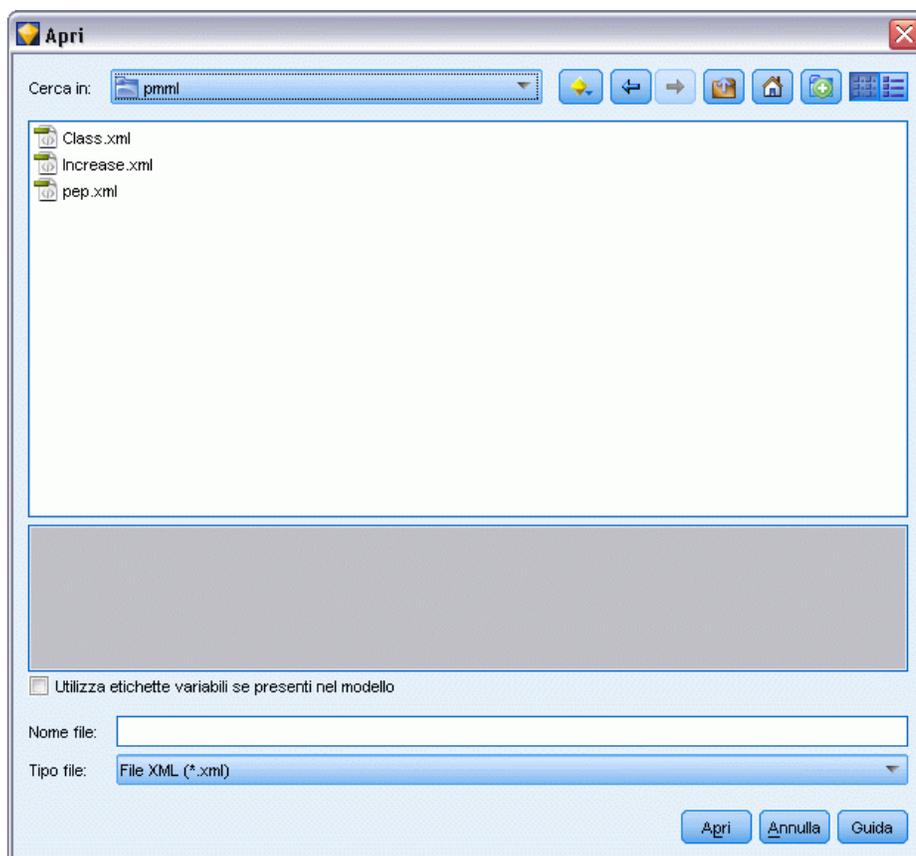
- Nella palette Modelli, fare clic con il pulsante destro del mouse sulla palette e selezionare Importa PMML dal menu.

Figura 10-9
Importazione di un modello in formato PMML



- ▶ Selezionare il file da importare e specificare le opzioni per le etichette delle variabili in base alle proprie esigenze.
- ▶ Fare clic su Apri.

Figura 10-10
Selezione del file XML per un modello salvato tramite PMML



Utilizza etichette variabili se presenti nel modello. Nel formato PMML è possibile specificare nomi ed etichette di variabili (per esempio ID referente per *RefID*) per le variabili contenute nel dizionario dei dati. Selezionare questa opzione per utilizzare le etichette delle variabili, se presenti nel PMML esportato.

Se sono state selezionate le opzioni delle etichette di variabile ma non sono presenti etichette di variabile nel PMML, i nomi delle variabili verranno utilizzati normalmente.

Tipi di modello che supportano il formato PMML

Esportazione PMML

SPSS Modeler modelli. I seguenti modelli creati in IBM® SPSS® Modeler possono essere esportati come PMML 4.0:

- C&R Tree
- QUEST
- CHAID
- Regressione lineare
- Rete neurale
- C5.0
- Regressione logistica
- Genlin
- SVM
- Rete di Bayes
- Apriori
- Carma
- Sequenza
- K-Means
- Kohonen
- TwoStep
- KNN
- Modelli Statistics

Il seguente modello creato in SPSS Modeler può essere esportato come PMML 3.2:

- Elenco decisionale

Modelli nativi di database. Per i modelli generati utilizzando algoritmi nativi di database, l'esportazione PMML è disponibile solo per i modelli IBM InfoSphere Warehouse. I modelli creati utilizzando Analysis Services di Microsoft o Oracle Data Miner non possono essere esportati. Si noti inoltre che i modelli IBM esportati come PMML non possono essere reimportati in SPSS Modeler. [Per ulteriori informazioni, vedere l'argomento Panoramica sulla modellazione di database in il capitolo 2 in IBM SPSS Modeler 14.2 Guida alla modellazione in-database.](#)

Importazione PMML

SPSS Modeler può importare e calcolare il punteggio di modelli PMML generati dalle versioni correnti di tutti i prodotti IBM® SPSS® Statistics, compresi i modelli esportati da SPSS Modeler nonché i modelli o le trasformazioni PMML generate da SPSS Statistics 17.0 o versioni successive. In sostanza, sono importabili tutti i modelli PMML di cui il modulo di gestione del punteggio è in grado di calcolare il punteggio, con le seguenti eccezioni:

- Non è possibile importare modelli Apriori, CARMA e Rilevamento anomalie.
- Non è possibile sfogliare i modelli PMML dopo averli importati in SPSS Modeler, anche se è possibile utilizzarli per il calcolo del punteggio. Questa eccezione riguarda anche i modelli che sono stati originariamente esportati da SPSS Modeler. Per evitare tale limitazione, esportare il modello come file di modello generato (*.gm) anziché come PMML.
- I modelli IBM InfoSphere Warehouse esportati come PMML non possono essere importati.
- Si è verificata una validazione limitata all'importazione, ma la validazione completa viene eseguita al tentativo di calcolare il punteggio del modello. Pertanto è possibile che l'importazione riesca e che, invece, il calcolo del punteggio fallisca o produca risultati non corretti.

Progetti e report

Introduzione ai progetti

Un **progetto** è un gruppo di file correlati a un'attività di data mining e include gli stream di dati, i grafici, i modelli generati, i report e qualsiasi altro elemento creato in IBM® SPSS® Modeler. I progetti di SPSS Modeler non forniscono semplicemente un metodo di organizzazione dell'output, ma anche numerose altre funzionalità. Grazie ai progetti, è possibile:

- Annotare ogni oggetto nel file di progetto.
- Utilizzare il modello CRISP-DM per l'elaborazione del data mining in base alle specifiche esigenze aziendali. I progetti includono anche una Guida in linea di CRISP-DM che fornisce informazioni dettagliate ed esempi pratici relativi al data mining con CRISP-DM.
- Aggiungere al progetto oggetti esterni a SPSS Modeler, per esempio una presentazione di PowerPoint che illustra gli obiettivi del data mining o white paper relativi agli algoritmi che si prevede di utilizzare.
- Creare report completi o semplici report di aggiornamento basati sulle proprie annotazioni. I report possono essere creati in formato HTML, in modo da poter essere facilmente pubblicati nella rete Intranet dell'organizzazione.

Nota: se il riquadro dei progetti non è visibile nella finestra di SPSS Modeler, fare clic su Progetto nel menu Visualizza.

Gli oggetti aggiunti a un progetto possono essere visualizzati in due modi: mediante la **scheda Classi** o la **scheda CRISP-DM**. Tutti gli elementi aggiunti a un progetto vengono inseriti in entrambe le visualizzazioni ed è possibile passare da una visualizzazione all'altra per ottenere l'organizzazione ottimale.

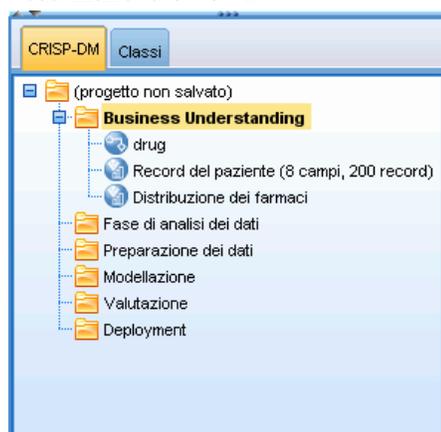
Figura 11-1
Visualizzazione CRISP-DM e Classi di un file di progetto



Visualizzazione CRISP-DM

I progetti di IBM® SPSS® Modeler supportano lo standard CRISP-DM (Cross-Industry Standard Process for Data Mining) e consentono pertanto di organizzare l'elaborazione del data mining aziendale mediante un metodo consolidato del settore e non proprietario. Il modello CRISP-DM prevede sei fasi dell'elaborazione, dalla fase iniziale (raccolta dei requisiti aziendali) a quella finale (deployment dei risultati). Alcune fasi non prevedono in genere l'utilizzo di SPSS Modeler, tuttavia il riquadro dei progetti le include tutte e sei allo scopo di fornire una posizione centralizzata per l'archiviazione e la registrazione di tutti i materiali associati al progetto. Per esempio, la fase Business Understanding è in genere incentrata sulla raccolta dei requisiti aziendali e sulla definizione degli obiettivi insieme ai colleghi, anziché sull'utilizzo dei dati in SPSS Modeler. Il riquadro dei progetti consente di archiviare le note delle riunioni nella cartella *Business Understanding*, in modo da poterle riutilizzare e inserire nei report in momenti successivi.

Figura 11-2
visualizzazione CRISP-DM



Nella visualizzazione di CRISP-DM del riquadro dei progetti è disponibile una Guida in linea che fornisce informazioni sull'intero ciclo di vita del data mining. Da SPSS Modeler è possibile accedere a questa Guida in linea facendo clic su Guida in linea di CRISP-DM nel menu Aiuto.

Nota: se il riquadro dei progetti non è visibile nella finestra, fare clic su Progetto nel menu Visualizza.

Impostazione della fase di default del progetto

Gli oggetti aggiunti a un progetto vengono inseriti in una fase di default di CRISP-DM. Ciò implica che è necessario organizzare manualmente gli oggetti in base alla fase del data mining in cui vengono utilizzati. È consigliabile impostare come cartella di default la fase di lavoro corrente.

Per selezionare la fase di default da utilizzare:

- ▶ Nella visualizzazione CRISP-DM, fare clic con il pulsante destro del mouse sulla cartella corrispondente alla fase di default che si desidera impostare.
- ▶ Fare clic su Imposta come default nel menu.

La cartella di default verrà visualizzata in grassetto.

Visualizzazione Classi

Nella visualizzazione Classi del riquadro dei progetti, le operazioni di IBM® SPSS® Modeler sono organizzate per categorie in base al tipo di oggetti creati. È possibile inserire gli oggetti salvati nelle seguenti categorie:

- Stream
- Nodi
- Modelli
- Tabelle, grafici, report
- Altro (file non di SPSS Modeler, per esempio presentazioni o white paper relativi ai processi di data mining)

Figura 11-3
visualizzazione Classi



Gli oggetti aggiunti nella visualizzazione Classi vengono inseriti anche nella cartella della fase di default della visualizzazione CRISP-DM.

Nota: se il riquadro dei progetti non è visibile nella finestra, fare clic su Progetto nel menu Visualizza.

Creazione di un progetto

Un progetto non è altro che un file contenente i riferimenti a tutti i file associati al progetto. Gli elementi di un progetto vengono pertanto salvati sia singolarmente, sia come un riferimento nel file di progetto (.cpj). Questa struttura di riferimenti comporta quanto segue:

- Gli elementi di un progetto devono essere salvati singolarmente prima di essere aggiunti a un progetto. In caso contrario, verrà richiesto di salvarli prima di aggiungerli al progetto corrente.

- Gli oggetti aggiornati singolarmente, per esempio gli stream, vengono aggiornati anche nel file di progetto.
- Se si spostano o si eliminano manualmente gli oggetti (per esempio gli stream, i nodi e gli oggetti di output) dal file system, i collegamenti nel file di progetto non saranno più validi.

Creazione di un nuovo progetto

La procedura di creazione di un nuovo progetto nella finestra di IBM® SPSS® Modeler è molto semplice. È possibile iniziare subito la creazione se non è aperto alcun progetto oppure chiudere il progetto corrente e iniziare da un progetto vuoto.

- ▶ Nel menu principale, fare clic su:
File > Progetto > Nuovo progetto...

Aggiunta di oggetti a un progetto

Dopo avere creato o aperto un progetto, è possibile aggiungere oggetti quali stream di dati, nodi e report utilizzando diversi metodi.

Aggiunta di oggetti dallo strumento Manager

È possibile aggiungere stream o output mediante lo strumento Manager disponibile nella parte in alto a destra della finestra di IBM® SPSS® Modeler.

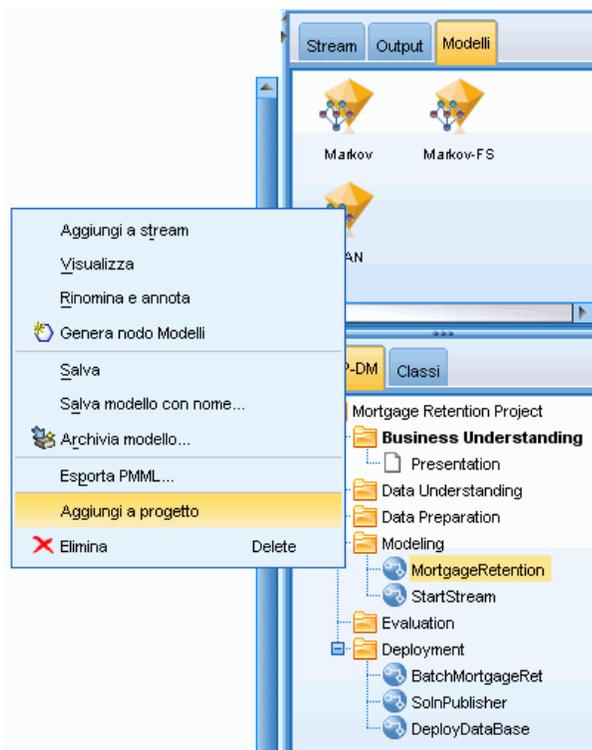
- ▶ Selezionare un oggetto, per esempio una tabella o uno stream, da una delle schede della finestra dei manager.
- ▶ Fare clic con il pulsante destro del mouse e selezionare **Aggiungi a progetto**.

Se l'oggetto è stato salvato in precedenza, verrà aggiunto automaticamente alla cartella degli oggetti appropriata (nella visualizzazione **Classi**) o alla cartella della fase di default (nella visualizzazione **CRISP-DM**).

- ▶ In alternativa, è possibile trascinare gli oggetti dai manager al riquadro dei progetti.

Nota: è possibile che venga richiesto di salvare prima l'oggetto. Durante il salvataggio, assicurarsi che l'opzione **Aggiungi file a progetto** della finestra di dialogo **Salva** sia selezionata. In questo modo, l'oggetto verrà aggiunto automaticamente al progetto dopo il salvataggio.

Figura 11-4
Aggiunta di elementi a un progetto



Aggiunta di nodi dall'area di disegno

Tramite la finestra di dialogo Salva, è possibile aggiungere i singoli nodi dall'area di disegno dello stream.

- ▶ Selezionare un nodo nell'area di disegno.
- ▶ Fare clic con il pulsante destro del mouse e selezionare Salva nodo. In alternativa, nel menu principale, fare clic su:
Modifica > Nodo > Salva nodo...
- ▶ Nella finestra di dialogo Salva selezionare Aggiungi file a progetto.
- ▶ Specificare un nome per il nodo e fare clic su Salva.

Il file verrà salvato e aggiunto al progetto. I nodi vengono aggiunti nella cartella *Nodi* della visualizzazione Classi e nella cartella della fase di default nella visualizzazione CRISP-DM.

Aggiunta di file esterni

È possibile aggiungere a un progetto un'ampia gamma di oggetti esterni a SPSS Modeler, che possono risultare utili per la gestione dell'intero processo di data mining all'interno di SPSS Modeler. Per esempio, è possibile archiviare in un progetto collegamenti a dati, note, presentazioni e grafici. Nella visualizzazione CRISP-DM è possibile aggiungere i file esterni

nella cartella desiderata, mentre nella visualizzazione Classi questi possono essere salvati solo nella cartella *Altro*.

Per aggiungere file esterni a un progetto:

- ▶ Trascinare i file dal desktop nel progetto.
 - o*
 - ▶ Con il pulsante destro del mouse, fare clic sulla cartella di destinazione nella visualizzazione CRISP-DM o Classi.
 - ▶ Fare clic su Aggiungi a cartella nel menu di scelta rapida.
 - ▶ Selezionare un file nella finestra di dialogo e fare clic su Apri.
- Nei progetti di SPSS Modeler verrà aggiunto un riferimento all'oggetto selezionato.

Trasferimento di progetti in IBM SPSS Collaboration and Deployment Services Repository

È possibile trasferire un intero progetto (compresi tutti i file dei componenti) nel IBM® SPSS® Collaboration and Deployment Services Repository in una sola operazione. Gli oggetti che sono già archiviati nella cartella di destinazione non saranno spostati. Questa funzione può effettuare anche l'operazione inversa, ovvero trasferire interi progetti dal IBM SPSS Collaboration and Deployment Services Repository al file system locale.

Nota: per l'accesso a un repository IBM® SPSS® Collaboration and Deployment Services è necessaria una licenza distinta. Per ulteriori informazioni, vedere <http://www.ibm.com/software/analytics/spss/products/deployment/cds/>

Trasferimento di un progetto

Verificare che il progetto da trasferire sia aperto nel riquadro dei progetti.

Per trasferire un progetto:

- ▶ Con il pulsante destro del mouse, fare clic sulla cartella dei progetti principale e selezionare Trasferisci progetto.
- ▶ Se richiesto, accedere a IBM SPSS Collaboration and Deployment Services Repository.
- ▶ Specificare la nuova cartella per il progetto e fare clic su OK.

Impostazione delle proprietà del progetto

È possibile personalizzare il contenuto e la documentazione di un progetto nella finestra di dialogo Proprietà progetto. Per accedere alle proprietà del progetto:

- ▶ Fare clic con il pulsante destro del mouse su un oggetto o su una cartella nel riquadro dei progetti e selezionare Proprietà progetto.

- Fare clic sulla scheda Progetto per inserire le informazioni di base del progetto.

Figura 11-5
Impostazione delle proprietà del progetto

Oggetti del progetto	Conteggi
Stream	1
Nodi	0
Supernodi	0
Modelli generati	0
Tabelle, grafici e report	2
Altro	0
TOTALE	3

Data di creazione. Mostra la data di creazione del progetto (non modificabile).

Tabella riassuntiva È possibile creare un riepilogo del progetto di data mining, che verrà visualizzato nel report del progetto.

Contenuto. Elenca il tipo e il numero dei componenti a cui viene fatto riferimento nel file di progetto (non modificabile).

Salva oggetti non salvati come. Specifica se gli oggetti non salvati devono essere salvati nel file system locale o archiviati nel repository. [Per ulteriori informazioni, vedere l'argomento Informazioni su IBM SPSS Collaboration and Deployment Services Repository in il capitolo 9 a pag. 156.](#)

Aggiorna i riferimenti a oggetti quando si carica il progetto. Selezionare questa opzione per aggiornare i riferimenti del progetto ai relativi componenti. *Nota:* i file aggiunti a un progetto non vengono salvati nel file di progetto, bensì vengono archiviati nel progetto sotto forma di riferimenti ai file e pertanto lo spostamento o l'eliminazione di un file comporterà la rimozione del relativo oggetto dal progetto.

Annotazione di un progetto

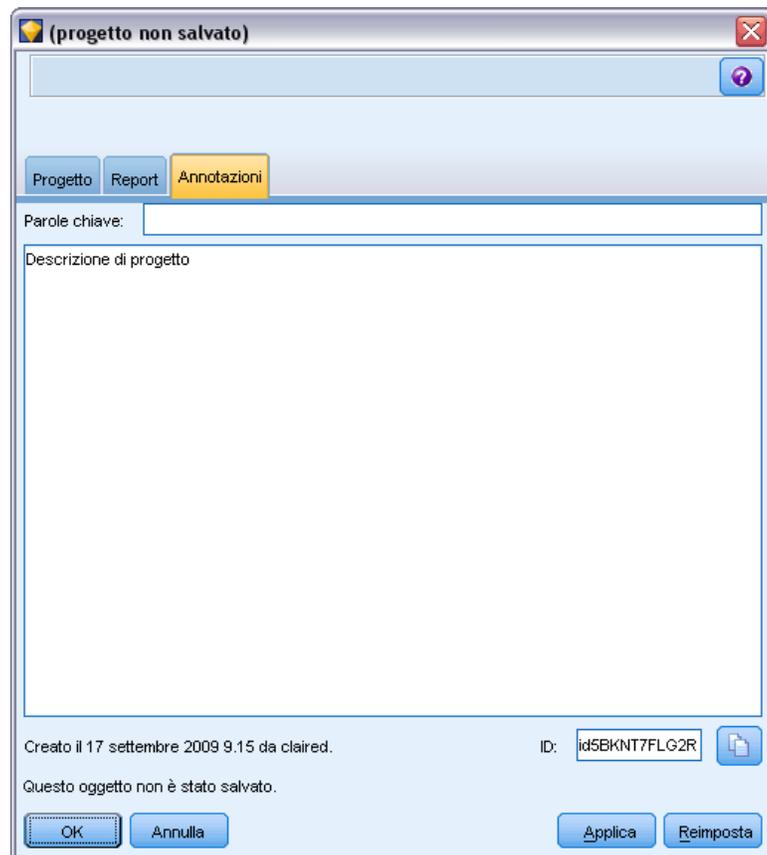
Il riquadro dei progetti offre diversi metodi per l'annotazione dei processi di data mining. Le annotazioni a livello di progetto consentono di tenere traccia degli obiettivi e delle scelte generali, mentre le annotazioni a livello di cartella o di nodo forniscono informazioni più dettagliate. Nella scheda Annotazioni è possibile inserire informazioni a livello di progetto, per esempio l'esclusione dei dati con dati mancanti irrecuperabili oppure le ipotesi formulate durante l'analisi dei dati.

Per annotare un progetto:

- ▶ Selezionare la cartella del progetto nella visualizzazione CRISP-DM o Classi.
- ▶ Fare clic sulla cartella con il pulsante destro del mouse e selezionare Proprietà progetto.
- ▶ Fare clic sulla scheda Annotazioni.

Figura 11-6

Scheda Annotazioni della finestra di dialogo Proprietà progetto



- ▶ Digitare le parole chiave e il testo per descrivere il progetto.

