

Guida rapida di IBM SPSS Statistics 20



Nota: Prima di utilizzare queste informazioni e il relativo prodotto, leggere le informazioni generali disponibili in Note legali a pag. 162.

Questa versione si applica a IBM® SPSS® Statistics 20 e a tutte le successive versioni e modifiche fino a eventuali disposizioni contrarie indicate in nuove versioni.

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

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

Materiali concessi in licenza - Proprietà di IBM

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

Tutti i diritti riservati.

Prefazione

Nella *Guida rapida di IBM SPSS Statistics 20* viene fornita una vasta gamma di esercitazioni che consentono all'utente di acquisire familiarità con i vari componenti di IBM® SPSS® Statistics. Questa guida è stata creata per essere utilizzata con tutte le versioni dei sistemi operativi per il software, inclusi: Windows, Macintosh e Linux. È possibile eseguire le esercitazioni sia seguendo l'ordine predefinito sia passando direttamente agli argomenti sui quali si desiderano maggiori informazioni. È possibile utilizzare questa guida come documentazione aggiuntiva rispetto all'esercitazione in linea fornita con il modulo Core SPSS Statistics oppure ignorare l'esercitazione in linea e iniziare dalle esercitazioni riportate di seguito.

IBM SPSS Statistics 20

IBM® SPSS® Statistics 20 è un sistema completo per l'analisi dei dati. SPSS Statistics è in grado di elaborare dati provenienti da quasi tutti i tipi di file e di utilizzarli per generare rapporti con tabelle, grafici, grafici di distribuzioni e trend, statistiche descrittive e complesse analisi statistiche.

SPSS Statistics rende l'analisi statistica accessibile a qualsiasi utente e ancora più pratica per l'utente esperto. I menu e le finestre di dialogo consentono di eseguire analisi complesse senza dover digitare una singola riga di sintassi. L'Editor dei dati offre uno strumento simile a un foglio di calcolo semplice ed efficiente per l'immissione di dati e la visualizzazione dei file di dati di lavoro.

Risorse Internet

Nel sito Web di IBM Corp. (<http://www.ibm.com/support>) sono disponibili risposte alle domande frequenti e l'accesso a file di dati e ad altre informazioni utili.

Il gruppo di discussione SPSS USENET (non sponsorizzato da IBM Corp.) è aperto a chiunque sia interessato. L'indirizzo del gruppo è *comp.soft-sys.stat.spss*.

È inoltre possibile sottoscrivere l'elenco dei messaggi di posta elettronica del gruppo USENET. Per la sottoscrizione, inviare un messaggio di posta elettronica all'indirizzo *listserv@listserv.uga.edu*. Il testo del messaggio dovrà essere il seguente: *subscribe SPSSX-L nome cognome*. Sarà quindi possibile pubblicare i messaggi nell'elenco inviando un messaggio di posta elettronica all'indirizzo *listserv@listserv.uga.edu*.

Pubblicazioni aggiuntive

Presso Prentice Hall è disponibile il volume *IBM SPSS Statistics Statistical Procedures Companion* di Marija Norusis, che include una panoramica delle procedure di IBM® SPSS® Statistics Base e inoltre della regressione logistica e dei modelli lineari generalizzati. Prentice Hall ha pubblicato inoltre il volume *IBM SPSS Statistics Advanced Statistical Procedures Companion* che contiene una panoramica delle procedure nei moduli Advanced e Regression.

IBM SPSS Statistics Opzioni

Le seguenti opzioni sono disponibili come componenti avanzati aggiuntivi per il sistema completo IBM® SPSS® Statistics (non per la versione per studenti):

Statistics Base fornisce un'ampia gamma di procedure statistiche per analisi e report di base, inclusi conteggi, tavole di contingenza e statistiche descrittive, cubi OLAP e report con informazioni sui dati. Offre inoltre un'ampia varietà di tecniche di riduzione, classificazione e segmentazione di dimensioni, quali l'analisi fattoriale, l'analisi cluster, l'analisi del vicino più vicino e l'analisi della funzione discriminante. In più, SPSS Statistics Base propone un'ampia gamma di algoritmi per il confronto delle medie, nonché tecniche predittive quali il test t, l'analisi della varianza, la regressione lineare e la regressione ordinale.

Advanced Statistics offre tecniche spesso utilizzate nella ricerca biomedica e sperimentale avanzata. Include procedure per modelli lineari generalizzati (GLM), modelli misti lineari, analisi della varianza, analisi loglineare, regressione ordinale, tavole di sopravvivenza attuariali, analisi della sopravvivenza di Kaplan-Meier e regressione di Cox di base ed estesa.