Proprietà cartella e Annotazioni

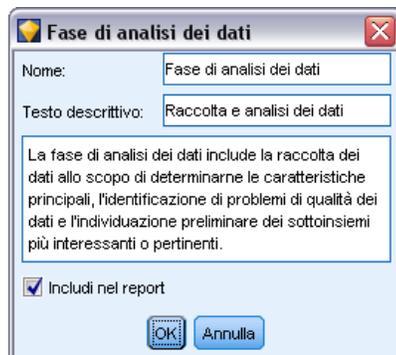
È possibile annotare le singole cartelle del progetto, sia nella visualizzazione CRISP-DM che nella visualizzazione Classi. Le annotazioni all'interno della visualizzazione CRISP-DM consentono di documentare in modo estremamente efficace gli obiettivi dell'organizzazione relativi a ogni fase del data mining. Per esempio, mediante le annotazioni è possibile inserire per la cartella *Business Understanding* un testo del tipo "L'obiettivo aziendale di questo studio è quello di ridurre il tasso di abbandono dei clienti più importanti". Il testo potrà quindi essere aggiunto automaticamente al report di progetto selezionando l'opzione *Includi nel report*.

Per annotare una cartella:

- ▶ Selezionare una cartella nel riquadro dei progetti.
- ▶ Fare clic sulla cartella con il pulsante destro del mouse e selezionare *Proprietà cartella*.

Nella visualizzazione CRISP-DM, le annotazioni delle cartelle includono un riepilogo dell'obiettivo associato a ogni fase e indicazioni sull'esecuzione delle principali attività del data mining. Le annotazioni possono essere rimosse o modificate.

Figura 11-7
Cartella di progetto con annotazione CRISP-DM



Nome. In quest'area è visualizzato il nome del campo selezionato.

Testo descrittivo. Crea i testi descrittivi personalizzati che saranno visualizzati quando si posiziona il puntatore del mouse su una cartella del progetto. Questi testi sono utili nella visualizzazione CRISP-DM, per esempio se si desidera fornire una breve panoramica degli obiettivi di ogni fase o contrassegnare lo stato di una fase come "In corso" o "Completata".

Campo dell'annotazione. In questo campo è possibile inserire annotazioni più lunghe che possono essere raccolte nel report del progetto. Le annotazioni della visualizzazione CRISP-DM includono una descrizione di ogni fase del data mining, che può tuttavia essere personalizzata in base alle specifiche esigenze del progetto.

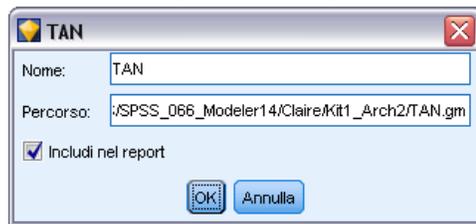
Includi nel report. Per includere l'annotazione nei report, selezionare l'opzione *Includi nel report*.

Proprietà oggetto

È possibile visualizzare le proprietà degli oggetti e decidere se includere i singoli oggetti nel report del progetto. Per accedere alle proprietà dell'oggetto:

- ▶ Fare clic con il pulsante destro del mouse su un oggetto nel riquadro dei progetti.
- ▶ Fare clic su Proprietà oggetto nel menu.

Figura 11-8
Finestra di dialogo Proprietà oggetto



Nome. Visualizza il nome dell'oggetto salvato.

Percorso. Visualizza la posizione dell'oggetto salvato.

Includi nel report. Selezionare questa opzione per inserire le informazioni sull'oggetto in un report generato.

Chiusura di un progetto

Quando si esce da IBM® SPSS® Modeler o si apre un nuovo progetto, il file di progetto esistente (.cpj) viene chiuso.

Alcuni file associati al progetto, quali stream, nodi o grafici, potrebbero essere ancora aperti. Se si desidera lasciare questi file aperti, rispondere No al messaggio ... Salvare e chiudere i file?

Se dopo avere chiuso un progetto si salva o si modifica un file associato, la versione aggiornata del file verrà inserita nel progetto alla successiva apertura. Per evitare che ciò accada, rimuovere il file dal progetto o salvarlo con un nome di file diverso.

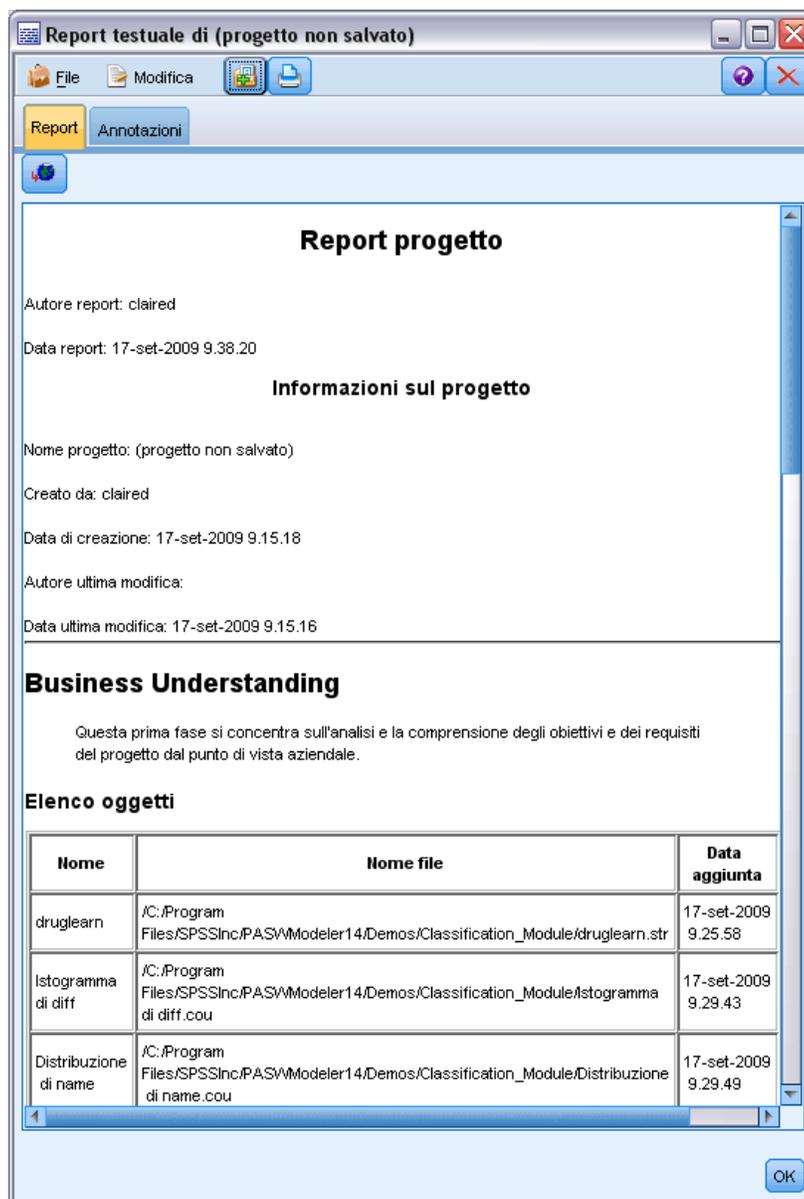
Generazione di report

La possibilità di creare report basati sugli elementi e sulle annotazioni del progetto è una delle funzioni più utili dei progetti e rappresenta un componente essenziale per l'efficacia del data mining, come illustrato nella metodologia CRISP-DM. Il report può essere generato direttamente in uno dei vari formati di file o in una finestra di output sullo schermo per consentire la visualizzazione immediata. Il report può essere quindi stampato, salvato o visualizzato in un browser Web. I report salvati possono essere distribuiti ad altri utenti della propria organizzazione.

Accade di frequente che i report vengano generati più volte dai file di progetto durante il processo di data mining, allo scopo di distribuirli alle persone coinvolte nel progetto. I report includono le informazioni sugli oggetti ai quali viene fatto riferimento nel file di progetto e

le eventuali annotazioni create ed è possibile generarli in base alla visualizzazione Classi o CRISP-DM.

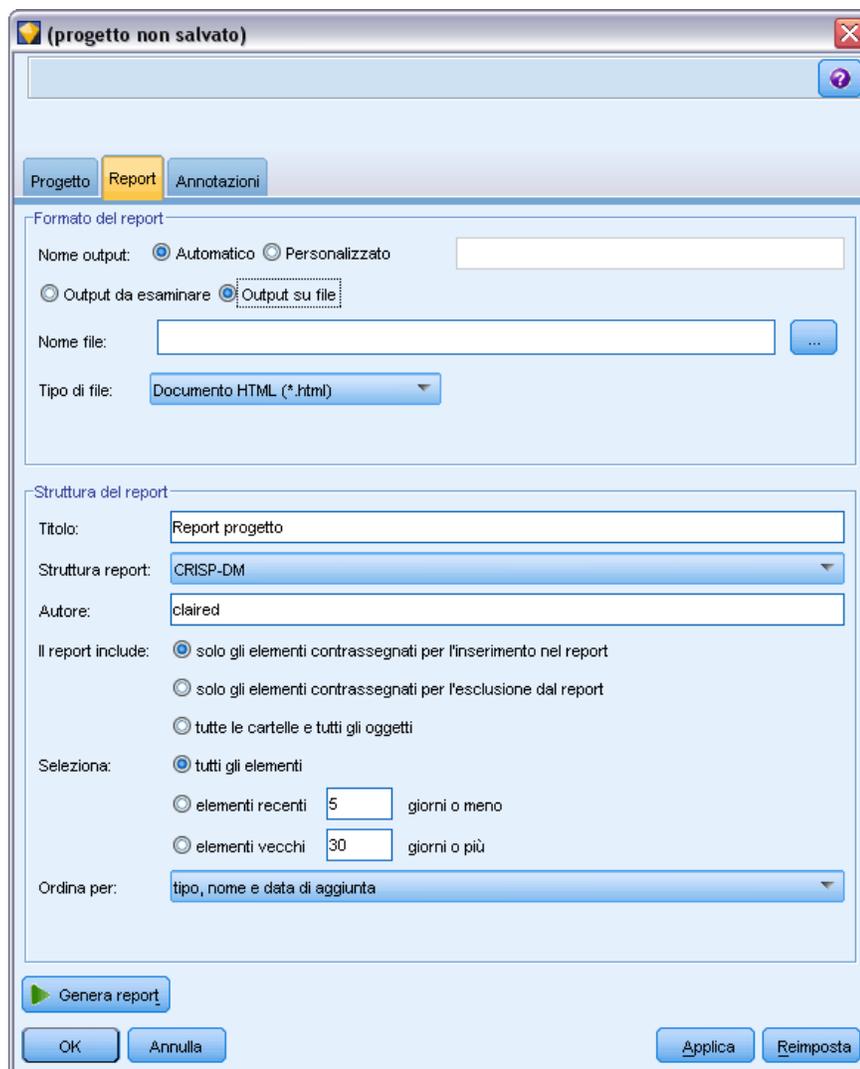
Figura 11-9
Finestra del report generato



Per generare un report:

- ▶ Selezionare la cartella del progetto nella visualizzazione CRISP-DM o Classi.
- ▶ Fare clic sulla cartella con il pulsante destro del mouse e selezionare Report progetto.
- ▶ Impostare le opzioni per il report e fare clic su Genera report.

Figura 11-10
Selezione delle opzioni per il report



Le opzioni presenti nella finestra di dialogo del report consentono di generare il tipo di report desiderato in diversi modi:

Nome output. Indicare il nome della finestra di output se si decide di inviare l'output del report allo schermo. È possibile specificare un nome personalizzato o lasciare che IBM® SPSS® Modeler assegni automaticamente un nome alla finestra.

Output su schermo. Selezionare questa opzione per generare il report e visualizzarlo in una finestra di output. Si noti che dalla finestra di output è possibile esportare il report in vari formati di file.

Output su file. Selezionare questa opzione per generare il report e salvarlo come file del tipo specificato nell'elenco Tipo di file.

Nome file. Specificare un nome di file per il report generato. I file vengono salvati per default nella directory `\bin` di SPSS Modeler. Per specificare una destinazione diversa, utilizzare il pulsante con i puntini di sospensione (...).

Tipo di file. I tipi di file disponibili sono:

- **Documento HTML.** Il report viene salvato in un unico file HTML. Se contiene dei grafici, questi vengono salvati sotto forma di file PNG a cui il file HTML fa riferimento. Quando si pubblica il report su Internet, accertarsi di avere caricato sia il file HTML che le immagini a cui fa riferimento.
- **File di testo.** Il report viene salvato in un unico file di testo. Se contiene dei grafici, nel report vengono inclusi solo dei riferimenti al nome del file e al percorso.
- **Documento di MS Word.** Il report viene salvato come un singolo documento, con gli eventuali grafici incorporati direttamente nel documento.
- **Documento di MS Excel.** Il report viene salvato come un singolo foglio di calcolo, con gli eventuali grafici incorporati direttamente nel foglio.
- **Documento di MS PowerPoint.** Ogni fase viene mostrata su una nuova diapositiva. Gli eventuali grafici sono incorporati direttamente nelle diapositive di PowerPoint.
- **Oggetto di output.** Quando viene aperto in SPSS Modeler, questo file (`.cou`) genera una finestra di output identica a quella generata con l'opzione Output su schermo nel Formato del report.

Nota: per esportare in un file di Microsoft Office, è necessario che l'applicazione corrispondente sia installata.

Titolo. Indicare un titolo per il report.

Struttura report. Selezionare CRISP-DM o Classi. La visualizzazione CRISP-DM fornisce un report di stato che contiene informazioni sintetiche di carattere generale, nonché dettagli su ogni fase del data mining. La visualizzazione Classi è basata su oggetti ed è più appropriata per la registrazione interna di dati e stream.

Autore. Viene visualizzato il nome utente di default, che può tuttavia essere modificato.

Il report include. Selezionare un metodo per includere gli oggetti nel report. Selezionare tutte le cartelle e tutti gli oggetti per inserire tutti gli elementi aggiunti al file di progetto. È inoltre possibile inserire gli elementi in base alla selezione di Includi nel report nelle proprietà degli oggetti. In alternativa, per verificare gli elementi non inseriti nel report, specificare l'inserimento unicamente degli elementi contrassegnati per l'esclusione (nei quali Includi nel report non è selezionato).

Seleziona. Questa opzione consente di aggiornare il progetto selezionando solo elementi recenti per il report oppure di tenere traccia dei vecchi problemi a volte non risolti impostando i parametri per elementi vecchi. Selezionare tutti gli elementi per non considerare il parametro del tempo per il report.

Ordina per. È possibile selezionare una combinazione delle seguenti caratteristiche degli oggetti per ordinarli all'interno di una cartella:

- **Tipo.** Raggruppa gli oggetti in base al tipo.
- **Nome.** Dispone gli oggetti in ordine alfabetico.
- **Data di aggiunta.** Ordina gli oggetti in base alla data in cui sono stati aggiunti al progetto.

Salvataggio ed esportazione dei report generati

I report generati sullo schermo vengono visualizzati in una nuova finestra di output. Gli eventuali grafici inclusi nel report vengono visualizzati come immagini in linea.

Nel report è riportato il numero totale dei nodi presenti in ogni stream. I numeri sono elencati sotto le seguenti intestazioni, che utilizzano la terminologia di IBM® SPSS® Modeler anziché quella di CRISP-DM:

- **Lettori di dati.** Nodi di input. [Per ulteriori informazioni, vedere l'argomento Osservazioni generali in il capitolo 2 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)
- **Scrittori di dati.** Nodi di esportazione. [Per ulteriori informazioni, vedere l'argomento Panoramica sui nodi di esportazione in il capitolo 7 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)
- **Generatori di modelli.** Nodi Modelli o di creazione. [Per ulteriori informazioni, vedere l'argomento Panoramica sui nodi Modelli in il capitolo 3 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)
- **Applicatori di modello.** Modelli generati, detti anche insiemi di modelli. [Per ulteriori informazioni, vedere l'argomento Insiemi di modelli in il capitolo 3 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)
- **Generatori di output.** Nodi Grafico o di output. [Per ulteriori informazioni, vedere l'argomento Caratteristiche comuni dei nodi Grafico in il capitolo 5 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#) [Per ulteriori informazioni, vedere l'argomento Panoramica sui nodi Output in il capitolo 6 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)
- **Altro.** Tutti gli altri nodi relativi al progetto. Per esempio, quelli disponibili nella scheda Oper su campi o Oper su record nella palette nodi.

Per salvare un report:

- ▶ Nel menu File, fare clic su Salva.
 - ▶ Digitare un nome per il file.
- Il report viene salvato come oggetto di output.

Per esportare un report:

- ▶ Nel menu File, fare clic su Esporta e sul tipo di file per l'esportazione.
 - ▶ Digitare un nome per il file.
- Il report viene salvato nel formato prescelto.

I tipi di file per l'esportazione sono i seguenti:

- HTML
- Testo
- Microsoft Word
- MS Excel
- MS PowerPoint

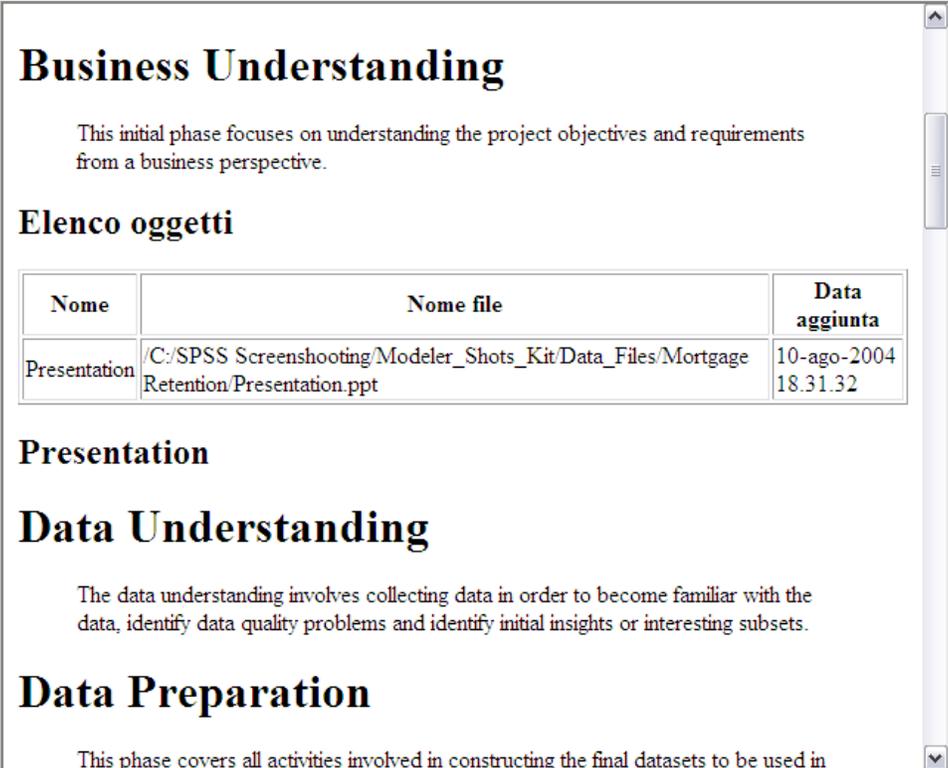
Nota: per esportare in un file di Microsoft Office, è necessario che l'applicazione corrispondente sia installata.

I pulsanti nella parte superiore della finestra consentono di:

- Stampare il report.
- Visualizzare il report in formato HTML in un browser Web esterno.

Figura 11-11

Report visualizzato in un browser Web



The screenshot shows a web browser window displaying a report. The report is titled "Business Understanding" and contains several sections: "Business Understanding", "Elenco oggetti", "Presentation", "Data Understanding", and "Data Preparation". The "Elenco oggetti" section contains a table with three columns: "Nome", "Nome file", and "Data aggiunta". The table has one row with the following data: "Presentation", "/C:/SPSS Screenshooting/Modeler_Shots_Kit/Data_Files/Mortgage Retention/Presentation.ppt", and "10-ago-2004 18.31.32".

Business Understanding

This initial phase focuses on understanding the project objectives and requirements from a business perspective.

Elenco oggetti

Nome	Nome file	Data aggiunta
Presentation	/C:/SPSS Screenshooting/Modeler_Shots_Kit/Data_Files/Mortgage Retention/Presentation.ppt	10-ago-2004 18.31.32

Presentation

Data Understanding

The data understanding involves collecting data in order to become familiar with the data, identify data quality problems and identify initial insights or interesting subsets.

Data Preparation

This phase covers all activities involved in constructing the final datasets to be used in

Personalizzazione di IBM SPSS Modeler

Personalizzazione delle opzioni di IBM SPSS Modeler

È possibile eseguire diverse operazioni per personalizzare IBM® SPSS® Modeler in base alle esigenze. La personalizzazione consiste soprattutto nell'impostazione di specifiche opzioni utente, quali l'allocazione di memoria, le directory di default e l'utilizzo di suono e colore. È inoltre possibile personalizzare la palette dei nodi situata nella parte inferiore della finestra di SPSS Modeler.

Impostazione delle opzioni di IBM SPSS Modeler

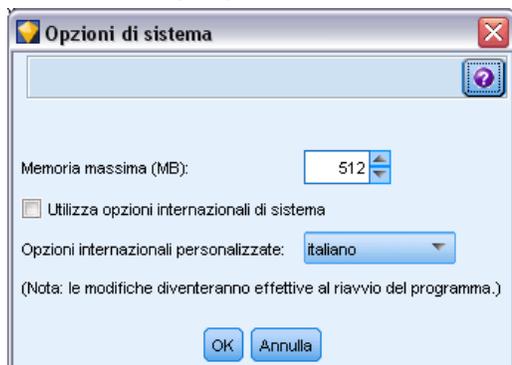
Le opzioni per IBM® SPSS® Modeler possono essere personalizzate e impostate in vari modi:

- Per impostare le opzioni di sistema, per esempio l'utilizzo della memoria e le opzioni internazionali, fare clic su Opzioni di sistema nel menu Strumenti > Opzioni.
- Per impostare le opzioni utente, per esempio i caratteri, gli avvisi e le ottimizzazioni, fare clic su Opzioni utente nel menu Strumenti > Opzioni.
- Per specificare la posizione delle applicazioni che interagiscono con SPSS Modeler, scegliere Applicazioni di supporto dal menu Strumenti > Opzioni
- Per specificare le directory di default utilizzate in SPSS Modeler, fare clic su Imposta directory o Imposta directory server nel menu File.

Opzioni di sistema

Per specificare la lingua preferita o le impostazioni internazionali per IBM® SPSS® Modeler, fare clic su Opzioni di sistema nel menu Strumenti > Opzioni. In questa finestra è inoltre possibile impostare la memoria massima utilizzabile da SPSS Modeler. Si noti che le modifiche apportate in questa finestra di dialogo diventeranno effettive solo al successivo riavvio di SPSS Modeler.

Figura 12-1
Finestra di dialogo Opzioni di sistema



Memoria massima. Selezionare questa opzione per stabilire un limite, in megabyte, per l'utilizzo della memoria da parte di SPSS Modeler. In alcune piattaforme, SPSS Modeler limita le dimensioni delle elaborazioni per ridurre l'utilizzo di risorse in computer con risorse limitate o forti carichi di lavoro. Se si stanno elaborando grandi quantità di dati, potrebbe derivarne un errore di "memoria esaurita". Per migliorare il carico di memoria, specificare una nuova soglia.

Utilizza opzioni internazionali di sistema. Questa opzione è selezionata per default e impostata su inglese (Stati Uniti). Deselezionarla per specificare un'altra lingua dall'elenco delle lingue e delle impostazioni internazionali disponibili.

Gestione della memoria

Oltre all'impostazione Memoria massima specificata nella finestra di dialogo Opzioni di sistema, sono disponibili svariate soluzioni per ottimizzare l'utilizzo della memoria:

- Impostare una cache in un qualsiasi nodo non terminale. In questo modo, all'esecuzione dello stream di dati, questi vengono letti dalla cache anziché essere recuperati dalla sorgente. Questa impostazione consente di ridurre il carico della memoria in caso di grandi insiemi di dati. [Per ulteriori informazioni, vedere l'argomento Opzioni di copia dei dati nella cache per i nodi in il capitolo 5 a pag. 47.](#)
- Configurare l'opzione Numero di membri massimo per campi nominali nella finestra di dialogo Proprietà stream. Questa opzione consente di specificare il numero massimo di membri per i campi nominali dopo il quale il livello di misurazione diventa *Senza tipo*. [Per ulteriori informazioni, vedere l'argomento Impostazione delle opzioni per gli stream in il capitolo 5 a pag. 51.](#)
- Liberare memoria facendo clic nell'angolo inferiore destro della finestra di IBM® SPSS® Modeler, nel punto in cui vengono visualizzate la memoria utilizzata da SPSS Modeler e la quantità di memoria allocata (xxMB/xxMB). Quando si fa clic al suo interno, tale area assume un colore più scuro, quindi i valori relativi all'allocazione diminuiscono. Quando il colore dell'area torna normale, in SPSS Modeler è stata liberata tutta la memoria disponibile.

Impostazione delle directory di default

Per specificare la directory di default utilizzata per i browser dei file e l'output, scegliere Imposta directory o Imposta directory server dal menu File.

- **Imposta directory.** Utilizzare questa opzione per impostare la directory che viene utilizzata. La directory di default in uso dipende dal percorso di installazione della versione di IBM® SPSS® Modeler o dal percorso della riga di comando utilizzato per avviare SPSS Modeler. In modalità locale, la directory di lavoro corrisponde al percorso utilizzato per tutte le operazioni e i file di output lato client (se ad essi viene fatto riferimento con percorsi relativi).
- **Imposta directory server.** L'opzione Imposta directory server, del menu File, viene attivata a ogni connessione del server remoto. Utilizzarla per specificare la directory di default per tutti i file del server e i file di dati specificati per l'input o l'output. La directory di default del server è *\$CLEO/data*, dove *\$CLEO* è la directory di installazione della versione Server di SPSS Modeler. Tramite la riga di comando, è possibile ignorare questa impostazione di default utilizzando il flag *-server_directory* con l'argomento della riga di comando *modelerclient*.

Impostazione delle opzioni utente

Per impostare le opzioni generali per IBM® SPSS® Modeler, scegliere Opzioni utente dal menu Strumenti > Opzioni. Queste opzioni si applicano a tutti gli stream utilizzati in SPSS Modeler.

Per impostare i seguenti tipi di opzioni, fare clic sulla scheda corrispondente:

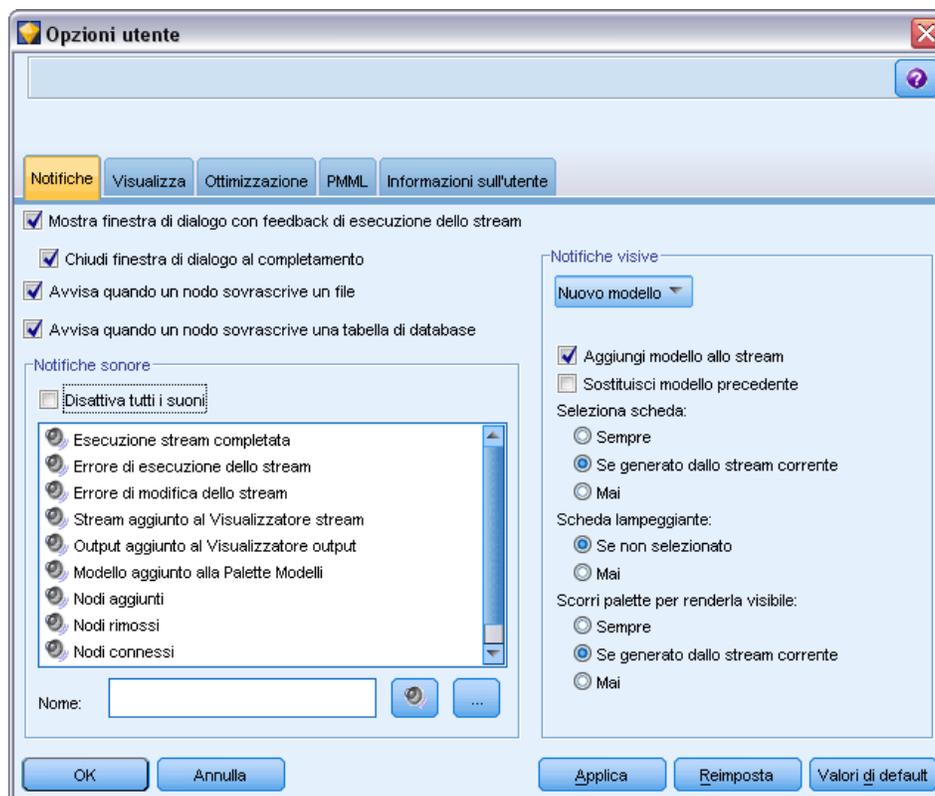
- **Notifiche**, per impostare per esempio opzioni relative alla sovrascrittura di modelli e ai messaggi di errore.
- **Visualizza**, per impostare per esempio opzioni relative ai colori di sfondo e dei grafici.
- **Ottimizzazione**, per impostare per esempio opzioni relative alla generazione SQL e la riscrittura degli stream.
- **Opzioni di esportazione PMML** utilizzate per l'esportazione di modelli in PMML (Predictive Model Markup Language).
- **Informazioni su autore o utente**, quali nome, iniziali e indirizzo di posta elettronica. Queste informazioni possono essere visualizzate nella scheda Annotazioni per i nodi e gli altri oggetti che vengono creati.

Per impostare opzioni specifiche per gli stream, per esempio i separatori decimali, i formati di data e ora e gli script degli stream, utilizzare la finestra di dialogo Proprietà stream disponibile dai menu File e Strumenti.

Impostazione delle opzioni di notifica

La scheda Notifiche della finestra di dialogo Opzioni utente consente di impostare varie opzioni relative all'occorrenza e al tipo di avvisi e finestre di conferma in IBM® SPSS® Modeler. È anche possibile specificare il comportamento delle schede Output e Modelli nel riquadro dei manager, al momento della generazione di nuovi output e modelli.

Figura 12-2
Finestra di dialogo Opzioni utente, scheda Notifiche



Mostra finestra di dialogo con feedback di esecuzione dello stream. Selezionare questa opzione per visualizzare una finestra di dialogo con un indicatore di avanzamento dopo tre secondi dall'inizio dell'esecuzione di uno stream. La finestra di dialogo comprende anche i dettagli degli oggetti di output creati dallo stream.

- **Chiudi finestra di dialogo al completamento.** Per default, la finestra di dialogo viene chiusa al termine dell'esecuzione dello stream. Deselezionare questa casella di controllo per continuare a visualizzare la finestra di dialogo una volta terminato lo stream.

Avvisa quando un nodo sovrascrive un file. Selezionare questa opzione se si desidera che venga visualizzato un messaggio di errore quando le operazioni dei nodi sovrascrivono un file esistente.

Avvisa quando un nodo sovrascrive una tabella di database. Selezionare questa opzione se si desidera che venga visualizzato un messaggio di errore quando le operazioni dei nodi sovrascrivono una tabella di database esistente.

Notifica acustica

Mediante l'elenco, specificare se utilizzare i suoni per notificare all'utente un evento o un errore. Sono disponibili numerosi suoni. Utilizzare il pulsante Esegui (altoparlante) per riprodurre il suono selezionato. Utilizzare il pulsante con i puntini di sospensione (...) per cercare e selezionare un suono.

Nota: i file .wav utilizzati per creare i suoni in SPSS Modeler sono memorizzati nella directory /media/sounds dell'installazione.

- **Disattiva tutti i suoni.** Selezionare questa opzione per disattivare la notifica acustica per tutti gli eventi.

Notifica visiva

Le opzioni riportate in questo gruppo sono utilizzate per specificare il comportamento delle schede Output e Modelli del riquadro dei manager nella parte superiore destra dello schermo quando vengono generati nuovi elementi. Selezionare Nuovo modello o Nuovo output dall'elenco per specificare il comportamento della scheda corrispondente.

Per Nuovo modello sono disponibili le seguenti opzioni:

Aggiungi modello allo stream. Se selezionata (impostazione predefinita), aggiunge un nuovo modello allo stream e alla scheda Modelli non appena viene creato il modello. Nello stream, il modello compare con un collegamento al nodo Modelli a partire dal quale è stato creato. Se si deseleziona questa casella, il modello viene aggiunto solo alla scheda Modelli.

Sostituisci modello precedente. Se selezionata (impostazione predefinita), sovrascrive un modello esistente dello stream nella scheda Modelli e nell'area di disegno dello stream. Se la casella è deselezionata, il modello viene aggiunto ai modelli esistenti sulla scheda e nell'area di disegno. Si noti che questa impostazione viene sovrascritta dall'impostazione di sostituzione del modello nel collegamento a un modello. [Per ulteriori informazioni, vedere l'argomento Sostituzione di un modello in il capitolo 3 in IBM SPSS Modeler 14.2 Nodi Modelli.](#)

Per Nuovo output sono disponibili le seguenti opzioni:

Avvisa quando output superano [n]. Selezionare se visualizzare un avviso quando il numero di elementi nella scheda Output supera una quantità predefinita. La quantità di default è 20, che può essere modificata se necessario.

Le seguenti opzioni sono disponibili in tutti i casi:

Seleziona scheda. Scegliere se passare alla scheda Output o Modelli quando l'oggetto corrispondente viene generato durante l'esecuzione dello stream.

- Selezionare Sempre per passare alla scheda corrispondente nel riquadro dei manager.
- Selezionare Se generato dallo stream corrente per attivare la scheda corrispondente per gli oggetti generati dallo stream attualmente visibile nell'area di disegno.
- Selezionare Mai per impedire che il software passi alla scheda corrispondente per notificare gli output o i modelli generati.

Scheda lampeggiante. Selezionare se far lampeggiare la scheda Output o Modelli nel riquadro dei manager ogni volta che vengono generati nuovi modelli o output.

- Selezionare Se non selezionato per far lampeggiare la scheda corrispondente (se non è già selezionata) ogni volta che nel riquadro dei manager vengono generati nuovi oggetti.
- Selezionare Mai per impedire che il software faccia lampeggiare la scheda corrispondente per notificare l'utente in merito alla generazione di oggetti.

Scorri palette per renderla visibile (solo Nuovo modello). Specificare se far scorrere automaticamente la scheda Modelli nel riquadro dei manager per rendere visibile il modello più recente.

- Selezionare Sempre per attivare lo scorrimento.
- Selezionare Se generato dallo stream corrente per far scorrere solo gli oggetti generati dallo stream attualmente visibile nell'area di disegno.
- Selezionare Mai per impedire che il software faccia scorrere automaticamente la scheda Modelli.

Apri finestra (solo Nuovo output). Selezionare se aprire automaticamente una finestra di output alla generazione.

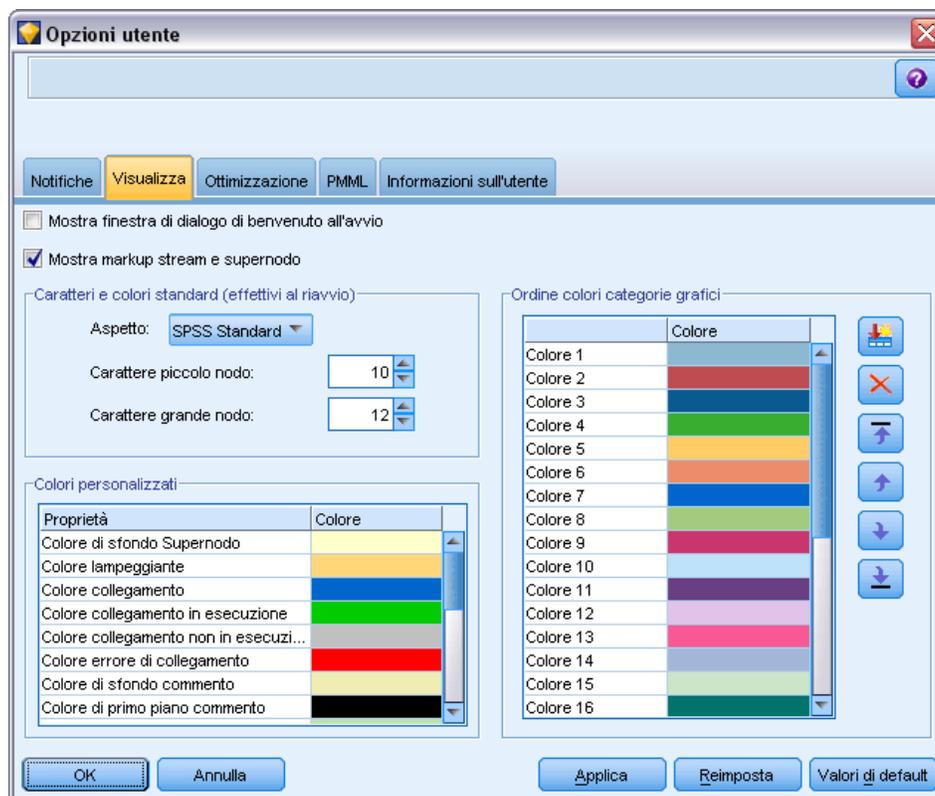
- Selezionare Sempre per aprire sempre una nuova finestra di output.
- Selezionare Se generato dallo stream corrente per aprire una nuova finestra per l'output generato dallo stream attualmente visibile nell'area di disegno.
- Selezionare Mai per impedire che il software apra automaticamente nuove finestre per l'output generato.

Fare clic su Valori di default per ripristinare le impostazioni predefinite del sistema relative a questa scheda.

Impostazione delle opzioni di visualizzazione

Tramite la scheda Visualizza della finestra di dialogo Opzioni utente, è possibile impostare le opzioni di visualizzazione dei caratteri e dei colori in IBM® SPSS® Modeler.

Figura 12-3
Finestra di dialogo Opzioni utente, scheda Visualizza



Mostra finestra di dialogo di benvenuto all'avvio. Selezionare per visualizzare la finestra di dialogo di benvenuto all'avvio. Tale finestra contiene le opzioni per avviare l'esercitazione Esempi di applicazioni, aprire uno stream di esempio o uno stream o progetto esistente oppure creare un nuovo stream.

Mostra markup stream e supernodo. Selezionare per far sì che l'eventuale markup su stream e i Supernodi venga visualizzato per default. Il markup comprende i commenti agli stream, i collegamenti a modelli e l'evidenziazione dei rami di calcolo del punteggio.

Caratteri e colori standard (effettivi al riavvio). Le opzioni presenti in questa casella di controllo sono utilizzate per specificare le caratteristiche dello schermo e lo schema di colori di SPSS Modeler nonché la dimensione dei caratteri visualizzati. Le opzioni selezionate verranno applicate solo previa chiusura e riavvio di SPSS Modeler.

- **Aspetto.** Consente di scegliere uno schema di colori e caratteristiche dello schermo standard. È possibile scegliere fra:
 - SPSS Standard (default), una struttura comune a tutti i prodotti IBM SPSS.
 - SPSS Classic, una struttura ben nota agli utenti delle precedenti versioni di SPSS Modeler.
 - Windows, un layout Windows che può essere utile per aumentare il contrasto delle palette e dell'area di disegno dello stream.

- **Carattere piccolo nodo.** Specificare la dimensione del carattere da utilizzare nelle palette dei nodi e per i nodi di piccole dimensioni visualizzati nell'area di disegno dello stream.
- **Carattere grande nodo.** Specificare la dimensione del carattere da utilizzare quando nell'area di disegno dello stream sono visualizzati nodi di grandi dimensioni (standard).

Nota: la dimensione dei nodi per uno stream può essere specificata nella scheda Layout della finestra di dialogo Proprietà stream.

Colori personalizzati. Questa tabella elenca i colori selezionati utilizzati per diversi elementi di visualizzazione. Per ciascun elemento elencato nella tabella, è possibile cambiare il colore corrente facendo doppio clic sulla riga corrispondente nella colonna Colore e selezionando un colore dall'elenco. Per specificare un colore personalizzato, scorrere l'elenco fino in fondo e fare clic sulla voce Colore....

Ordine colori categorie grafici. In questa tabella sono riportati i colori correntemente selezionati e utilizzati per la visualizzazione nei grafici appena creati. L'ordine dei colori riflette l'ordine in cui verranno utilizzati nel grafico. Se, per esempio, un campo nominale utilizzato come sovrapposizione colore contiene quattro valori univoci, verranno utilizzati solo i primi quattro colori elencati qui. Per ciascun elemento elencato nella tabella, è possibile cambiare il colore corrente facendo doppio clic sulla riga corrispondente nella colonna Colore e selezionando un colore dall'elenco. Per specificare un colore personalizzato, scorrere l'elenco fino in fondo e fare clic sulla voce Colore.... Le modifiche apportate non influenzano i grafici creati in precedenza.

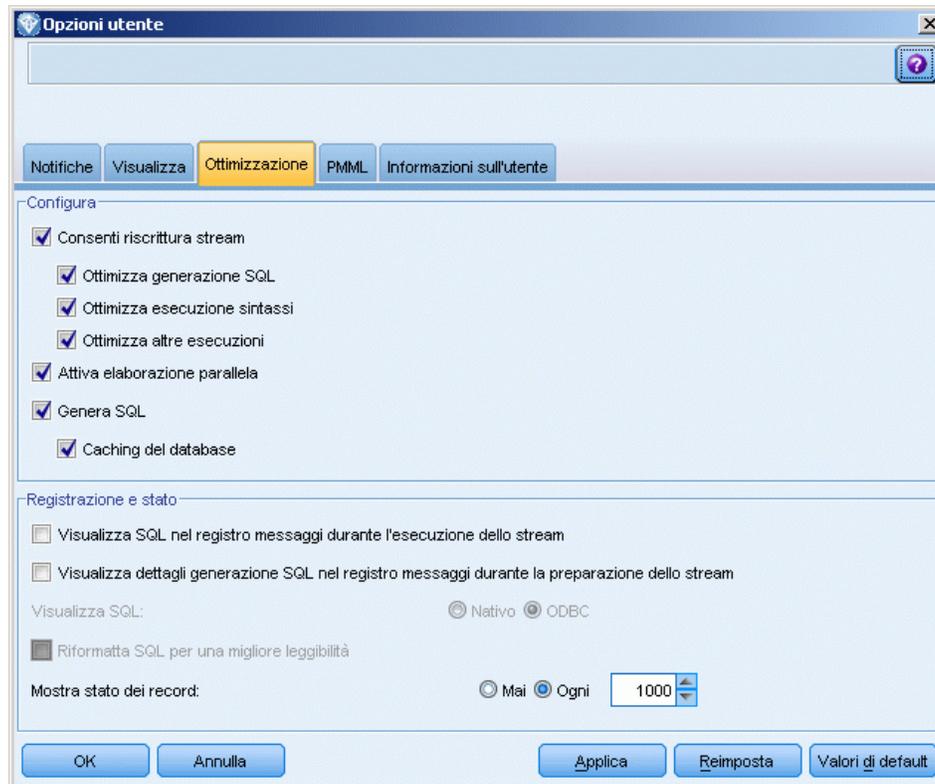
Fare clic su Valori di default per ripristinare le impostazioni predefinite del sistema relative a questa scheda.

Impostazione delle opzioni di ottimizzazione

La scheda Ottimizzazione della finestra di dialogo Opzioni utente consente di ottimizzare la performance degli stream. Si noti che le impostazioni relative alle prestazioni e all'ottimizzazione in IBM® SPSS® Modeler Server (se utilizzato) prevalgono sulle corrispondenti impostazioni nel client. *Questa scheda è disponibile solo se sul computer IBM® SPSS® Modeler è stata attivata la connettività di SPSS Modeler Server.*

Nota: le funzionalità di modellazione in-database e ottimizzazione SQL richiedono che sul computer SPSS Modeler sia attivata la connettività SPSS Modeler Server. Con questa impostazione attivata, è possibile accedere agli algoritmi di database, restituire codice SQL direttamente da SPSS Modeler e accedere a SPSS Modeler Server. Per verificare lo stato corrente della licenza, fare clic su Informazioni su nel menu Aiuto di SPSS Modeler. [Per ulteriori informazioni, vedere l'argomento Connessione a IBM SPSS Modeler Server in il capitolo 3 a pag. 9.](#)

Figura 12-4
Finestra di dialogo Opzioni utente, scheda Ottimizzazione



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 14.2, vedere le matrici di compatibilità del prodotto sul sito aziendale dedicato al supporto tecnico all'indirizzo <http://www.ibm.com/support>.