Il bootstrap è un metodo utilizzato per derivare delle stime affidabili su errori standard e intervalli di confidenza per stime quali media, mediana, proporzione, rapporto odd, coefficiente di correlazione o coefficiente di regressione.

Categories consente di eseguire procedure di scaling ottimale, inclusa l'analisi delle corrispondenze.

Complex Samples consente ai ricercatori che eseguono sondaggi, indagini di mercato, ricerche in ambito medico o dell'opinione pubblica, nonché agli scienziati che si avvalgono della metodologia basata sui sondaggi campione, di incorporare nell'analisi dei dati i disegni di campioni complessi creati.

Conjoint consente di misurare in modo realistico l'influenza degli attributi di prodotti individuali sulle preferenze dei consumatori e dei cittadini. Mediante Conjoint, è possibile misurare in modo semplice l'effetto dello scambio di ogni attributo di prodotti nel contesto di un insieme di attributi di prodotti —e verificare la reazione dei consumatori all'atto di compiere una decisione di acquisto.

Custom Tables consente di creare una vasta gamma di rapporti con tabelle, incluse le tabelle complesse con intestazioni e colonne, e consente di visualizzare dati a risposta multipla.

Data Preparation fornisce un'istantanea immediata dei dati. In particolare consente di applicare le regole di convalida che identificano i valori di dati non validi. È possibile creare regole che escludono i valori fuori intervallo, i valori mancanti o vuoti. È possibile anche salvare le variabili che registrano violazioni alle singole regole nonché il numero totale di violazioni alle regole per caso. Viene fornito un set limitato di regole predefinite da copiare o modificare.

Decision Trees crea un modello di classificazione basato su alberi. Classifica i casi in gruppi o prevede i valori di una variabile dipendente (di destinazione) in base ai valori di variabili (predittore) indipendenti. La procedura offre strumenti di validazione per l'analisi di classificazione confermativa ed esplorativa.

Direct marketing consente alle organizzazioni di realizzare campagne il più possibile efficaci attraverso tecniche specificamente ideate per il direct marketing.

Exact Tests consente di calcolare valori p esatti nei test statistici, soprattutto quando, a causa di campioni piccoli o mal distribuiti, i test comuni potrebbero risultare non corretti. Questa opzione è disponibile solo nei sistemi operativi Windows.

Forecasting consente di eseguire previsioni complete e analisi di serie storiche con più modelli per la stima delle curve, modelli di livellamento e metodi per la stima delle funzioni di autoregressione.

Missing Values descrive i modelli dei dati mancanti, fornisce una stima delle medie e di altre statistiche e assegna i valori per le osservazioni mancanti.

Neural Networks può essere utilizzato per prendere decisioni aziendali tramite la previsione della domanda relativa a un prodotto in funzione del prezzo e di altre variabili oppure classificando i clienti in base alle abitudini di acquisto e alle caratteristiche demografiche. Le reti neurali sono strumenti per la creazione di modelli basati su dati non lineari. Possono essere utilizzate per creare modelli di relazioni complesse tra gli input e gli output o per individuare motivi nei dati.

Regression fornisce tecniche per l'analisi dei dati che non rientrano nei modelli statistici lineari tradizionali. Include procedure per l'analisi Probit, la regressione logistica, il metodo dei minimi quadrati ponderati, la regressione con il metodo dei minimi quadrati ponderati e la regressione non lineare generalizzata.

Amos™ (analysis of moment structures) utilizza modelli di equazioni strutturali per confermare e spiegare modelli concettuali che riguardano gli atteggiamenti, le percezioni e altri fattori comportamentali.

Corsi di formazione

IBM Corp. organizza corsi di formazione su IBM® SPSS® Statistics che includono esercitazioni pratiche. I corsi verranno offerti periodicamente nelle maggiori città statunitensi ed europee. Per ulteriori informazioni su questi seminari, andare a <http://www.ibm.com/software/analytics/spss/training/>.

Supporto tecnico

Ai clienti che richiedono la manutenzione, viene messo a disposizione un servizio di supporto tecnico. I clienti possono contattare il supporto tecnico per richiedere assistenza per l'utilizzo dei prodotti IBM Corp. o per l'installazione di uno degli ambienti hardware supportati. Per contattare il supporto tecnico, visitare il sito Web di 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.

IBM SPSS Statistics 20 Versione per studenti

La versione per studenti di IBM® SPSS® Statistics 20 è una versione più limitata ma ugualmente efficace di SPSS Statistics.