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

- **Ottimizza generazione SQL.** Questa opzione consente di riordinare i nodi nello stream in modo da permettere di eseguire il push back di un maggior numero di operazioni tramite la generazione SQL per l'esecuzione nel database. Quando trova un nodo che non può essere reso in codice SQL, l'ottimizzatore controlla più avanti l'eventuale presenza di nodi a valle che è possibile rendere in SQL e spostare in modo sicuro davanti al nodo del problema senza compromettere la semantica dello stream. Non solo il database può eseguire le operazioni in modo più efficiente di SPSS Modeler, ma questi pushback o "rinvii" hanno anche la possibilità di ridurre la dimensione dell'insieme di dati restituito a SPSS Modeler per l'elaborazione. Di conseguenza, si riduce il traffico di rete e si sveltiscono le operazioni stream. Si noti che è necessario selezionare la casella di controllo Genera SQL per consentire l'ottimizzazione SQL. Per ulteriori informazioni, vedere l'argomento Ottimizzazione SQL in il capitolo 6 in *IBM SPSS Modeler Server 14.2 Guida della performance e amministrazione*.

- **Ottimizza esecuzione sintassi.** Questo metodo di riscrittura degli stream aumenta l'efficienza delle operazioni che incorporano più di un nodo contenente la sintassi IBM® SPSS® Statistics. L'ottimizzazione si ottiene combinando i comandi di sintassi all'interno di un'unica operazione, invece di eseguirli singolarmente come operazioni separate.
- **Ottimizza altre esecuzioni.** Questo metodo di riscrittura degli stream aumenta l'efficienza delle operazioni (quelle che non possono essere delegate al database). L'ottimizzazione viene raggiunta riducendo la quantità dei dati nello stream non appena possibile. Pur mantenendo l'integrità dei dati, lo stream viene riscritto per spingere le operazioni più vicino alla sorgente dati, riducendo in tal modo i dati a valle per operazioni dispendiose quali le unioni.

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

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

- **Caching del database.** Per gli stream che generano SQL da eseguire nel database, i dati possono essere inseriti nella cache midstream in una tabella temporanea nel database anziché nel file system. Insieme all'ottimizzazione SQL, ciò consente di ottenere significativi vantaggi in termini di prestazioni. Per esempio, l'output di uno stream che unisce più tabelle per creare una visualizzazione data mining può essere inserito nella cache e riutilizzato in un secondo tempo in base alle esigenze. Con l'inserimento nella cache del database attivato, è sufficiente fare clic con il pulsante destro del mouse su qualsiasi nodo non terminale per inserire i dati nella cache in quel punto e la cache viene creata automaticamente direttamente nel database alla successiva esecuzione dello stream. Ciò consente la generazione di SQL per i nodi a valle, migliorando ulteriormente le prestazioni. In alternativa, se lo si desidera, è possibile disattivare questa opzione, per esempio se i criteri o le autorizzazioni precludono la scrittura dei dati nel database. Se l'inserimento nella cache del database o l'ottimizzazione SQL non sono attivati, la cache verrà scritta nel file system. [Per ulteriori informazioni, vedere l'argomento Opzioni di copia dei dati nella cache per i nodi in il capitolo 5 a pag. 47.](#)

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

Visualizza SQL nel registro dei messaggi durante l'esecuzione degli stream. Specifica se l'SQL generato durante l'esecuzione dello stream viene passato al registro del messaggio.

Visualizza SQL nel registro dei messaggi durante la preparazione degli stream. Durante l'anteprima degli stream, specifica se un'anteprima di SQL da generare viene passata al registro dei messaggi.

Visualizza SQL. Specifica se l'SQL visualizzato nel registro debba contenere funzioni SQL native o funzioni ODBC standard nel formato {fn FUNC(...)} generato da SPSS Modeler. Le prime si basano sulla funzionalità dei driver ODBC che non è possibile implementare. Per esempio, questo controllo non ha effetto su SQL Server.

Riformattare SQL per una migliore leggibilità. Specifica se l'SQL visualizzato nel registro debba essere formattato per migliorarne la leggibilità.

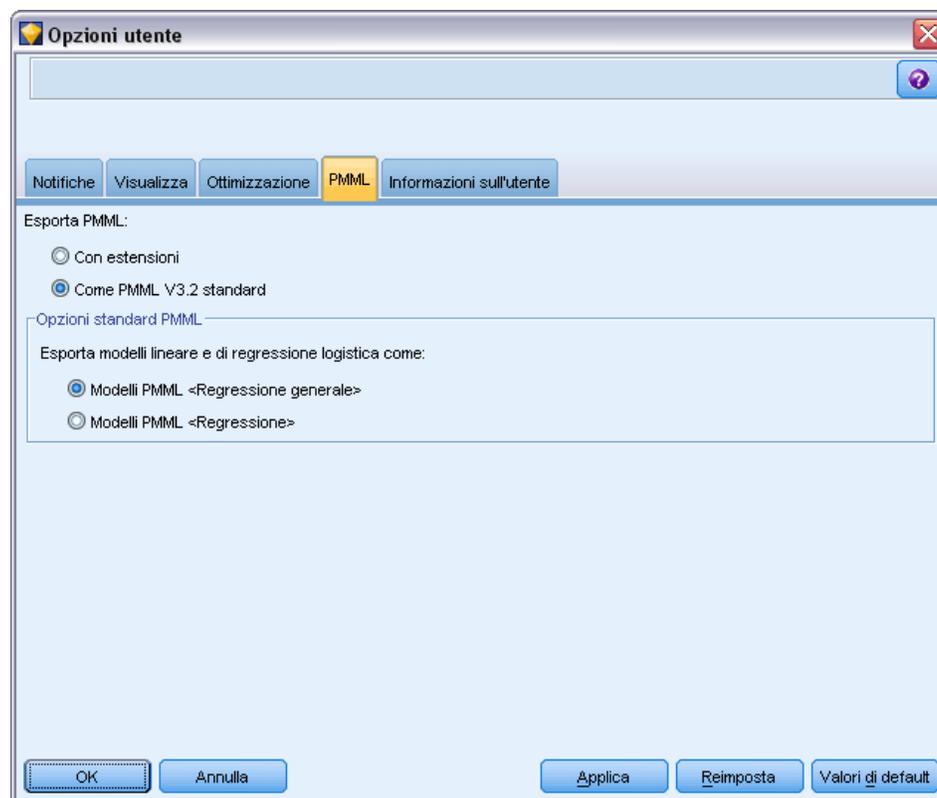
Mostra stato dei record. Specifica quando devono essere registrati i record al loro arrivo ai nodi terminali. Specificare il numero utilizzato per aggiornare lo stato ogni N record.

Fare clic su Valori di default per ripristinare le impostazioni predefinite del sistema relative a questa scheda.

Impostazione delle opzioni di esportazione PMML

Dalla scheda PMML è possibile controllare le modalità di esportazione dei modelli in PMML (Predictive Model Markup Language) di IBM® SPSS® Modeler. [Per ulteriori informazioni, vedere l'argomento Importazione ed esportazione di modelli come PMML in il capitolo 10 a pag. 206.](#)

Figura 12-5
Finestra di dialogo Opzioni utente, scheda PMML



Esporta PMML. Qui è possibile configurare le variazioni di PMML che funzionano meglio con l'applicazione di destinazione.

- Selezionare **Con estensioni** per consentire le estensioni PMML nei casi speciali in cui non esiste un equivalente PMML standard. Si noti che nella maggior parte dei casi questo produrrà gli stessi risultati di PMML standard.
- Selezionare **Come PMML V3.2 standard** per esportare PMML nella modalità più vicina possibile allo standard PMML.

Opzioni PMML standard. Se viene selezionata l'opzione **Come PMML V3.2 standard**, è possibile scegliere uno dei due metodi validi per esportare modelli di regressione lineare e logistica:

- Come Modelli PMML <Regressione generale>
- Come Modelli PMML <Regressione><

Impostazione delle informazioni sull'utente

Informazioni su utente/autore. Le informazioni immesse qui possono essere visualizzate nella scheda **Annotazioni dei nodi** e degli altri oggetti creati.

Personalizzazione della palette dei nodi

Gli stream vengono creati mediante i nodi. La palette dei nodi disponibile nella parte inferiore della finestra di IBM® SPSS® Modeler contiene tutti i nodi utilizzabili per la creazione degli stream. [Per ulteriori informazioni, vedere l'argomento Palette nodi in il capitolo 3 a pag. 15.](#)

La palette dei nodi può essere riorganizzata in due modi:

- Con la personalizzazione del **Manager palette**. [Per ulteriori informazioni, vedere l'argomento Personalizzazione del Manager palette a pag. 237.](#)
- Con la modifica del modo in cui le schede della palette che contengono sottopalette vengono visualizzate nella palette dei nodi. [Per ulteriori informazioni, vedere l'argomento Creazione di una sottopalette a pag. 242.](#)

Figura 12-6

Scheda **Oper su record** della palette dei nodi



Personalizzazione del Manager palette

È possibile personalizzare il **Manager palette** per soddisfare esigenze specifiche di utilizzo di IBM® SPSS® Modeler. Per esempio, se si analizzano di frequente dati di serie storica di un database, è possibile decidere di tenere contemporaneamente disponibili in un'unica scheda della palette il nodo di input Database, il nodo Intervalli di tempo, il nodo Serie storica e il nodo Grafico plot tempo. Il **Manager palette** consente di effettuare con facilità queste modifiche, creando schede della palette personalizzate nella palette dei nodi.

Il Manager palette consente di effettuare varie attività:

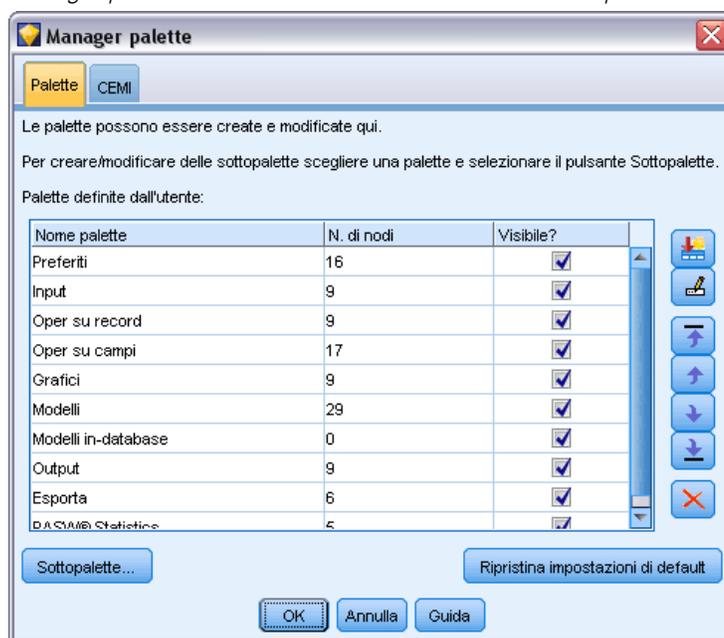
- Controllare quali schede della palette vengono visualizzate nella palette dei nodi sotto l'area di disegno.
- Controllare l'ordine in cui le schede della palette vengono visualizzate nella palette dei nodi.
- Creare e modificare le proprie schede della palette e le eventuali sottopalette associate.
- Modificare le selezioni di nodi di default nella scheda Preferiti.

Per accedere al Manager palette:

- Fare clic su Gestisci palette nel menu Strumenti.

Figura 12-7

Manager palette che mostra le schede visualizzate nella palette dei nodi



Nome palette. Viene elencata ogni scheda della palette disponibile, visualizzata o meno nella palette dei nodi, incluse le schede della palette create dall'utente. [Per ulteriori informazioni, vedere l'argomento Creazione di una scheda della palette a pag. 239.](#)

N. di nodi. Il numero di nodi visualizzato in ogni scheda della palette. Un numero elevato indica che potrebbe essere più utile creare sottopalette per suddividere i nodi sulla scheda. [Per ulteriori informazioni, vedere l'argomento Creazione di una sottopalette a pag. 242.](#)

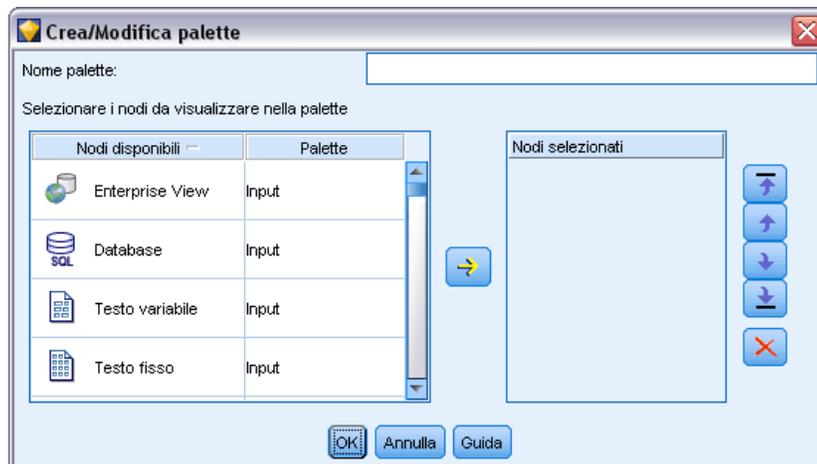
Visibile? Selezionare questo campo per visualizzare la scheda della palette nella palette dei nodi. [Per ulteriori informazioni, vedere l'argomento Visualizzazione delle schede della palette nella palette dei nodi a pag. 240.](#)

Sottopalette. Per selezionare le sottopalette da visualizzare su una scheda della palette, evidenziare il Nome palette richiesto e fare clic su questo pulsante per visualizzare la finestra di dialogo Sottopalette. [Per ulteriori informazioni, vedere l'argomento Creazione di una sottopalette a pag. 242.](#)

Ripristina valori predefiniti. Per rimuovere completamente tutte le modifiche e aggiunte effettuate a palette e sottopalette e ripristinare le impostazioni di default delle palette, fare clic su questo pulsante.

Creazione di una scheda della palette

Figura 12-8
Creazione di una scheda della palette nella finestra di dialogo Crea/Modifica palette



Per creare una scheda della palette personalizzata:

- ▶ Scegliere Manager palette dal menu Strumenti.
- ▶ A destra della colonna *Visibile?* fare clic sul pulsante Aggiungi palette a progetto. Viene visualizzata la finestra di dialogo Crea/Modifica palette.
- ▶ Digitare un Nome palette univoco.
- ▶ Nell'area Nodi disponibili, selezionare il nodo da aggiungere alla scheda della palette.
- ▶ Fare clic sul pulsante con la freccia a destra di aggiunta nodo per spostare il nodo evidenziato nell'area Nodi selezionati. Ripetere la procedura fino ad avere aggiunto tutti i nodi desiderati.

Al termine dell'aggiunta, è possibile modificare l'ordine in cui i nodi selezionati appaiono nella scheda della palette:

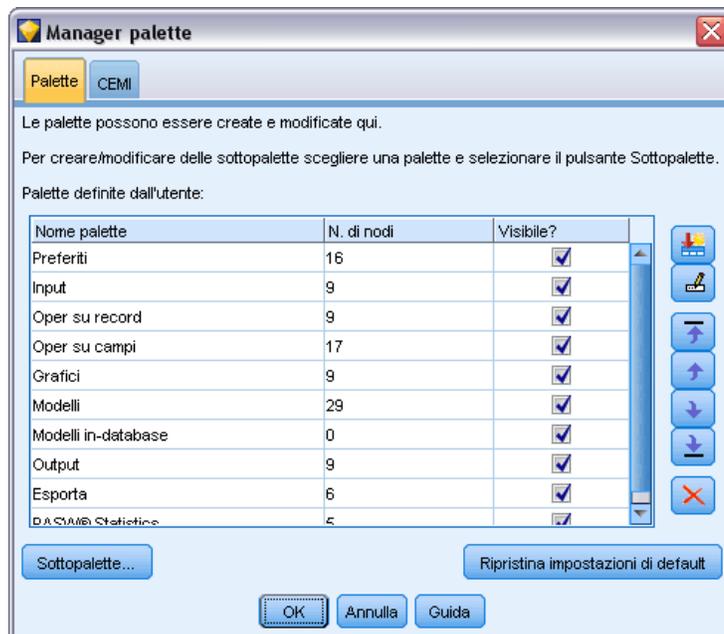
- ▶ Utilizzare i pulsanti con le frecce per spostare un nodo verso l'alto o verso il basso di una riga.
- ▶ Utilizzare i pulsanti con le frecce e una riga per spostare un nodo all'inizio o alla fine dell'elenco.
- ▶ Per rimuovere un nodo da una palette, evidenziarlo e fare clic sul pulsante Elimina a destra dell'area Nodi selezionati.

Visualizzazione delle schede della palette nella palette dei nodi

Potrebbero essere disponibili delle opzioni di IBM® SPSS® Modeler che non vengono mai utilizzate; in tal caso, è possibile utilizzare il Manager palette per nascondere le schede che contengono questi nodi.

Figura 12-9

Manager palette che mostra le schede visualizzate nella palette dei nodi



Per selezionare le schede da visualizzare nella palette dei nodi:

- ▶ Scegliere Manager palette dal menu Strumenti.
- ▶ Utilizzare le caselle di controllo disponibili nella colonna *Visibile?* per scegliere se includere o nascondere ogni scheda della palette.

Per rimuovere in modo permanente una scheda della palette dalla palette dei nodi, evidenziarla e fare clic sul pulsante Elimina a destra della colonna *Visibile?*. Una volta eliminata, una scheda della palette non può essere recuperata.

Nota: non è possibile eliminare le schede della palette di default fornite con SPSS Modeler, fatta eccezione per la scheda Preferiti.

Modifica dell'ordine di visualizzazione nella palette dei nodi

Dopo avere selezionato le schede della palette da visualizzare, è possibile modificare l'ordine in cui vengono visualizzate nella palette dei nodi:

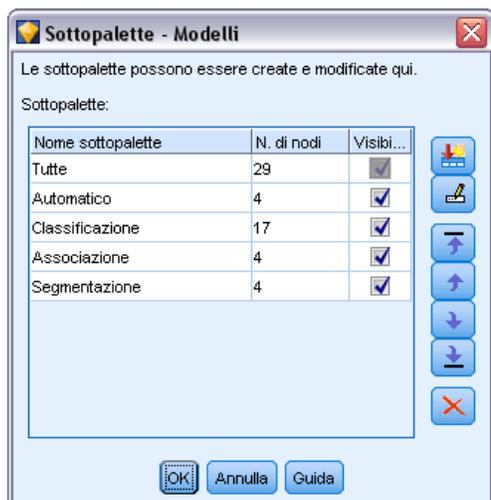
- ▶ Utilizzare i pulsanti con le frecce per spostare una scheda della palette verso l'alto o verso il basso di una riga. Lo spostamento in alto corrisponde alla sinistra della palette dei nodi e viceversa.

- Utilizzare i pulsanti con le frecce e una riga per spostare una scheda della palette all'inizio o alla fine dell'elenco. Le schede all'inizio dell'elenco appariranno a sinistra nella palette dei nodi.

Visualizzazione delle sottopalette su una scheda della palette

Così come è possibile controllare quali schede della palette vengono visualizzate nella palette dei nodi, è possibile controllare quali sottopalette sono disponibili per la relativa scheda della palette padre.

Figura 12-10
Sottopalette disponibili per la scheda della palette Modelli



Per selezionare le sottopalette da visualizzare in una scheda della palette:

- Scegliere Manager palette dal menu Strumenti.
- Selezionare la palette desiderata.
- Fare clic sul pulsante Sottopalette. Viene visualizzata la finestra di dialogo Sottopalette.
- Utilizzare le caselle di controllo disponibili nella colonna *Visibile?* per scegliere se includere o meno ogni sottopalette nella scheda della palette. La sottopalette Tutte viene sempre visualizzata e non è possibile eliminarla.
- Per rimuovere in modo permanente una sottopalette dalla scheda della palette, evidenziare la sottopalette e fare clic sul pulsante Elimina a destra della colonna *Visibile?*.

Nota: non è possibile eliminare le sottopalette di default fornite con la scheda della palette Modelli.

Modifica dell'ordine di visualizzazione nella scheda della palette

Dopo avere selezionato le sottopalette da visualizzare, è possibile modificare l'ordine in cui vengono visualizzate nella scheda della palette genitore:

- Utilizzare i pulsanti con le frecce semplici per spostare una sottopalette verso l'alto o verso il basso di una riga.

- Utilizzare i pulsanti con le frecce formate da una riga per spostare un nodo all'inizio o alla fine dell'elenco.

Le sottopalette create vengono visualizzate nella palette dei nodi quando si seleziona la relativa scheda della palette padre. [Per ulteriori informazioni, vedere l'argomento Modifica della visualizzazione di una scheda della palette a pag. 243.](#)

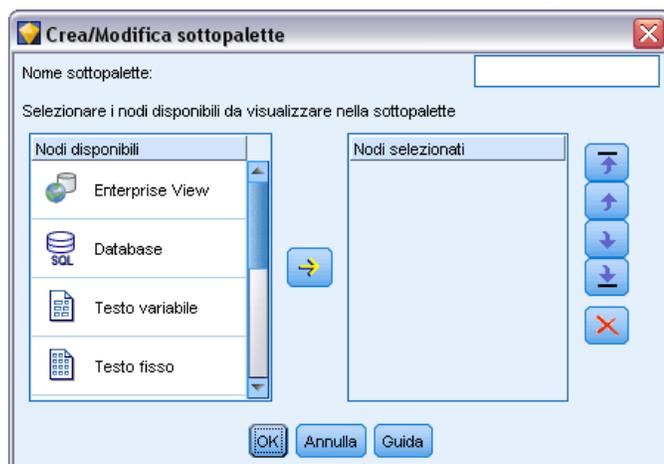
Creazione di una sottopalette

Poiché è possibile aggiungere tutti i nodi esistenti alle schede della palette personalizzate create, potrebbe accadere di selezionare un numero di nodi tale da richiedere lo scorrimento dello schermo per visualizzarli tutti. Per evitare lo scorrimento, è possibile creare delle sottopalette in cui collocare i nodi scelti per la scheda della palette. Per esempio, se si è creata una scheda della palette contenente i nodi utilizzati più frequentemente per creare gli stream, è possibile creare quattro sottopalette che suddividono le selezioni per nodo Input, Operazioni su campi, Modelli e Output.

Nota: è possibile selezionare nodi per le sottopalette solo tra i nodi aggiunti alla scheda della palette genitore.

Figura 12-11

Creazione di una sottopalette nella finestra di dialogo Crea/Modifica palette



Per creare una sottopalette:

- Scegliere Manager palette dal menu Strumenti.
- Selezionare la palette a cui si desidera aggiungere le sottopalette.
- Fare clic sul pulsante Sottopalette. Viene visualizzata la finestra di dialogo Sottopalette.
- A destra della colonna *Visibile?* fare clic sul pulsante Aggiungi sottopalette. Viene visualizzata la finestra di dialogo Crea/Modifica sottopalette.
- Digitare un Nome sottopalette univoco.
- Nell'area Nodi disponibili, selezionare il nodo da aggiungere alla sottopalette.

- ▶ Fare clic sul pulsante con la freccia a destra di aggiunta nodo per spostare il nodo selezionato nell'area Nodi selezionati.
- ▶ Al termine, scegliere OK per tornare alla finestra di dialogo Sottopalette.

Le sottopalette create vengono visualizzate nella palette dei nodi quando si seleziona la relativa scheda della palette padre. [Per ulteriori informazioni, vedere l'argomento Modifica della visualizzazione di una scheda della palette a pag. 243.](#)

Modifica della visualizzazione di una scheda della palette

A causa dell'elevato numero di nodi disponibili in IBM® SPSS® Modeler, questi potrebbero non essere tutti visibili su schermi di piccole dimensioni se non scorrendo a destra o sinistra la palette dei nodi, particolarmente nel caso della scheda della palette Modelli. Per ridurre la necessità di scorrimento, è possibile scegliere di visualizzare solo i nodi contenuti in una sottopalette, se disponibile. [Per ulteriori informazioni, vedere l'argomento Creazione di una sottopalette a pag. 242.](#)

Per cambiare i nodi visualizzati in una scheda della palette, selezionare la scheda e, dal menu a sinistra, selezionare se visualizzare tutti i nodi o solo i nodi contenuti in una specifica sottopalette.

Figura 12-12
Scheda della palette Modelli che mostra la sottopalette Classificazione



Gestione dei nodi CEMI

CEMI non è attualmente più supportato ed è stato sostituito da CLEF, che offre una serie di funzioni molto più versatili e semplici da usare. Per ulteriori informazioni, vedere il documento *IBM® SPSS® Modeler 14.2 CLEF Developer's Guide* fornito con questa versione.

Considerazioni sulle performance degli stream e dei nodi

È possibile progettare gli stream per ottimizzare le performance disponendo i nodi nella configurazione più efficiente e attivandone la cache quando è necessario, nonché prestando attenzione alle altre considerazioni esposte in questa sezione.

A parte le considerazioni illustrate in questa sede, in genere è possibile ottenere ulteriori e sensibili miglioramenti delle performance utilizzando il database in modo efficiente, in particolare mediante l'ottimizzazione SQL. [Per ulteriori informazioni, vedere l'argomento *Panoramica delle performance in il capitolo 5 in IBM SPSS Modeler Server 14.2 Guida della performance e amministrazione.*](#)

Ordine dei nodi

Anche quando non si fa uso dell'ottimizzazione SQL, l'ordine dei nodi in uno stream può incidere sulle prestazioni. L'obiettivo generale è ridurre al minimo l'elaborazione a valle; pertanto, i nodi che riducono la quantità di dati devono essere collocati nella parte iniziale dello stream. IBM® SPSS® Modeler Server può applicare automaticamente alcune regole di riordinamento in fase di compilazione per spostare in avanti alcuni nodi quando si ha la garanzia che l'operazione è priva di rischi. Questa funzione è attivata di default. Contattare l'amministratore di sistema per verificare che sia attivata nella propria installazione di Clementine.

Quando si utilizza l'ottimizzazione SQL, è consigliabile migliorarne la disponibilità e l'efficienza. Poiché l'ottimizzazione si interrompe quando lo stream contiene un'operazione che non può essere eseguita nel database, è preferibile raggruppare le operazioni ottimizzate da SQL nella parte iniziale dello stream. Questa strategia mantiene una parte più consistente dell'elaborazione nel database, trasferendo una mole di dati inferiore a IBM® SPSS® Modeler.

Le seguenti operazioni possono essere eseguite nella maggior parte dei database. Tentare di raggrupparle all'*inizio* dello stream:

- Unione per chiave (join)
- Select
- Aggrega
- Sort
- Campione
- Accoda
- Operazioni distinte in modalità *includi*, in cui vengono selezionati tutti i campi
- Operazioni di riempimento

- Operazioni di derivazione elementari mediante calcoli aritmetici standard o manipolazione delle stringhe (a seconda delle operazioni supportate dal database)
- Creazione di flag

Le seguenti operazioni non possono essere eseguite nella maggior parte dei database e devono quindi essere collocate all'interno dello stream *dopo* le operazioni riportate nell'elenco precedente:

- Operazioni su dati non contenuti nel database, per esempio file piatti
- Unione per ordine
- Balance
- Operazioni distinte in modalità *scarta* o in cui solo un sottoinsieme di campi è selezionato come distinto
- Tutte le operazioni che richiedono di accedere ai dati di record diversi da quello in corso di elaborazione
- Derivazioni dei campi stato e conteggio
- Operazioni del nodo Cronologia
- Operazioni con le funzioni "@" (di serie storica)
- Modalità di verifica del tipo *Avvisa* e *Termina*
- Costruzione, applicazione e analisi di modelli

Nota: i modelli di alberi decisionali, insiemi di regole, regressione lineare e generati da fattori possono generare stringhe SQL e possono quindi essere spostati nuovamente nel database.

- Output di dati a qualunque destinazione diversa dal database che sta elaborando i dati

Cache dei nodi

Per ottimizzare l'esecuzione dello stream, è possibile impostare una **cache** in qualsiasi nodo diverso da un nodo terminale. La cache configurata per un nodo viene riempita con i dati che passano attraverso il nodo alla successiva esecuzione dello stream di dati. Da questo momento in poi, i dati vengono letti dalla cache, che è archiviata in una directory temporanea, anziché dalla sorgente dati.

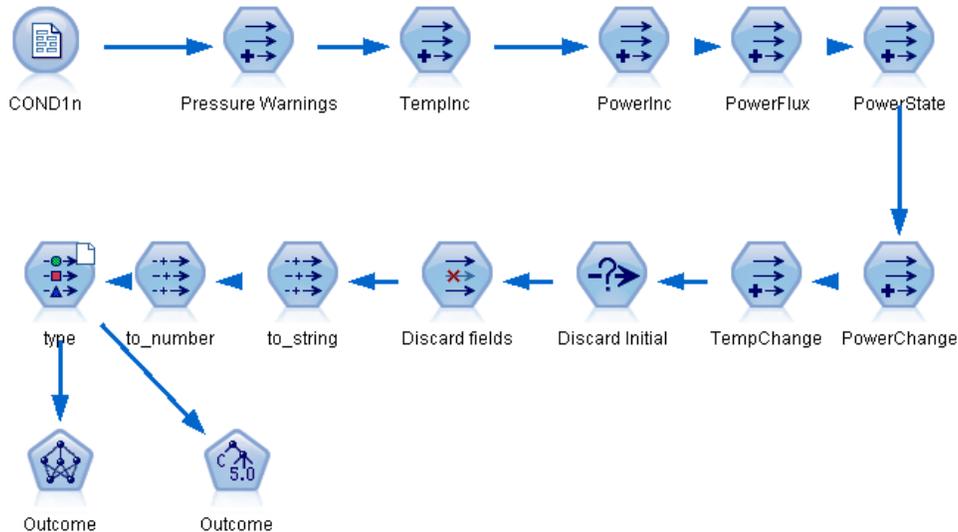
La copia dei dati nella cache è utile soprattutto dopo un'operazione che richiede tempo, per esempio un ordinamento, un'unione o un'aggregazione. Si supponga, per esempio, che esista un nodo di input impostato in modo da leggere i dati relativi alle vendite da un database e un nodo Aggregazione che riepiloga le vendite in base alla posizione. È possibile impostare una cache nel nodo Aggregazione anziché nel nodo di input, in modo tale da memorizzarvi i dati aggregati anziché l'intero insieme di dati.

Nota: l'impostazione di una cache nei nodi di input, che consente semplicemente di archiviare una copia dei dati originali non appena vengono letti in IBM® SPSS® Modeler, nella maggior parte dei casi non migliorerà le prestazioni.

I nodi con la cache attivata sono indicati da una piccola icona di documento nell'angolo in alto a destra. Quando i dati vengono inseriti nella cache del nodo, l'icona del documento diventa di colore verde.

Figura 13-1

Copia dei dati nella cache del nodo Tipo per memorizzare i campi appena derivati



Per attivare una cache:

- ▶ Nell'area di disegno dello stream, fare clic con il pulsante destro del mouse sul nodo e fare clic su Cache nel menu.
- ▶ Fare clic su Attiva nel sottomenu.
- ▶ Per disattivare la cache, fare clic con il pulsante destro del mouse sul nodo e fare clic su Disattiva nel sottomenu.

Inserimento nella cache di nodi nel database

Per gli stream eseguiti nel database, i dati possono essere inseriti nella cache midstream in una tabella temporanea nel database anziché nel file system. Insieme all'ottimizzazione SQL, ciò consente di ottenere significativi vantaggi in termini di prestazioni. Per esempio, l'output di uno stream che unisce più tabelle per creare una visualizzazione data mining può essere inserito nella cache e riutilizzato in un secondo tempo in base alle esigenze. Generando automaticamente codice SQL per tutti i nodi a valle, è possibile migliorare ulteriormente le prestazioni.

Se si utilizza l'inserimento nella cache del database con stringhe di lunghezza superiore a 256 caratteri, assicurarsi che sia presente un nodo Tipo a monte del nodo per cui è attivata la cache e che i valori dei campi vengano letti. In questo modo la colonna corrispondente nella tabella temporanea viene impostata sulla larghezza corretta per adattarsi alle stringhe.

Per sfruttare l'inserimento nella cache del database, è necessario che siano attivate l'ottimizzazione SQL e l'inserimento nella cache del database. Si noti che le impostazioni di ottimizzazione del Server prevalgono su quelle del Client. [Per ulteriori informazioni, vedere l'argomento Impostazione delle opzioni di ottimizzazione in il capitolo 12 a pag. 233.](#)

Con l'inserimento nella cache del database attivato, è sufficiente fare clic con il pulsante destro del mouse su qualsiasi nodo non terminale per inserire i dati nella cache in quel punto e la cache verrà creata automaticamente direttamente nel database alla successiva esecuzione dello stream. Se l'inserimento nella cache del database o l'ottimizzazione SQL non sono attivati, la cache verrà scritta nel file system.

Nota: i seguenti database supportano tabelle temporanee per l'inserimento di dati nella cache: DB2, Netezza, Oracle, SQL Server e Teradata. Altri database utilizzano una tabella normale per l'inserimento di dati nella cache. È possibile personalizzare il codice SQL per database specifici - contattare l'assistenza.

Performance: Nodi di elaborazione

Ordina. Il nodo Ordina deve leggere tutto l'insieme di dati di input prima di poterlo ordinare. I dati vengono salvati in memoria fino a un certo limite e i dati eccedenti vengono riversati su disco. L'algoritmo di ordinamento è di tipo combinato: i dati vengono letti e inseriti in memoria fino al limite e ordinati con un algoritmo di ordinamento rapido di tipo ibrido. Se i dati rientrano tutti nella memoria, l'ordinamento è terminato. In caso contrario viene applicato un algoritmo di ordinamento di unione. I dati ordinati vengono scritti su file e il successivo blocco di dati viene letto e inserito in memoria, ordinato e scritto su disco. L'operazione viene ripetuta fino a quando tutti i dati sono stati letti; i blocchi ordinati vengono quindi uniti. L'unione può richiedere ripetuti passaggi sui dati archiviati su disco. Al culmine dell'elaborazione, il nodo Ordina avrà due copie complete su disco dell'insieme di dati, una ordinata e una non ordinata.

Il tempo di esecuzione complessivo dell'algoritmo è nell'ordine di $N \cdot \log(N)$, in cui N è il numero di record. L'ordinamento in memoria è più veloce dell'unione da disco: pertanto, il tempo di esecuzione effettivo può essere ridotto allocando una maggiore quantità di memoria all'ordinamento. L'algoritmo alloca a sé stesso una porzione della RAM fisica controllata dall'opzione di configurazione di IBM® SPSS® Modeler Server *Moltiplicatore utilizzo memoria*. Per aumentare la memoria utilizzata per l'ordinamento, espandere la RAM fisica o incrementare il valore di questa opzione. Si noti che, quando la porzione di memoria utilizzata supera la quantità allocata per l'esecuzione del processo provocando il paging di parte della memoria su disco, si ha una diminuzione delle prestazioni perché l'algoritmo di ordinamento accede alla memoria secondo uno schema diretto che può determinare un paging eccessivo. L'algoritmo di ordinamento è utilizzato da vari nodi oltre al nodo Ordina, ma per tutti valgono le stesse regole riguardo alle prestazioni.

Discretizza. Il nodo Discretizza legge tutto l'insieme dei dati di input per calcolare i limiti dei bin prima di assegnare i record ai vari bin. Durante il calcolo dei limiti l'insieme di dati viene inserito nella cache e in seguito viene letto nuovamente per l'allocazione. Quando il metodo di discretizzazione è *a larghezza fissa* o *media+deviazione standard*, l'insieme di dati viene memorizzato direttamente su disco. Questi metodi hanno un tempo di esecuzione lineare e richiedono spazio su disco sufficiente per salvare l'intero insieme di dati. Quando il metodo di discretizzazione è per *ranghi* o *blocchi*, l'insieme di dati viene ordinato con l'algoritmo di ordinamento descritto in precedenza e, una volta ordinato, viene utilizzato come cache. A causa dell'ordinamento, il tempo di esecuzione di questi metodi è pari a $M \cdot N \cdot \log(N)$, dove M è il numero di campi nei bin e N è il numero dei record; lo spazio su disco necessario è pari al doppio delle dimensioni dell'insieme di dati.

Generando un nodo Nuovo campo in base ai bin generati è possibile migliorare le performance nei passaggi successivi. Le operazioni di derivazione sono molto più rapide della discretizzazione. [Per ulteriori informazioni, vedere l'argomento Anteprema dei bin generati in il capitolo 4 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)

Unione per chiave (join). Il nodo Unione, quando il metodo di unione è per *chiave* (equivalente al join di un database), ordina ogni insieme di dati di input in base ai campi chiave. Questa parte della procedura ha un tempo di esecuzione di $M*N*\log(N)$, dove M è il numero degli input e N è il numero di record nell'input più grande; l'operazione richiede una quantità di spazio su disco sufficiente per salvare tutti gli insiemi di dati di input più una seconda copia dell'insieme di dati più grande. Il tempo di esecuzione dell'operazione di unione è proporzionale alle dimensioni dell'insieme dei dati di output, che dipende dalla frequenza della corrispondenza delle chiavi. Nel peggiore dei casi, in cui l'output è il prodotto Cartesiano degli input, il tempo di esecuzione può quasi avvicinarsi a NM . Si tratta di un'eventualità rara: quasi tutte le operazioni di join hanno un numero molto inferiore di corrispondenze delle chiavi. Se un insieme di dati è relativamente più grande degli altri, o se i dati in ingresso sono già ordinati in base a un campo chiave, è possibile migliorare le prestazioni di questo nodo mediante la scheda Ottimizzazione. [Per ulteriori informazioni, vedere l'argomento Impostazioni di ottimizzazione dell'unione in il capitolo 3 in IBM SPSS Modeler 14.2 Nodi di input, elaborazione e output.](#)

Aggregazione. Quando l'opzione *Chiavi contigue* non è impostata, questo nodo legge (senza memorizzare) tutto l'insieme dei dati di input prima di generare l'output aggregato. Nei casi più estremi, in cui le dimensioni dei dati aggregati raggiungono un limite (determinato dall'opzione di configurazione *Moltiplicatore utilizzo memoria* di SPSS Modeler Server), il resto dell'insieme di dati viene ordinato ed elaborato come se l'opzione *Chiavi contigue* fosse impostata. Quando questa opzione è impostata non vi è alcun salvataggio dei dati, perché i record di output aggregati vengono generati durante la lettura dei dati di input.

Elimina dupl. Il nodo Elimina duplicati memorizza tutti i campi chiave univoci dell'insieme dei dati di input; nei casi in cui tutti i campi sono campi chiave e tutti i record sono univoci, il nodo memorizza tutto l'insieme di dati. Per default, il nodo Elimina duplicati ordina i dati nei campi chiave e quindi seleziona (o scarta) il primo record distinto di ogni gruppo. Per gli insiemi di dati più piccoli con un numero ridotto di chiavi distinte o per quelli che sono stati già ordinati, è possibile selezionare delle opzioni per migliorare la velocità e l'efficienza dell'elaborazione.

Tipo. In alcuni casi, il nodo Tipo inserisce i dati di input nella cache durante la lettura dei valori; la cache viene utilizzata per l'elaborazione a valle. Benché la cache richieda spazio su disco sufficiente per archiviare tutto l'insieme di dati, il suo utilizzo riduce i tempi di elaborazione.

Valutazione. Il nodo Valutazione deve ordinare i dati di input per calcolare i blocchi. L'operazione di ordinamento viene ripetuta per tutti i modelli valutati poiché i punteggi e l'ordine dei record che ne deriva sono diversi caso per caso. Il tempo di esecuzione è pari a $M*N*\log(N)$, dove M è il numero dei modelli e N è il numero di record.

Performance: Nodi Modelli

Rete neurale e Kohonen. Gli algoritmi di addestramento delle reti neurali (compreso l'algoritmo Kohonen) effettuano molti passaggi sui dati di addestramento. I dati vengono salvati in memoria fino a un certo limite e i dati eccedenti vengono riversati su disco. L'accesso ai dati di

addestramento dal disco è dispendioso in termini di risorse perché il metodo di accesso è diretto e può provocare un'eccessiva attività del disco. È possibile disattivare l'uso dell'archiviazione su disco per questi algoritmi, forzando il salvataggio di tutti i dati in memoria, selezionando l'opzione *Ottimizza per velocità* nella scheda *Modello* della finestra di dialogo del nodo. Si noti che, se la quantità di memoria necessaria per archiviare i dati supera la memoria di lavoro del processo server, viene eseguito il paging di una parte della memoria su disco e le prestazioni ne risentiranno di conseguenza.

Quando è selezionata l'ottimizzazione per la memoria, una percentuale della RAM fisica viene allocata all'algoritmo secondo il valore dell'opzione di configurazione di IBM® SPSS® *Modeler Server* *Percentuale limite memoria di modellazione*. Per utilizzare una maggiore quantità di memoria per l'addestramento delle reti neurali, espandere la RAM o incrementare il valore di questa opzione; tenere presente, tuttavia, che l'impostazione di un valore troppo alto causerà il paging su disco.

Il tempo di esecuzione degli algoritmi delle reti neurali dipende dal livello di precisione richiesto e può essere controllato impostando una condizione di arresto nella finestra di dialogo del nodo.

K-Means. L'algoritmo di cluster K-Means dispone delle stesse opzioni di controllo dell'utilizzo della memoria degli algoritmi delle reti neurali. Le prestazioni sui dati salvati su disco sono tuttavia migliori perché l'accesso ai dati è sequenziale.

Performance: espressioni CLEM

Le funzioni sequenza CLEM (“funzioni @”) che esaminano a ritroso lo stream di dati devono memorizzare una quantità di dati tale da consentire di risalire quanto più possibile all'indietro nello stream. Per le operazioni in cui il grado di ricerca a ritroso è illimitato, è necessario salvare tutti i valori del campo. Per operazione illimitata si intende un'operazione in cui il valore della distanza non è un numero intero vero e proprio, per esempio *@OFFSET(Sales, Month)*. Il valore della distanza è il nome del campo *Mese*, il cui valore è sconosciuto fino al momento dell'esecuzione. Il server deve salvare tutti i valori del campo *Vendite* per poter garantire risultati precisi. Quando si conosce il limite superiore è consigliabile specificarlo come argomento aggiuntivo, per esempio come *@OFFSET(Sales, Month, 12)*. Questa operazione comunica al server di archiviare solo i 12 valori più recenti del campo *Vendite*. Le funzioni di sequenza, siano esse limitate o meno, impediscono quasi sempre la generazione di stringhe SQL.

Accesso facilitato in IBM SPSS Modeler

Panoramica delle funzionalità per l'accesso facilitato in IBM SPSS Modeler

In questa versione sono state potenziate le funzionalità di accesso facilitato per tutti gli utenti e sono state introdotte caratteristiche specifiche per gli utenti con difficoltà visive o motorie. In questa sezione viene descritto come utilizzare funzionalità di accesso facilitato quali le utilità di lettura dello schermo e i tasti di scelta rapida.

Tipi di funzionalità per l'accesso facilitato

Questa applicazione per il data mining può essere utilizzata in numerosi modi, per far fronte alle esigenze di tutti gli utenti. Per esempio, è possibile creare stream, specificare opzioni e leggere l'output senza dover utilizzare il mouse. Negli argomenti seguenti verranno descritti i tasti di scelta rapida disponibili per le varie operazioni. Inoltre, IBM® SPSS® Modeler supporta un'ampia gamma di utilità di lettura dello schermo, per esempio JAWS per Windows. È possibile, inoltre, ottimizzare le combinazioni di colori in modo da aumentare il contrasto sullo schermo. Questi argomenti verranno illustrati più avanti in questa sezione.

Accesso facilitato per utenti con problemi di vista

In IBM® SPSS® Modeler sono disponibili numerose proprietà che consentono di ottimizzare e semplificare l'uso del software.

Opzioni di visualizzazione

È possibile selezionare i colori per la visualizzazione dei grafici. È possibile, inoltre, decidere di utilizzare le impostazioni di Windows anche per Clementine. Questo accorgimento, infatti, può servire ad aumentare il contrasto sullo schermo.

- ▶ Per impostare le opzioni di visualizzazione, fare clic su Opzioni utente nel menu Strumenti.
- ▶ Fare clic sulla scheda Visualizza. Le opzioni disponibili in questa scheda riguardano la combinazione di colori del software, i colori dei grafici e le dimensioni dei caratteri per i nodi.

Utilizzo di suoni per la notifica

L'attivazione o disattivazione di suoni consente di controllare la modalità di notifica per determinate operazioni nel software. Per esempio, è possibile attivare suoni per eventi quali la creazione o eliminazione di nodi o la generazione di nuovi modelli o output.

- ▶ Per impostare le opzioni di notifica, fare clic su Opzioni utente nel menu Strumenti.
- ▶ Fare clic sulla scheda Notifiche.

Controllo dell'apertura automatica di nuove finestre

La scheda Notifiche della finestra di dialogo Opzioni utente viene inoltre utilizzata per specificare se si desidera che i nuovi output generati, quali tabelle e grafici, vengano aperti in una finestra separata. Può essere più conveniente disattivare questa opzione e aprire una finestra di output solo se richiesto.

- ▶ Per impostare le opzioni, fare clic su Opzioni utente nel menu Strumenti.
- ▶ Fare clic sulla scheda Notifiche.
- ▶ Nella finestra di dialogo selezionare Nuovo output dall'elenco nel gruppo Notifiche visive.
- ▶ In Apri finestra, selezionare Mai.

Dimensioni dei nodi

È possibile visualizzare i nodi con le dimensioni standard o con dimensioni ridotte. Se necessario, è possibile regolare queste dimensioni in base alle proprie esigenze.

- ▶ Per impostare le opzioni per le dimensioni dei nodi, fare clic su Proprietà stream nel menu File.
- ▶ Fare clic sulla scheda Layout.
- ▶ Nell'elenco Dimensione icona, selezionare Standard.

Accesso facilitato per utenti non vedenti

Gli utenti non vedenti possono avvalersi di un'utilità per la lettura dello schermo, quale JAWS per Windows. In IBM® SPSS® Modeler sono disponibili varie impostazioni che consentono di ottimizzare l'uso delle utilità di lettura dello schermo.

Opzioni di visualizzazione

In genere le prestazioni delle utilità di lettura dello schermo sono migliori se il contrasto dello schermo è elevato. Se è già stato impostato un contrasto elevato per Windows, è possibile scegliere di utilizzare le impostazioni di Windows anche per Clementine.

- ▶ Per impostare le opzioni di visualizzazione, fare clic su Opzioni utente nel menu Strumenti.
- ▶ Fare clic sulla scheda Visualizza.

Utilizzo di suoni per la notifica

L'attivazione o disattivazione di suoni consente di controllare la modalità di notifica per determinate operazioni nel software. Per esempio, è possibile attivare suoni per eventi quali la creazione o eliminazione di nodi o la generazione di nuovi modelli o output.

- ▶ Per impostare le opzioni di notifica, fare clic su Opzioni utente nel menu Strumenti.
- ▶ Fare clic sulla scheda Notifiche.

Controllo dell'apertura automatica di nuove finestre

La scheda Notifiche della finestra di dialogo Opzioni utente viene inoltre utilizzata per specificare se si desidera che i nuovi output generati vengano aperti in una finestra separata. Può essere più conveniente disattivare questa opzione e aprire una finestra di output quando necessario.

- ▶ Per impostare le opzioni, fare clic su Opzioni utente nel menu Strumenti.
- ▶ Fare clic sulla scheda Notifiche.
- ▶ Nella finestra di dialogo selezionare Nuovo output dall'elenco nel gruppo Notifiche visive.
- ▶ In Apri finestra, selezionare Mai.

Accesso facilitato mediante la tastiera

La funzionalità del prodotto è accessibile da tastiera. Di base, è possibile premere Alt più il tasto pertinente per attivare i menu delle finestre (per esempio, premendo Alt+F si accede al menu File) oppure premere il tasto Tab per spostarsi fra i controlli delle finestre di dialogo. Tuttavia, ogni finestra principale del prodotto presenta problemi speciali ma sono anche disponibili utili suggerimenti per l'esplorazione delle finestre di dialogo.

In questa sezione verranno descritte le funzionalità per l'accesso facilitato mediante la tastiera, grazie alle quali è possibile svolgere con efficienza ogni operazione, dall'apertura di uno stream all'utilizzo delle finestre di dialogo dei nodi, fino agli interventi sull'output. Inoltre, verranno forniti elenchi dei tasti di scelta rapida, per rendere ancora più efficiente l'esplorazione.

Tasti di scelta rapida per l'esplorazione della finestra principale

La finestra principale di IBM® SPSS® Modeler è il luogo in cui si svolge la maggior parte del lavoro di data mining. L'area principale è detta **area di disegno dello stream** e serve per creare ed eseguire gli stream di dati. Nella parte inferiore della finestra sono presenti le **palette dei nodi**, che contengono tutti i nodi disponibili. Le palette sono strutturate in schede, che corrispondono al tipo di operazione di data mining per ogni gruppo di nodi. Per esempio, i nodi che consentono di portare i dati in SPSS Modeler sono raggruppati nella scheda Input, mentre i nodi che consentono di derivare campi o applicarvi filtri sono raggruppati nella scheda Oper su campi (abbreviazione di Operazioni sui campi).

Nel lato destro della finestra sono presenti vari strumenti per la gestione degli stream, dell'output e dei progetti. La metà superiore a destra contiene i **manager** e presenta tre schede, rispettivamente per la gestione degli stream, dell'output e dei modelli generati. È possibile accedere a questi oggetti selezionando la scheda e un oggetto dall'elenco. La metà inferiore a

destra contiene il **riquadro dei progetti**, che consente di organizzare il lavoro in progetti. In quest'area sono presenti due schede, che consentono di accedere a due visualizzazioni diverse di un progetto. Nella **visualizzazione Classi** gli oggetti sono ordinati per tipo, mentre nella **visualizzazione CRISP-DM** gli oggetti sono ordinati in base alla fase di data mining pertinente, per esempio Data Preparation o Modeling. Questi aspetti della finestra di SPSS Modeler sono trattati in dettaglio nella Guida in linea e nel Manuale dell'utente.

Di seguito è riportata una tabella con i tasti di scelta rapida che consentono di spostarsi nella finestra principale di SPSS Modeler e di creare stream. I tasti di scelta rapida per le finestre di dialogo e l'output sono riportati negli argomenti successivi. Questi tasti di scelta rapida sono disponibili solo nella finestra principale.

Tasti di scelta rapida per la finestra principale

Tasto di scelta rapida	Funzione
Ctrl+F5	Consente di spostare lo stato attivo sulle palette di nodi.
Ctrl+F6	Consente di spostare lo stato attivo sull'area di disegno dello stream.
Ctrl+F7	Consente di spostare lo stato attivo nel riquadro dei manager.
Ctrl+F8	Consente di spostare lo stato attivo nel riquadro dei progetti.

Tasti di scelta rapida per nodi e stream

Tasto di scelta rapida	Funzione
Ctrl+N	Consente di creare una nuova area di disegno dello stream vuota.
Ctrl+O	Consente di visualizzare la finestra di dialogo Apri, da cui è possibile selezionare e aprire uno stream esistente.
Ctrl+tasti numerici	Consente di spostare lo stato attivo sulla scheda corrispondente in una finestra o riquadro. In una finestra o in un riquadro con schede, per esempio, Ctrl+1 consente di passare alla prima scheda partendo da sinistra, Ctrl+2 alla seconda e così via.
Ctrl+Freccia giù	Nella palette dei nodi, consente di spostare lo stato attivo da una scheda della palette al primo nodo sotto a tale scheda.
Ctrl+Freccia su	Nella palette dei nodi, consente di spostare lo stato attivo da un nodo alla scheda della palette corrispondente.
Enter	Quando è selezionato un nodo in una palette dei nodi (inclusi i modelli rifiniti nella palette dei modelli generati), consente di aggiungere il nodo all'area di disegno dello stream. Se nell'area di disegno è già selezionato un nodo, premendo Invio verrà aperta la finestra di dialogo del nodo.
Ctrl+Invio	Quando nella palette è selezionato un nodo, consente di aggiungere tale nodo all'area di disegno dello stream senza selezionarlo, spostando lo stato attivo sul primo nodo della palette.
Alt+Invio	Quando nella palette è selezionato un nodo, consente di aggiungere tale nodo all'area di disegno dello stream e di selezionarlo, spostando lo stato attivo sul primo nodo della palette.
Maiusc+Barra spaziatrice	Quando un nodo o un commento ha lo stato attivo nella palette, consente di alternare tra la selezione e la deselection del nodo o del commento. Se vengono selezionati anche altri nodi o commenti, questo comporta la loro deselection.

Tasto di scelta rapida	Funzione
Ctrl+Maiusc+Barra spaziatrice	Quando un nodo o commento ha lo stato attivo nello stream o nella palette, consente di alternare tra la selezione e la deselection del nodo o del commento. Questo non incide su altri eventuali nodi o commenti selezionati.
Freccia destra/sinistra	Se l'area di disegno dello stream ha lo stato attivo, consente di spostare l'intero stream orizzontalmente sullo schermo. Se una scheda della palette ha lo stato attivo, consente di spostarsi tra le schede. Se un nodo della palette ha lo stato attivo, consente di spostarsi tra i nodi della palette.
Freccia su/giù	Se l'area di disegno dello stream ha lo stato attivo, consente di spostare l'intero stream verticalmente sullo schermo. Se un nodo della palette ha lo stato attivo, consente di spostarsi tra i nodi della palette. Se una palette secondaria ha lo stato attivo, consente di spostarsi tra le altre palette secondarie della scheda della palette.
Alt+Freccia destra/sinistra	Consente di spostare i nodi e i commenti selezionati nell'area di disegno dello stream orizzontalmente nella direzione della freccia.
Alt+Freccia su/giù	Consente di spostare i nodi e i commenti selezionati nell'area di disegno dello stream verticalmente nella direzione della freccia.
Ctrl+A	Consente di selezionare tutti i nodi di uno stream.
Ctrl+Q	Quando un nodo ha lo stato attivo, consente di selezionare il nodo stesso e tutti i nodi a valle e di deselectionare tutti i nodi a monte.
Ctrl+W	Quando un nodo ha lo stato attivo, consente di deselectionare il nodo stesso e tutti i nodi a valle selezionati.
Ctrl+Alt+D	Consente di duplicare il nodo selezionato.
Ctrl+Alt+L	Quando nello stream è selezionato un insieme di modelli, apre una finestra di dialogo Inserisci che consente di caricare un modello salvato da un file .nod nello stream.
Ctrl+Alt+R	Visualizza la scheda Annotazioni per un nodo selezionato per consentire di rinominare il nodo.
Ctrl+Alt+U	Consente di creare un nodo di input Input utente.
Ctrl+Alt+C	Consente di attivare/disattivare la cache per un nodo.
Ctrl+Alt+F	Consente di svuotare la cache di un nodo.
Tab	Nell'area di disegno dello stream, consente di spostarsi tra tutti i commenti e i nodi di input dello stream corrente. In una sottopalette selezionata, consente di alternare lo stato attivo tra l'etichetta della sottopalette e il primo nodo della sottopalette.
Maiusc+Tab	Esegue la medesima operazione del tasto Tab ma in ordine inverso.
Ctrl+Tab	Con lo stato attivo nell'area di disegno dello stream, nel riquadro dei manager o nel riquadro dei progetti, sposta lo stato attivo sulla palette dei nodi.
Qualsiasi tasto alfabetico	Con lo stato attivo su un nodo dello stream corrente, consente di assegnare lo stato attivo e di passare al nodo successivo il cui nome inizia con la lettera del tasto premuto.
F1	Consente di aprire la Guida in linea in corrispondenza di un argomento relativo all'elemento attivo.
F2	Consente di iniziare il processo di connessione per un nodo selezionato nell'area di disegno. Utilizzare il tasto Tab per passare al nodo richiesto nell'area di disegno e premere Maiusc+Barra spaziatrice per terminare la connessione.
F3	Consente di eliminare tutte le connessioni per il nodo selezionato nell'area di disegno.

Tasto di scelta rapida	Funzione
F6	Consente di spostare lo stato attivo tra il riquadro dei manager, il riquadro dei progetti e le palette dei nodi.
F10	Apre il menu File.
Maiusc+F10	Apre il menu di scelta rapida per il nodo o lo stream.
Canc	Consente di eliminare il nodo selezionato dall'area di disegno.
Esc	Chiude il menu di scelta rapida o la finestra di dialogo.
Ctrl+Alt+X	Consente di espandere un Supernodo.
Ctrl+Alt+Z	Consente di eseguire lo Zoom avanti in un Supernodo.
Ctrl+Alt+Maiusc+Z	Consente di eseguire lo Zoom indietro dal Supernodo
Ctrl+E	Con lo stato attivo nell'area di disegno dello stream, consente di eseguire lo stream corrente.

In SPSS Modeler sono inoltre disponibili vari tasti di scelta rapida standard, come Ctrl+C, che consente di effettuare una copia. [Per ulteriori informazioni, vedere l'argomento Utilizzo dei tasti di scelta rapida in il capitolo 3 a pag. 22.](#)

Tasti di scelta rapida per finestre di dialogo e tabelle

Sono disponibili numerosi tasti di scelta rapida e tasti per la lettura dello schermo, che agevolano l'utilizzo di finestre di dialogo, tabelle e tabelle all'interno di finestre di dialogo. Di seguito è riportato un elenco completo.

Tasti di scelta rapida per le finestre di dialogo e per il generatore di espressioni

Tasto di scelta rapida	Funzione
Alt+4	Consente di chiudere tutte le finestre di dialogo o le finestre di output aperte. L'output può essere recuperato dalla scheda Output nel riquadro dei manager.
Ctrl+Fine	Se lo stato attivo è su un controllo nel generatore di espressioni, consente di spostare il punto di inserimento alla fine dell'espressione.
Ctrl+1	Nel generatore di espressioni consente di spostare lo stato attivo sul controllo di modifica dell'espressione.
Ctrl+2	Nel generatore di espressioni consente di spostare lo stato attivo sull'elenco di funzioni.
Ctrl+3	Nel generatore di espressioni consente di spostare lo stato attivo sull'elenco di campi.

Tasti di scelta rapida per le tabelle

Questi tasti di scelta rapida funzionano per le tabelle di output e per i controlli tabella in finestre di dialogo di nodi quali Tipo, Filtro e Unione. Di norma si utilizza il tasto Tab per spostarsi tra le celle delle tabelle e Ctrl+Tab per uscire. *Nota:* in alcuni casi può accadere che l'utilità di lettura

dello schermo non legga subito il contenuto di una cella. Premendo una o due volte i tasti freccia, si reimposta il software e si avvia la lettura.

Tasto di scelta rapida	Funzione
Ctrl+W	Per le tabelle, consente di leggere la descrizione breve della riga (row) selezionata. Per esempio, "I valori della riga selezionata 2 sono sesso, flag, m/f, e così via."
Ctrl+Alt+W	Per le tabelle, consente di leggere la descrizione lunga della riga (row) selezionata. Per esempio, "I valori della riga selezionata 2 sono campo = sesso, tipo = flag, sesso = m/f, e così via."
Ctrl+D	Per le tabelle, consente di leggere la descrizione breve dell'area selezionata. Per esempio, "La selezione è composta da una riga per sei colonne."
Ctrl+Alt+D	Per le tabelle, consente di leggere la descrizione lunga dell'area selezionata. Per esempio, "La selezione è composta da una riga per sei colonne. Le colonne selezionate sono Campo, Tipo, Mancante. La riga selezionata è 1".
Ctrl+T	Per le tabelle, consente di leggere una descrizione breve delle colonne selezionate. Per esempio, "Campi, Tipo, Mancante".
Ctrl+Alt+T	Per le tabelle, consente di leggere una descrizione lunga delle colonne selezionate. Per esempio, "Le colonne selezionate sono Campi, Tipo, Mancante".
Ctrl+R	Per le tabelle, consente di conoscere il numero di record nella tabella.
Ctrl+Alt+R	Per le tabelle, consente di conoscere il numero di record nella tabella, nonché i nomi delle colonne.
Ctrl+I	Per le tabelle, consente di leggere le informazioni, ovvero il contenuto, della cella con lo stato attivo.
Ctrl+Alt+I	Per le tabelle, consente di leggere la descrizione lunga delle informazioni della cella con lo stato attivo, ovvero nome della colonna e contenuto della cella.
Ctrl+G	Per le tabelle, consente di leggere informazioni brevi, di carattere generale, sulla selezione.
Ctrl+Alt+G	Per le tabelle, consente di leggere informazioni lunghe, di carattere generale, sulla selezione.
Ctrl+Q	Per le tabelle, consente di attivare e disattivare rapidamente (quick) le descrizioni delle celle delle tabelle. Ctrl+Q consente di leggere descrizioni lunghe, per esempio "Sesso=Femmina" mentre ci si sposta nella tabella utilizzando i tasti freccia. Se si seleziona nuovamente Ctrl+Q, si passa alle descrizioni brevi (contenuto della cella).

Tasti di scelta rapida per i commenti

Quando si opera con i commenti su schermo, è possibile utilizzare i seguenti tasti di scelta rapida.

Tasto di scelta rapida	Funzione
Alt+C	Consente di mostrare/nascondere la funzione commenti.
Alt+M	Consente di inserire un nuovo commento se i commenti sono visualizzati o di visualizzare i commenti se sono nascosti.
Tab	Nell'area di disegno dello stream, consente di spostarsi tra tutti i commenti e i nodi di input dello stream corrente.

Tasto di scelta rapida	Funzione
Enter	Quando un commento ha lo stato attivo, indica l'inizio della modifica.
Alt+Invio o Ctrl+Tab	Termina la modifica e salva le modifiche apportate.
Esc	Annulla la modifica. I cambiamenti apportati durante la modifica vengono persi.
Alt+Maiusc+Freccia su	Riduce l'altezza dell'area di testo di una cella di griglia (o un pixel) se è attivato (o disattivato) il blocco sulla griglia.
Alt+Maiusc+Freccia giù	Aumenta l'altezza dell'area di testo di una cella di griglia (o un pixel) se è attivato (o disattivato) il blocco sulla griglia.
Alt+Maiusc+Freccia sinistra	Riduce la larghezza dell'area di testo di una cella di griglia (o un pixel) se è attivato (o disattivato) il blocco sulla griglia.
Alt+Maiusc+Freccia destra	Aumenta la larghezza dell'area di testo di una cella di griglia (o un pixel) se è attivato (o disattivato) il blocco sulla griglia.

Tasti di scelta rapida per il Visualizzatore cluster e il Visualizzatore del modello

Per spostarsi all'interno delle finestre del Visualizzatore cluster e del Visualizzatore del modello è possibile utilizzare dei tasti di scelta rapida.

Comandi generici - Visualizzatore cluster e Visualizzatore del modello

Tasto di scelta rapida	Funzione
Tab	Consente di spostare lo stato attivo sul comando successivo di quella schermata.
Maiusc+Tab	Consente di spostare lo stato attivo sul comando precedente di quella schermata.
Freccia giù	Se un elenco a discesa ha lo stato attivo, consente di aprire l'elenco o di spostarsi sulla voce successiva al suo interno. Se un menu ha lo stato attivo, consente di spostarsi sulla voce di menu seguente. Se il grafico in anteprima ha lo stato attivo, consente di spostarsi al grafico successivo (o al primo grafico se lo stato attivo è sull'ultima anteprima).
Freccia su	Se un elenco a discesa è aperto, consente di spostarsi sulla voce precedente al suo interno. Se un menu ha lo stato attivo, consente di spostarsi sulla voce di menu precedente. Se il grafico in anteprima ha lo stato attivo, consente di spostarsi al grafico precedente (o all'ultimo grafico se lo stato attivo è sulla prima anteprima).
Enter	Consente di chiudere un elenco a discesa aperto o di effettuare una selezione in un menu aperto.
F6	Consente di alternare lo stato attivo tra il riquadro destro e sinistro della finestra.
Freccia destra e sinistra	Se una scheda ha lo stato attivo, consente di spostarsi sulla scheda precedente o successiva. Se un menu ha lo stato attivo, consente di spostarsi sul menu precedente o successivo.
Alt+lettera	Consente di selezionare il pulsante o il menu nel cui nome è sottolineata la lettera.
Esc	Consente di chiudere un menu o un elenco a discesa aperto.

Solo Visualizzatore cluster

Il Visualizzatore cluster dispone di una visualizzazione Cluster che contiene una griglia cluster-per-funzioni. Per ulteriori informazioni, vedere l'argomento [Visualizzazione cluster in il capitolo 11 in IBM SPSS Modeler 14.2 Nodi Modelli](#).

Per scegliere la visualizzazione Cluster anziché la visualizzazione Riepilogo del modello:

- ▶ Premere più volte il tasto Tab fino a quando viene selezionato il pulsante Visualizza.
- ▶ Premere due volte il tasto Freccia giù per selezionare Cluster.

Da qui è possibile selezionare una singola cella all'interno della griglia:

- ▶ Premere più volte il tasto Tab fino ad arrivare sull'ultima icona della barra degli strumenti di visualizzazione.

Figura A-1

Icona Mostra struttura di visualizzazione



- ▶ Premere un'altra volta il tasto Tab, seguito dalla Barra spaziatrice e da un tasto freccia.

Attualmente sono disponibili i seguenti tasti di scelta rapida:

Tasto di scelta rapida	Funzione
Tasto freccia	Consente di spostare lo stato attivo tra le singole celle della griglia. La visualizzazione della distribuzione delle celle nel riquadro di destra cambia via via che lo stato attivo passa da una cella all'altra.
Ctrl+, (virgola)	Consente di selezionare o deselegionare tutta la colonna nella griglia in cui una cella ha lo stato attivo. Per aggiungere una colonna alla selezione, spostarsi su una cella in quella colonna con i tasti freccia e premere nuovamente Ctrl+.
Tab	Consente di spostare lo stato attivo al di fuori della griglia e sul comando successivo della schermata.
Maiusc+Tab	Consente di spostare lo stato attivo al di fuori della griglia e sul comando precedente della schermata.
F2	Consente di accedere alla modalità di modifica (solo celle di descrizione e delle etichette).
Enter	Consente di salvare le modifiche apportate e di uscire dalla modalità di modifica (solo celle di descrizione e delle etichette).
Esc	Consente di uscire dalla modalità di modifica senza salvare le modifiche (solo celle di descrizione e delle etichette).

Esempio di utilizzo dei tasti di scelta rapida: Creazione di stream

Per rendere più chiara la procedura per la creazione di stream mediante la tastiera o un'utilità di lettura dello schermo, di seguito è riportato un esempio che illustra come creare uno stream senza utilizzare il mouse. In questo esempio verrà creato uno stream contenente un nodo Testo Variabile, un nodo Nuovo campo e un nodo Istogramma:

- ▶ **Avviare SPSS Modeler.** Quando IBM® SPSS® Modeler viene avviato, lo stato attivo è sulla scheda Preferiti della palette dei nodi.
- ▶ **Ctrl+Freccia giù.** Consente di spostarsi dalla scheda nel corpo della scheda.
- ▶ **Freccia destra.** Consente di spostare lo stato attivo sul nodo Testo Variabile.
- ▶ **Barra spaziatrice.** Consente di selezionare il nodo Testo Variabile.
- ▶ **Ctrl+Invio.** Consente di aggiungere il nodo Testo variabile all'area di disegno dello stream. Inoltre, la selezione viene mantenuta sul nodo Testo Variabile, in modo che il nodo che verrà aggiunto successivamente risulti connesso ad esso.
- ▶ **Tab.** Consente di spostare lo stato attivo sulla palette dei nodi.
- ▶ **Freccia destra 4 volte.** Consente di passare al nodo Nuovo campo.
- ▶ **Barra spaziatrice.** Consente di selezionare il nodo Nuovo campo.
- ▶ **Alt+Invio.** Consente di aggiungere il nodo Nuovo campo all'area di disegno e di spostare la selezione sul nodo Nuovo campo. Ora questo nodo è pronto per essere connesso al successivo nodo che verrà aggiunto.
- ▶ **Tab.** Consente di spostare lo stato attivo sulla palette dei nodi.
- ▶ **Freccia destra 5 volte.** Consente di spostare lo stato attivo sul nodo Istogramma nella palette.
- ▶ **Barra spaziatrice.** Consente di selezionare il nodo Istogramma.
- ▶ **Invio.** Consente di aggiungere il nodo allo stream e di spostare lo stato attivo sull'area di disegno dello stream.

Passare all'esempio successivo o salvare lo stream se si desidera provare il prossimo esempio in un secondo tempo.

Esempio di utilizzo dei tasti di scelta rapida: Modifica dei nodi

In questo esempio verrà utilizzato lo stream creato nell'esempio precedente. Lo stream è costituito da un nodo Testo Variabile, da un nodo Nuovo campo e da un nodo Istogramma. La procedura parte con lo stato attivo sul terzo nodo dello stream, vale a dire sul nodo Istogramma.

- ▶ **Ctrl+Freccia sinistra 2 volte.** Consente di spostare nuovamente lo stato attivo sul nodo Testo Variabile.
- ▶ **Invio.** Consente di aprire la finestra di dialogo Testo Variabile. Con il tasto Tab spostarsi fino al campo File e digitare il percorso e il nome di un file di testo per selezionare il file. Con la

combinazione di tasti Ctrl+Tab spostarsi nella parte inferiore della finestra di dialogo fino a raggiungere il pulsante OK e premere Invio per chiudere la finestra di dialogo.

- ▶ **Ctrl + Freccia destra.** Consente di spostare lo stato attivo sul secondo nodo, vale a dire il nodo Nuovo campo.
- ▶ **Invio.** Consente di aprire la finestra di dialogo del nodo Nuovo campo. Premere Tab per selezionare i campi e specificare le condizioni per la derivazione. Premere Ctrl+Tab per spostarsi sul pulsante OK e premere Invio per chiudere la finestra di dialogo.
- ▶ **Ctrl + Freccia destra.** Consente di spostare lo stato attivo sul terzo nodo, vale a dire il nodo Istogramma.
- ▶ **Invio.** Consente di aprire la finestra di dialogo del nodo Istogramma. Premere Tab per selezionare i campi e specificare le opzioni per il grafico. Per gli elenchi a discesa, premere il tasto Freccia giù per aprire l'elenco ed evidenziare una voce, quindi premere Invio per selezionarla. Con il tasto Tab spostarsi sul pulsante OK e premere Invio per chiudere la finestra di dialogo.

A questo punto è possibile aggiungere altri nodi o eseguire lo stream corrente. Di seguito sono riportati alcuni suggerimenti da tenere presente quando si creano stream:

- Quando si collegano manualmente i nodi, utilizzare F2 per creare il punto di inizio di una connessione, il tasto Tab per spostarsi sul punto di fine e Maiusc+Barra spaziatrice per finalizzare la connessione.
- Utilizzare F3 per eliminare tutte le connessioni per il nodo selezionato nell'area di disegno.
- Dopo aver creato uno stream, utilizzare Ctrl+E per eseguire lo stream corrente.

È disponibile un elenco completo dei tasti di scelta rapida. [Per ulteriori informazioni, vedere l'argomento Tasti di scelta rapida per l'esplorazione della finestra principale a pag. 252.](#)

Uso di un'utilità di lettura dello schermo

Sul mercato sono disponibili numerose utilità di lettura dello schermo. IBM® SPSS® Modeler è configurato per supportare JAWS per Windows tramite il Java Access Bridge installato insieme a SPSS Modeler. Se JAWS è installato, per utilizzare il prodotto è sufficiente avviare JAWS prima di avviare SPSS Modeler.

Per la natura specifica e unica della rappresentazione grafica del processo di data mining in SPSS Modeler, la rappresentazione visiva di grafici e diagrammi è eccellente. È possibile, tuttavia, analizzare i dati e prendere decisioni sulla base di output e modelli anche visualizzandoli in forma testuale, tramite un'utilità di lettura dello schermo.

Nota: sui computer client a 64 bit alcune funzioni delle tecnologie per l'accesso facilitato non funzionano, in quanto Java Access Bridge non è progettato per il funzionamento a 64 bit.

Utilizzo del file di dizionario di IBM SPSS Modeler

In SPSS Modeler è disponibile un file di dizionario (*Awt.JDF*) da utilizzare con JAWS. Per utilizzare questo file:

- ▶ Passare alla sottodirectory */accessibility* dell'installazione di SPSS Modeler e copiare il file di dizionario (*Awt.JDF*).
- ▶ Copiarlo nella directory con gli script per JAWS.

Se nel sistema sono presenti altre applicazioni JAVA, è possibile che esista già un altro file denominato *Awt.JDF*. In questo caso, per poter utilizzare il file di dizionario potrebbe essere necessario modificarlo manualmente.

Utilizzo di un'utilità di lettura dello schermo con output HTML

In IBM® SPSS® Modeler possono verificarsi alcune difficoltà per quanto riguarda la visualizzazione di output HTML mediante un'utilità di lettura dello schermo. I problemi si verificano con i seguenti tipi di output:

- Output visualizzato nella scheda Avanzate dei nodi Regressione, Regressione logistica e Fattoriale
- Output del nodo Report

In ognuna di queste finestre o finestre di dialogo è presente uno strumento sulla barra degli strumenti che consente di visualizzare l'output nel browser predefinito, in cui è disponibile il supporto per le utilità di lettura dello schermo. È quindi possibile utilizzare l'utilità per la lettura dello schermo per trasmettere le informazioni relative all'output.

Accesso facilitato nella finestra Albero interattivo

La visualizzazione standard di un modello di albero decisionale nella finestra Albero interattivo potrebbe causare problemi alle utilità per la lettura dello schermo. Per accedere a una versione di accesso facilitato, nei menu Albero interattivo fare clic su:

Visualizza > Finestra accessibile

Viene visualizzata una vista simile alla mappa dell'albero standard che JAWS può leggere correttamente. È possibile spostarsi in alto, in basso, verso destra, verso sinistra utilizzando i tasti freccia standard. Durante l'esplorazione della finestra accessibile, si sposta anche lo stato attivo nella finestra Albero interattivo. Utilizzare la barra spaziatrice per modificare la selezione, oppure utilizzare Ctrl+Barra spaziatrice per estendere la selezione corrente.

Suggerimenti per l'utilizzo

Esistono vari accorgimenti per rendere più facilmente accessibile l'ambiente di IBM® SPSS® Modeler. Di seguito sono riportati alcuni suggerimenti di carattere generale per l'utilizzo di SPSS Modeler.

- **Uscire da caselle di testo estese.** Utilizzare Ctrl+Tab per uscire da caselle di testo estese. *Nota:* Ctrl+Tab consente, inoltre, di uscire dalle tabelle.
- **Utilizzo di Tab anziché dei tasti freccia.** Quando si selezionano le opzioni di una finestra di dialogo è necessario utilizzare Tab per spostarsi tra i pulsanti d'opzione. In questo contesto, infatti, i tasti freccia non funzionano.
- **Elenchi a discesa.** Negli elenchi a discesa delle finestre di dialogo è possibile utilizzare il tasto Esc o la Barra spaziatrice per selezionare una voce e chiudere l'elenco. È possibile, inoltre, utilizzare il tasto Esc per chiudere gli elenchi a discesa che rimangono aperti dopo che si è passati a un altro controllo mediante il tasto Tab.
- **Stato dell'esecuzione.** Quando si esegue uno stream o un database di grandi dimensioni, può verificarsi un ritardo di JAWS per quanto riguarda la lettura dello stato dello stream. Premere Ctrl periodicamente per aggiornare la lettura dello stato.
- **Utilizzo delle palette di nodi.** La prima volta che si accede a una scheda della palette di nodi, può accadere che JAWS non proponga il nome del nodo. In questo caso è possibile utilizzare Ctrl+Freccia destra, quindi Ctrl+Freccia sinistra per reimpostare l'utilità di lettura dello schermo e ottenere il nome del nodo.
- **Letture dei menu.** Può accadere, la prima volta che si apre un menu, che JAWS salti il primo comando. Se si sospetta un'evenienza simile, utilizzare Freccia giù, quindi Freccia su per conoscere il primo comando del menu.
- **Menu in cascata.** JAWS non è in grado di leggere il primo livello di un menu in cascata. Se si avverte un'interruzione della lettura mentre si scorre un menu, premere il tasto Freccia destra per conoscere i comandi del menu secondario.

Inoltre, se è stato installato IBM® SPSS® Text Analytics, i seguenti suggerimenti consentono di rendere l'interfaccia workbench interattiva ancora più accessibile.

- **Accedere alle finestre di dialogo.** Potrebbe essere necessario premere il tasto Tab per spostare lo stato attivo sul primo controllo di una finestra di dialogo.
- **Uscire da caselle di testo estese.** Utilizzare Ctrl+Tab per uscire da caselle di testo estese e passare al controllo successivo. *Nota:* Ctrl+Tab consente, inoltre, di uscire dalle tabelle.
- **Digitare la prima lettera per trovare un elemento in un elenco ad albero.** Quando si cerca un elemento nel riquadro delle categorie, nel riquadro dei risultati estratti o nell'albero delle librerie, è possibile digitare la prima lettera dell'elemento quando lo stato attivo si trova nel riquadro. In questo modo verrà selezionata l'occorrenza successiva di un elemento che inizia con la lettera immessa.
- **Elenchi a discesa.** Negli elenchi a discesa delle finestre di dialogo è possibile utilizzare la Barra spaziatrice per selezionare una voce e chiudere l'elenco.

Ulteriori suggerimenti per l'utilizzo vengono discussi ampiamente nei seguenti argomenti.

Interferenze con altri programmi

Durante il test di IBM® SPSS® Modeler con utilità di lettura dello schermo quali JAWS, il team di sviluppo ha scoperto che l'utilizzo di un server SMS (Systems Management Server) nell'organizzazione potrebbe interferire con la capacità di JAWS di leggere applicazioni basate su

Java, quali SPSS Modeler. Per risolvere il problema è sufficiente disattivare SMS. Per ulteriori informazioni su SMS, visitare il sito Web Microsoft.

JAWS e Java

Il livello di supporto per le applicazioni basate su Java è diverso nelle varie versioni di JAWS. Sebbene IBM® SPSS® Modeler supporti tutte le versioni recenti di JAWS, alcune versioni potrebbero presentare problemi di lieve entità se utilizzate in combinazione con sistemi basati su Java. Visitare il sito Web di JAWS per Windows all'indirizzo <http://www.FreedomScientific.com>.

Utilizzo dei grafici in IBM SPSS Modeler

Per le utilità di lettura dello schermo è difficile interpretare le visualizzazioni grafiche delle informazioni, per esempio istogrammi, grafici di valutazione, plot multipli e grafici a dispersione. Si tenga presente, tuttavia, che è possibile visualizzare grafici Web e nodi di distribuzione utilizzando il riepilogo in formato testo disponibile nella finestra di output.

Supporto Unicode

Supporto Unicode in IBM SPSS Modeler

IBM® SPSS® Modeler dispone del supporto completo per Unicode, sia nella versione IBM® SPSS® Modeler che in quella IBM® SPSS® Modeler Server. Consente quindi lo scambio di dati con altre applicazioni che supportano Unicode, inclusi i database multilingue, senza alcun rischio di perdita delle informazioni dovuta alla conversione in o da uno schema di codifica specifico di determinate opzioni internazionali.

- SPSS Modeler archivia i dati Unicode internamente ed è in grado di leggere e scrivere dati multilingue memorizzati nei database in formato Unicode senza alcun rischio di perdita.
- SPSS Modeler può leggere e scrivere file di testo con codifica UTF-8. Le operazioni di importazione ed esportazione di file di testo verranno eseguite di default nella codifica specifica delle opzioni internazionali, ma supporteranno la codifica alternativa UTF-8. Questa impostazione può essere specificata nei nodi di importazione ed esportazione file oppure è possibile modificare la codifica di default nella finestra di dialogo delle proprietà dello stream. [Per ulteriori informazioni, vedere l'argomento Impostazione delle opzioni per gli stream in il capitolo 5 a pag. 51.](#)
- I file di dati Statistics, SAS e di testo archiviati nella codifica specifica delle opzioni internazionali verranno convertiti in UTF-8 durante l'importazione e riconvertiti in fase di esportazione. Se, durante la scrittura in un file, sono presenti caratteri Unicode non disponibili nel set di caratteri specifico delle opzioni internazionali, tali caratteri verranno sostituiti e verrà visualizzato un avviso. Questa situazione dovrebbe verificarsi solo nel caso in cui i dati siano stati importati da una sorgente dati che supporta Unicode (un database o un file di testo UTF-8) e contengano caratteri di opzioni internazionali diverse o di più opzioni internazionali o set di caratteri.
- Le immagini IBM® SPSS® Modeler Solution Publisher sono codificate UTF-8 e possono essere perfettamente trasferite tra piattaforme e opzioni internazionali diverse.

Informazioni su Unicode

L'obiettivo dello standard Unicode consiste nel fornire un metodo uniforme per la codifica di testo multilingue, in modo da garantire una facile condivisione delle informazioni tra confini, opzioni internazionali e applicazioni. Lo standard Unicode, giunto ora alla versione 4.0.1, definisce un set di caratteri che rappresenta un superinsieme di tutti i set di caratteri comunemente utilizzati oggi e assegna a ciascun carattere un nome univoco e un punto di codice. I caratteri e i relativi punti di codice sono identici a quelli UCS (Universal Character Set), definiti dalla norma ISO-10646. Per ulteriori informazioni, fare riferimento alla [Home page Unicode \(http://www.unicode.org\)](http://www.unicode.org).



Note

This information was developed for products and services offered worldwide.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

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

For license inquiries regarding double-byte character set (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

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

Il seguente paragrafo non si applica per il Regno Unito o altri paesi in cui le presenti disposizioni non sono conformi alle leggi locali: SPSS INC., UNA SOCIETÀ IBM, FORNISCE IL PRESENTE DOCUMENTO “COSÌ COM'È” SENZA GARANZIA DI ALCUN TIPO, ESPRESSA O IMPLICITA, COMPRESE, TRA LE ALTRE, LE GARANZIE IMPLICITE DI NON VIOLAZIONE DEI DIRITTI ALTRUI, COMMERCIALIZZABILITÀ O IDONEITÀ PER UNO SCOPO SPECIFICO. Alcuni stati non consentono limitazioni di garanzie espresse o implicite in determinate transazioni, pertanto quanto sopra potrebbe non essere applicabile.

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

Eventuali riferimenti contenuti nelle presenti informazioni a siti Web non SPSS e non IBM vengono forniti per comodità dell'utente e non implicano in alcun modo l'approvazione di tali siti Web. I materiali presenti in tali siti Web non sono parte dei materiali di questo prodotto SPSS Inc. e l'utilizzo di tali siti Web è a rischio esclusivo dell'utente.

Al momento dell'invio di informazioni a IBM o SPSS, l'utente concede a IBM e SPSS un diritto non esclusivo a utilizzare o distribuire tali informazioni nel modo che ritiene appropriato senza incorrere in alcun obbligo nei confronti dell'utente.

Le informazioni relative ai prodotti non SPSS provengono dai fornitori di tali prodotti, dagli annunci pubblicati o da altre fonti pubbliche disponibili. SPSS non ha testato tali prodotti e non può confermare l'accuratezza delle dichiarazioni relative alle prestazioni, alla compatibilità o ad altre caratteristiche relative ai prodotti non SPSS. Le domande relative alle funzionalità dei prodotti non SPSS devono essere indirizzate ai fornitori di tali prodotti.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Software Group, Attention: Licensing, 233 S. Wacker Dr., Chicago, IL 60606, USA.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

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

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

Marchi

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

SPSS è un marchio di SPSS Inc., una società IBM, registrati in numerose giurisdizioni nel mondo.

Adobe, il logo Adobe, PostScript e il logo PostScript sono marchi o marchi registrati di Adobe Systems Incorporated negli Stati Uniti e/o in altri paesi.

IT Infrastructure Library è un marchio registrato della Central Computer and Telecommunications Agency ora parte dell'Office of Government Commerce.

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.

ITIL è un marchio registrato e un marchio comunitario registrato dell'Office of Government Commerce ed è registrato presso l'Ufficio Brevetti degli Stati Uniti.

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

Cell Broadband Engine è un marchio di Sony Computer Entertainment, Inc. negli Stati Uniti e/o in altri paesi ed è utilizzato su licenza.

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

Linear Tape-Open, LTO, the LTO Logo, Ultrium, and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and other countries.

I nomi di altri prodotti o servizi possono essere marchi di IBM, SPSS o altre società.

- accesso facilitato, 250, 263
 - esempio, 259
 - funzionalità in IBM SPSS Modeler, 250
 - suggerimenti in IBM SPSS Modeler, 261
- aggiornamento
 - nodi di input, 51
- aggiornamento modelli, 186, 191–192
- aggiunta
 - a un progetto, 213
- aggiunta di connessioni di IBM SPSS Modeler Server, 11–12
- alberi decisionali
 - accesso facilitato, 261
- analisi basata su alberi
 - applicazioni comuni, 26
- Analisi Campione/Sfidante, 158, 186
- and, operatore, 131
- annotazione
 - nodi, 69, 77
 - stream, 69, 77
- annotazioni
 - cartella, 219
 - conversione in commenti, 77
 - progetto, 218
- annulla, 19
- annullamento della mappatura dei campi, 83
- anteprima
 - dati nodo, 50
- apertura
 - modelli, 82
 - nodi, 82
 - output, 82
 - progetti, 213
 - stati, 82
 - stream, 82
- applicazioni, 26
- applicazioni comuni, 26
- applicazioni del data mining, 27
- Applicazioni predittive, 186
- apprendimento automatico, 25
- archiviazione di oggetti nel IBM SPSS Collaboration and Deployment Services Repository, 163
- area di disegno, 15
- area di disegno dello stream
 - impostazioni, 54
- assegnazione di un nome a nodi e stream, 77
- attivazione nodi , 43
- attribute, 25
- automazione, 97
- avvisi, 56
 - impostazione delle opzioni, 228

- barra degli strumenti, 19
- barra rovesciata nelle espressioni CLEM, 121
- blocco di nodi , 50

- blocco di oggetti del IBM SPSS Collaboration and Deployment Services Repository, 178

- cache
 - attivazione, 47, 227, 245
 - impostazione di una cache, 45
 - opzioni per i nodi, 47, 245
 - salvataggio, 49
 - svuotamento, 49, 51
- cache per nodo
 - attivazione, 47, 245
- cambiamento giorni, 53
- campi, 25, 119, 121
 - in espressioni CLEM, 113
 - visualizzazione di valori, 113
- campi del modello, 87
- campi essenziali, 83, 86
- campi obbligatori, 87
- campi ora
 - conversione , 146
- caratteri, 119, 121, 231
- caratteri speciali
 - rimozione da stringhe, 105
- caricamento
 - nodi, 82
 - stati, 82
- cartelle, IBM SPSS Collaboration and Deployment Services Repository, 177, 179
- caso, 25
- classi , 18, 211, 213
- CLEM, 110
 - controllo di espressioni, 115
 - esempi, 100
 - espressioni, 103, 119
 - funzioni, 112
 - generazione di espressioni, 111
 - introduzione, 24, 97
 - lingua, 119
 - tipi di dati, 120–121
- CLEM espressioni
 - parametri, 57, 104
- CLEM funzioni
 - bitwise, 134
 - casuali, 136
 - confronto, 129
 - datetime, 142
 - funzioni speciali, 154
 - globali, 152
 - informazione, 127
 - logiche, 131
 - numerico, 132
 - probability, 134
 - sequence, 146, 148
 - stringa, 136
 - trigonometriche, 133
 - valori mancanti, 95

- vuoti e nulli, 153
- CLEMfunzioni
 - conversione, 128
 - elenco funzioni disponibili, 125
- client
 - directory di default, 228
- codifica di default dello stream, 53
- codifica di testo, 53
- codifica UTF-8, 53, 264
- colori
 - impostazioni, 231
- comando set, 57, 104
- comma, 51
- commenti
 - elenco di tutti in uno stream, 75
 - scelte rapide da tastiera, 256–257
 - su nodi e stream, 69
- concatenazione di stringhe, 128
- condizioni, 103
- Conformità 508, 250
- connessioni
 - a IBM SPSS Collaboration and Deployment Services Repository, 159–160
 - a IBM SPSS Modeler Server, 9, 11–12
 - cluster di server, 12
- controllo di espressioni CLEM, 115
- convenzioni, 126
- Coordinator of Processes, 12
- COP, 12
- copia, 19
- creazione di palette personalizzate, 239
- creazione di sottopalette, 242
- Creazione guidata applicazioni predittive, 159, 198, 200
 - prerequisiti stream, 199
- CRISP-DM, 18, 211
 - visualizzazione progetti, 212

- data mining, 25
 - esempi di applicazioni, 37
 - strategia, 28
- date
 - conversione, 146
 - gestione, 146
- dati
 - anteprima, 50
- decimali
 - formati di visualizzazione, 53
- default
 - fase del progetto, 212
- deployment di scenari, 186
- descrizioni
 - annotazione di nodi, 77
- descrizioni stream, 64, 67
- directory
 - default, 228
- directory temp, 13
- disattiva nodi , 43, 45

- distribuzione chi-quadrato
 - funzioni di probabilità, 134
- distribuzione *f*
 - funzioni di probabilità, 134
- distribuzione normale
 - funzioni di probabilità, 134
- distribuzione *t*
 - funzioni di probabilità, 134
- diverso da, operatore, 129
- documentazione, 2
- DTD, 207

- elaborazione parallela
 - attivazione, 233
- elenchi, 119, 121
- elenco di tutti i commenti di uno stream, 75
- encoding, 53, 264
- esecuzione degli stream, 67
- esempi
 - cenni generali, 4
 - Guida alle applicazioni, 2
- esempi di applicazioni, 2
- esplorazione
 - scelte rapide da tastiera, 252
- esportazione
 - descrizioni stream, 67
 - PMML, 206, 209
- espressioni, 119
- espressioni CLEM
 - performance, 249
 - ricerca e sostituzione di testo, 115
- etichette
 - value, 207
 - variabile, 207
 - visualizzazione, 54
- etichette di versione, oggetto del IBM SPSS Collaboration and Deployment Services Repository, 184
- etichette, oggetto del IBM SPSS Collaboration and Deployment Services Repository, 184

- fattore, 261
- file di dati testuali
 - encoding, 264
- file di dizionario, 260
- file di output
 - salvataggio, 80
- file di PowerPoint, 213
- file di registro
 - visualizzazione SQL generato, 235
- file di stream di backup
 - ripristino, 79
- file SAS
 - encoding, 264
- Finestra Albero interattivo
 - accesso facilitato, 261
- finestra di dialogo all'avvio, 232

- finestra di dialogo di benvenuto, 232
- finestra principale, 15
- formati di data, 52, 122–123
- formati di ora, 52, 122–123
- formati di visualizzazione
 - decimali, 53
 - numeri, 53
 - scientifico, 53
 - simbolo di raggruppamento, 53
 - valuta, 53
- formati di visualizzazione numeri, 53
- formato di visualizzazione valute, 53
- funzione @BLANK, 95, 127, 153
- funzione @DIFF, 146, 148
- funzione @FIELD, 95, 154
- funzione @FIELDS_BETWEEN, 95, 108, 154
- funzione @FIELDS_MATCHING, 95, 108, 154
- funzione @INDEX, 146, 148
- funzione @LAST_NON_BLANK, 146, 148, 153
- funzione @MAX, 146, 148
- funzione @MEAN, 146, 148
- funzione @MIN, 146, 148
- funzione @MULTI_RESPONSE_SET, 109, 154
- funzione @NULL, 95, 127, 153
- funzione @OFFSET, 146, 148
 - considerazioni sulle performance, 249
- funzione @PARTITION_FIELD, 154
- funzione @PREDICTED, 154
- funzione @SDEV, 146, 148
- funzione @SINCE, 146, 148
- funzione @SUM, 146, 148
- funzione @TARGET, 154
- funzione @TESTING_PARTITION, 154
- funzione @THIS, 146, 148
- funzione @TODAY, 142
- funzione @TRAINING_PARTITION, 154
- funzione @VALIDATION_PARTITION, 154
- funzione abs, 132
- funzione allbutfirst, 136
- funzione allbutlast, 136
- funzione alphabefore, 136
- funzione arccos, 133
- funzione arccosh, 133
- funzione arcsin, 133
- funzione arcsinh, 133
- funzione arctan, 133
- funzione arctan2, 133
- funzione arctanh, 133
- funzione cdf_chisq, 134
- funzione cdf_f, 134
- funzione cdf_normal, 134
- funzione cdf_t, 134
- funzione cos, 133
- funzione cosh, 133
- funzione count_equal, 108, 129
- funzione count_greater_than, 108, 129
- funzione count_less_than, 108, 129
- funzione count_non_nulls, 129
- funzione count_not_equal, 108, 129
- funzione count_nulls, 95, 108, 129
- funzione count_substring, 136
- funzione date_before, 129
- funzione datetime_date, 128
- funzione DIFF, 148
- funzione div, 132
- funzione endstring, 136
- funzione exp, 132
- funzione first_index, 109, 129
- funzione first_non_null, 109, 129
- funzione first_non_null_index, 109, 129
- funzione fracof, 132
- funzione hasendstring, 136
- funzione hasmidstring, 136
- funzione hasstartstring, 136
- funzione INDEX, 148
- funzione integer_bitcount, 134
- funzione integer_leastbit, 134
- funzione integer_length, 134
- funzione intof, 132
- funzione is_date, 127
- funzione is_datetime, 127
- funzione is_integer, 127
- funzione is_number, 127
- funzione is_real, 127
- funzione is_string, 127
- funzione is_time, 127
- funzione is_timestamp, 127
- funzione isalphacode, 136
- funzione isendstring, 136
- funzione islowercode, 136
- funzione ismidstring, 136
- funzione isnumbercode, 136
- funzione isstartstring, 136
- funzione issubstring, 136
- funzione issubstring_count, 136
- funzione issubstring_lim, 136
- funzione isuppercode, 136
- funzione last_index, 109, 129
- funzione LAST_NON_BLANK, 148
- funzione last_non_null, 109, 129
- funzione last_non_null_index, 109, 129
- funzione length, 136
- funzione locchar, 136
- funzione locchar_back, 136
- funzione log, 132
- funzione log10, 132
- funzione lowertoupper, 136
- funzione matches, 136
- funzione max, 129
- funzione MAX, 148
- funzione max_index, 109, 129
- funzione max_n, 108, 129
- funzione MEAN, 146, 148
- funzione mean_n, 108, 132

- funzione member, 129
- funzione min, 129
- funzione MIN, 148
- funzione min_index, 109, 129
- funzione min_n, 108, 129
- funzione mod, 132
- funzione negate, 132
- funzione OFFSET, 148
- funzione oneof, 136
- funzione pi, 133
- funzione potenza (esponenziale), 132
- funzione random, 136
- funzione random0, 136
- funzione rem, 132
- funzione replace, 136
- funzione replicate, 136
- funzione round, 132
- funzione SDEV, 148
- funzione sdev_n, 108, 132
- funzione sign, 132
- funzione sin, 133
- funzione SINCE, 148
- funzione sinh, 133
- funzione skipchar, 136
- funzione skipchar_back, 136
- funzione soundex, 141
- funzione soundex_difference, 141
- funzione sqrt, 132
- funzione startstring, 136
- funzione stripchar, 136
- funzione strmember, 136
- funzione subsers, 136
- funzione substring, 136
- funzione substring_between, 136
- funzione SUM, 148
- funzione sum_n, 108, 132
- funzione tan, 133
- funzione tanh, 133
- funzione testbit, 134
- funzione THIS, 148
- funzione time_before, 129
- funzione to_date, 128, 142
- funzione to_dateline, 142
- funzione to_datetime, 128
- funzione to_integer, 128
- funzione to_number, 128
- funzione to_string, 128
- funzione to_time, 128, 142
- funzione to_timestamp, 128, 142
- funzione trim, 136
- funzione trim_start, 136
- funzione trimend, 136
- funzione undef, 153
- funzione unicode_char, 136
- funzione unicode_value, 136
- funzione uppertolower, 136
- funzione value_at, 109, 129
- funzioni, 122–123, 126–127, 146
 - @BLANK, 94
 - esempi, 100
 - @FIELD, 110, 154
 - gestione valori mancanti, 95
 - @GLOBAL_MAX, 152
 - @GLOBAL_MEAN, 152
 - @GLOBAL_MIN, 152
 - @GLOBAL_SDEV, 152
 - @GLOBAL_SUM, 152
 - in espressioni CLEM, 112
 - @PARTITION, 154
 - @PREDICTED, 110, 154
 - @TARGET, 110, 154
- funzioni bitwise, 134
- funzioni datetime
 - datetime_date, 142
 - datetime_day, 142
 - datetime_day_name, 142
 - datetime_day_short_name, 142
 - datetime_hour, 142
 - datetime_in_seconds, 142
 - datetime_minute, 142
 - datetime_month, 142
 - datetime_month_name, 142
 - datetime_month_short_name, 142
 - datetime_now datetime_second, 142
 - datetime_time , 142
 - datetime_timestamp, 142
 - datetime_weekday , 142
 - datetime_year, 142
- funzioni di confronto, 129
- funzioni di conversione, 128
- funzioni di data, 122–123
 - date_before, 129, 142
 - date_days_difference, 142
 - date_in_days, 142
 - date_in_months, 142
 - date_in_weeks, 142
 - date_in_years, 142
 - date_months_difference, 142
 - date_weeks_difference, 142
 - date_years_difference, 142
 - funzione @TODAY, 142
- funzioni di data e ora, 122–123
- funzioni di distribuzione, 134
- funzioni di informazione, 127
- funzioni di ora, 122–123
 - time_before, 129, 142
 - time_hours_difference, 142
 - time_in_hours, 142
 - time_in_mins, 142
 - time_in_secs, 142
 - time_mins_difference, 142
 - time_secs_difference, 142
- funzioni di probabilità, 134
- funzioni di sequenza, 146, 148

- funzioni globali, 152
- funzioni logiche, 131
- funzioni numeriche, 132
- funzioni speciali, 154
- funzioni stringa, 136
- funzioni trigonometriche, 133
- funzone to_real, 128

- Generatore di espressioni
 - accesso, 111
 - utilizzo, 111
- Generatore espressioni, 255
 - cenni generali, 110
 - ricerca e sostituzione di testo, 115
- Generazione SQL, 233
 - anteprima, 235
 - registrazione, 235
- gestione dei vuoti
 - CLEMfunzioni, 153
- gradi
 - unità di misura, 52
- grafici
 - aggiunta a progetti, 213
 - salvataggio dell'output, 80

- hassubstring, funzione, 136

- IBM InfoSphere Warehouse (ISW)
 - esportazione PMML, 209
- IBM SPSS Collaboration and Deployment Services, 158
- IBM SPSS Collaboration and Deployment Services Enterprise View, 158, 186
- IBM SPSS Collaboration and Deployment Services Repository, 156, 158
 - archiviazione oggetti, 163
 - Blocco e sblocco degli oggetti, 178
 - cartelle, 177, 179
 - connessione a, 159–160
 - eliminazione di oggetti e versioni, 178
 - esplorazione, 161
 - proprietà oggetto, 181
 - recupero di oggetti, 171
 - ricerca in, 174
 - single sign-on, 159
 - trasferimento di progetti in, 216
- IBM SPSS Modeler, 1, 14
 - cenni generali, 8, 226
 - documentazione, 2
 - esecuzione dalla riga di comando, 9
 - funzionalità di accesso facilitato, 250
 - guida introduttiva, 8
 - opzioni, 226
 - suggerimenti e tasti di scelta rapida, 88
- IBM SPSS Modeler Advantage, 158, 184
- IBM SPSS Modeler Server
 - ID utente, 9
 - nome dominio (Windows), 9
 - nome host, 9, 11
 - numero porta, 9, 11
 - password, 9
- IBM SPSS Modeler Server, accesso, 9
- IBM SPSS Text Analytics, 2
- icone
 - impostazione delle opzioni, 54
- ID utente
 - IBM SPSS Modeler Server, 9
- if then else, funzioni, 131
- importazione
 - PMML, 207, 209
- incolla, 19
- insieme di regole
 - valutazione, 51
- insiemi, 51
- insiemi di categorie multiple
 - in espressioni CLEM, 109
- insiemi di dicotomie multiple
 - in espressioni CLEM, 109
- insiemi di modelli, 68
 - definizione, 17
- insiemi di risposte multiple
 - in espressioni CLEM, 109, 113
- interi, 119–120
- internazionali
 - opzioni, 226
- interrompi esecuzione, 19
- introduzione, 119
 - IBM SPSS Modeler, 8, 226

- Java, 263
- JAWS, 250, 260, 262–263

- knowledge discovery, 25

- lettori di schermo, 252, 255, 260–261
 - esempio, 259
- lingua
 - opzioni, 226

- maggiore di, operatore, 129
- manager, 16
- mappatura dei dati, 86
- mappatura di campi, 83
- marchi, 266
- memoria
 - gestione, 226–227
- messaggi
 - visualizzazione SQL generato, 235
- messaggi di errore, 56
- metodo Differenza
 - conversione di confidenze in punteggi di propensione, 200

- modelli Rete neurale, 200
- metodo MaxSfumato
 - conversione di confidenze in punteggi di propensione, 200
 - modelli Rete neurale, 200
- minore di, operatore, 129
- modellazione
 - ramo, 69
- modelli, 68, 84
 - aggiornamento, 191–192
 - aggiunta a progetti, 213
 - archiviazione nel IBM SPSS Collaboration and Deployment Services Repository, 171
 - esportazione, 236
 - sostituzione, 230
- modelli di albero decisionale
 - punteggi di propensione, 200
- modelli di offerta, 200
- modelli di regressione logistica
 - punteggi di propensione, 200
- modelli di valore, 200
- modelli IBM SPSS Statistics, 37
- modelli Insieme di regole
 - punteggi di propensione, 200
- Modelli PMML
 - regressione lineare, 237
 - regressione logistica, 237
- modelli Rete neurale
 - punteggi di propensione, 200
- modelli Statistics, 37
- modello di elaborazione CRISP-DM, 28–29
- mouse
 - utilizzo in IBM SPSS Modeler, 22, 40

- nodi, 8
 - aggiunta, 40, 44
 - aggiunta a progetti, 213–214
 - aggiunta di commenti a, 69
 - annotazione, 69, 77
 - anteprima dei dati, 50
 - archiviazione nel IBM SPSS Collaboration and Deployment Services Repository, 170
 - attivazione, 43
 - blocco, 50
 - caricamento, 82
 - connessione in uno stream, 40
 - creazione di palette personalizzate, 239
 - creazione di sottopalette personalizzate, 242
 - disattivazione, 43, 45
 - disattivazione in uno stream, 43
 - uplicazione, 45
 - eliminazione, 40
 - eliminazione delle connessioni, 45
 - ignorare in uno stream, 42
 - impostazione delle opzioni, 45
 - introduzione, 39
 - modifica, 45
 - ordine dei, 244
 - performance, 247–248
 - personalizzazione delle schede della palette, 243
 - ricerca di, 62
 - rimozione dalla palette, 240
 - salvataggio, 79
 - visualizzazione nella palette, 240
- nodi di input, 39
 - aggiornamento, 51
 - mappatura dei dati, 84
- nodi di processo, 39
 - performance, 247
- nodi Modelli, 30, 39
 - performance, 248
 - personalizzazione della scheda della palette Modelli, 243
- nodi output, 39
- nodi porpora, 233
- nodo Aggregazione
 - performance, 248
- nodo Creazione regola
 - caricamento, 82
- nodo Discretizza
 - performance, 247
- nodo Elimina duplicati
 - performance, 248
- nodo Enterprise View, 186
- nodo Esplora
 - utilizzo nel data mining, 27
 - utilizzo nell'analisi, 25
- nodo File cache
 - caricamento, 82
- nodo finale, 39
- nodo K-Means
 - insiemi grandi, 51
 - performance, 249
- nodo Kohonen
 - insiemi grandi, 51
 - performance, 248
- nodo Ordina
 - performance, 247
- nodo Qualità
 - valori mancanti, 93
- nodo Rete neurale
 - insiemi grandi, 51
 - performance, 248
- nodo Riempimento
 - valori mancanti, 94
- nodo Selezione funzioni
 - valori mancanti, 93
- nodo Tipo
 - performance, 248
 - valori mancanti, 94
- nodo Unione
 - performance, 248
- nodo Valutazione
 - performance, 248

- nome dominio (Windows)
 - IBM SPSS Modeler Server, 9
- nome host
 - IBM SPSS Modeler Server, 9, 11
- nomi di nodo, 77
- nomi di stream, 77
- not, operatore, 131
- notazione scientifica
 - formato di visualizzazione, 53
- note legali, 265
- notifiche
 - impostazione delle opzioni, 228
- novità, 6–7
- nulli, 91, 106
- numeri, 107, 120
- numeri reali, 119–120
- numero porta
 - IBM SPSS Modeler Server, 9, 11

- oggetti
 - proprietà, 220
- oggetti di output
 - archiviazione nel IBM SPSS Collaboration and Deployment Services Repository, 170
- Operatori
 - unione di stringhe, 128
- opzioni, 226
 - display, 231
 - per IBM SPSS Modeler, 226
 - PMML, 236
 - proprietà stream, 51, 54, 56
 - utente, 228
- opzioni di deployment
 - scenari, 186
- opzioni utente, 228
- or, operatore, 131
- ottimizzazione, 233
- output, 16
- output HTML
 - lettore di schermo, 261

- palette, 15
 - personalizzazione, 237
- palette dei modelli generati, 16
- palette Modelli, 171
- parametri
 - creazione di modelli, 189
 - in espressioni CLEM, 113
 - prompt di runtime, 59
 - punteggio, 189
 - sessione, 57, 59–60, 104
 - stream, 57, 59–60, 104
 - type, 60
 - utilizzo in scenari, 189
- parametri sessione, 57, 59–60, 104
- parametri stream, 57, 59–60, 104

- parole chiave
 - annotazione di nodi, 77
- password
 - IBM SPSS Modeler Server, 9
- performance
 - cache per nodo, 47, 245
 - dei nodi di elaborazione, 247
 - dei nodi Modelli, 248
 - espressioni CLEM, 249
- period, 51
- personalizzazione delle schede della palette, 243
- PMML
 - esportazione di modelli, 206, 209
 - importazione di modelli, 207, 209
 - opzioni di esportazione, 236
- precedenza, 123
- precedenza degli operatori, 123
- PredictiveMarketing, 198
- probabilità
 - applicazioni predittive, 200
- progetti, 18, 211
 - aggiunta di oggetti, 214
 - annotazione, 218
 - archiviazione nel IBM SPSS Collaboration and Deployment Services Repository, 169
 - chiusura, 220
 - creazione, 213
 - creazione di nuovi, 214
 - generazione di report, 220
 - impostazione di proprietà, 216
 - impostazione di una cartella di default, 212
 - nel IBM SPSS Collaboration and Deployment Services Repository, 216
 - proprietà cartella, 219
 - proprietà oggetto, 220
 - visualizzazione Classi, 213
 - visualizzazione CRISP-DM, 212
- programmazione visuale, 14
- prompt, di runtime, 59
- proprietà
 - cartella progetto, 219
 - fasi dei report, 220
 - per stream di dati, 51
- proprietà oggetto, IBM SPSS Collaboration and Deployment Services Repository, 181
- pulsante centrale del mouse
 - simulazione, 22, 40
- punteggi di propensione
 - applicazioni predittive, 200
- punteggio
 - ramo, 69, 189–190, 195
- push back, 233

- radianti
 - unità di misura, 52
- rami, creazione di modelli e calcolo del punteggio, 69, 189–190, 195

- record, 25
 - valori mancanti, 93
- recupero di oggetti da IBM SPSS Collaboration and Deployment Services Repository, 171
- regressione, 261
- regressione lineare
 - esportazione come PMML, 237
- regressione logistica, 261
 - esportazione come PMML, 237
- report
 - aggiunta a progetti, 213
 - generazione, 220
 - impostazione di proprietà, 220
 - salvataggio dell'output, 80
- ricerca
 - di nodi in uno stream, 62
- ricerca di connessioni in COP, 12
- ricerca di oggetti nel IBM SPSS Collaboration and Deployment Services Repository, 174
- ricerca di testo, 115
- ridenominazione
 - nodi, 77
 - stream, 64
- ridimensionamento, 21
- riduzione a icona, 21
- riga di comando
 - avvio di IBM SPSS Modeler, 9
- riscrittura degli stream
 - attivazione, 233
- rumore nei dati, 27

- salvataggio
 - nodi, 79
 - oggetti di output, 80
 - più oggetti, 80
 - stati, 79–80
 - stream, 79
- sblocco di oggetti del IBM SPSS Collaboration and Deployment Services Repository, 178
- scelte rapide
 - tastiera, 22, 252, 255–257
 - utilizzo generale, 88
- scelte rapide da tastiera, 252, 255–257
- scenari, 184
 - definizione, 158
 - opzioni di deployment, 186
- scorrimento
 - impostazione delle opzioni, 54
- script, 24, 97
 - ricerca e sostituzione di testo, 115
- selezione della palette dei nodi, 240
- server
 - accesso a, 9
 - aggiunta di connessioni, 11
 - directory di default, 228
 - ricerca di server in COP, 12
- sessioni multiple di IBM SPSS Modeler, 14
- simbolo decimale, 51
 - formati di visualizzazione numeri, 53
- simbolo di raggruppamento
 - formati di visualizzazione numeri, 53
- single sign-on, 11, 160
- single sign-on, IBM SPSS Collaboration and Deployment Services Repository, 156, 159
- sistema
 - opzioni, 226
- sostituzione di modelli, 230
- sostituzione di testo, 115
- sottopalette
 - creazione, 242
 - rimozione dalla scheda della palette, 241
 - visualizzazione nella scheda della palette, 241
- sovrapposte, 25
- spazi
 - rimozione da stringhe, 105, 136
- spazio bianco
 - rimozione da stringhe, 105, 136
- SPSS Modeler Server, 1
- stampa, 23
 - stream, 45
- stati
 - caricamento, 82
 - salvataggio, 79–80
- Statisticsfile
 - encoding, 264
- stream, 8, 15
 - aggiunta a progetti, 213–214
 - aggiunta di commenti, 69
 - aggiunta di nodi, 40, 44
 - annotazione, 69, 77
 - archiviazione nel IBM SPSS Collaboration and Deployment Services Repository, 169
 - caricamento, 82
 - connessione di nodi, 40
 - creazione, 38
 - disattivazione nodi, 43
 - esecuzione, 67
 - file di backup, 79
 - ignorare i nodi, 42
 - opzioni, 51
 - opzioni di deployment, 186
 - ridenominazione, 64, 77
 - salvataggio, 79
- stream di dati
 - creazione, 38
- stringhe, 119, 121
 - corrispondenza, 105
 - manipolazione in espressioni CLEM, 105
 - sostituzione, 105
- strumento per la mappatura dei dati, 83–84
- suggerimenti
 - per l'accesso facilitato, 261
 - utilizzo generale, 88

Indice

Supernodo

parametri, 57, 104

Supporto Unicode, 264

tabelle, 255

aggiunta a progetti, 213

salvataggio dell'output, 80

taglia, 19

tasti di scelta rapida, 22

tipi di dati, 102

nei parametri, 60

tipo di deployment, 186

uguale a, operatore, 129

valori, 102

aggiunta a espressioni CLEM, 113

visualizzazione da un report del nodo Esplora, 113

valori data/ora, 107

valori globali

in espressioni CLEM, 113

valori mancanti, 92–93, 106

espressioni CLEM, 95

gestione, 91

in record, 93

riempimento, 91

valori mancanti definiti dall'utente, 91

valori mancanti di sistema, 91

vuoti, 91–92, 106

Zoom indietro da un Supernodo, 19