Caratteristiche

La versione per studenti contiene molti importanti strumenti per l'analisi dei dati forniti con IBM® SPSS® Statistics, compresi:

- Editor dei dati analogo a un foglio di calcolo per l'inserimento, la modifica e la visualizzazione di file di dati.
- Procedure statistiche, inclusi test t , analisi della varianza e tavole di contingenza.
- Grafici interattivi che consentono di modificare o aggiungere elementi di grafici in modo dinamico. Le modifiche vengono visualizzate nel momento stesso in cui sono specificate.
- Grafici standard ad alta risoluzione per una matrice completa di tabelle e grafici di presentazione e analitici.

Limitazioni

La versione per studenti, progettata come complemento alle attività scolastiche, può essere utilizzata solo da studenti e insegnanti e solo per scopi didattici. La versione per studenti di IBM® SPSS® Statistics 20 è soggetta alle limitazioni seguenti:

- I file di dati possono contenere al massimo 50 variabili.
- I file di dati non possono contenere più di 1.500 casi. Non sono disponibili i moduli SPSS Statistics aggiuntivi, quali Regression o Advanced Statistics.
- Non è disponibile la sintassi dei comandi SPSS Statistics. Ciò significa che non è possibile ripetere un'analisi salvando una serie di comandi in un blocco di sintassi o file "di lavoro", a differenza di quanto avviene nella versione completa di IBM® SPSS® Statistics.
- Non sono disponibili le caratteristiche di script e automazione. Ciò significa che non è possibile creare script per l'automazione delle attività più frequenti, a differenza di quanto avviene nella versione completa di SPSS Statistics.

Supporto tecnico per studenti

Gli studenti che utilizzano una versione accademica o grad pack di qualsiasi prodotto software IBM SPSS sono pregati di utilizzare le apposite pagine online per studenti [Solutions for Education](http://www.ibm.com/spss/rd/students/) (<http://www.ibm.com/spss/rd/students/>). Gli studenti che utilizzano una copia del software IBM SPSS fornita dall'università, sono pregati di contattare il coordinatore del prodotto IBM SPSS presso l'università.

Supporto tecnico per insegnanti

Gli insegnanti che utilizzano la versione per studenti in ambito scolastico, possono contattare il supporto tecnico per assistenza. Negli Stati Uniti e in Canada, telefonare al supporto tecnico al numero (312) 651-3410, o visitare il sito <http://www.ibm.com/support>.

Contenuto

1	Introduzione	1
	File di esempio	1
	Apertura di un file di dati.	1
	Esecuzione di un'analisi	3
	Visualizzazione di risultati.	7
	Creazione di grafici.	8
2	Lettura dei dati	10
	Struttura di base dei file di dati di IBM SPSS Statistics	10
	Lettura dei file di dati di IBM SPSS Statistics	10
	Lettura dei dati di un foglio elettronico	11
	Lettura dei dati di un database	13
	Lettura dei dati di un file di testo	19
3	Utilizzo dell'Editor dei dati	27
	Inserimento di dati numerici	27
	Inserimento di dati stringa	30
	Definizione di dati.	32
	Aggiunta delle etichette delle variabili	32
	Modifica del formato e del tipo di variabile.	33
	Aggiunta di etichette dei valori per variabili numeriche	33
	Aggiunta di etichette dei valori per variabili stringa	35
	Utilizzo di etichette dei valori definite per l'inserimento dei dati	36
	Gestione dei dati mancanti	37
	Valori mancanti per una variabile numerica	38
	Valori mancanti per una variabile stringa	40
	Copiare e incollare gli attributi delle variabili	41
	Definizione delle proprietà delle variabili categoriali	44
4	Utilizzo di più sorgenti dati	51
	Gestione di più sorgenti di dati	51
	Utilizzo di più insiemi di dati nella sintassi dei comandi	52
	Copia e inserimento di informazioni in più insiemi di dati.	53

Ridenominazione degli insiemi di dati	53
Soppressione di più insiemi di dati	54
5 <i>Esame delle statistiche riassuntive per singole variabili</i>	55
Livello di misurazione	55
Misure riassuntive per dati categoriali	55
Grafici per i dati categoriali	57
Misure riassuntive per variabili di scala	58
Istogrammi per variabili di scala	60
6 <i>Creazione e modifica di grafici</i>	62
Nozioni base per la creazione di grafici.	62
Utilizzo di Modelli del generatore di grafici.	63
Definizione delle variabili e delle statistiche	64
Aggiunta di testo	67
Per creare il grafico.	68
Nozioni base per la modifica di grafici	68
Selezione degli elementi dei grafici	69
Utilizzo della finestra Proprietà	70
Modifica dei colori delle barre	70
Formattazione dei numeri nelle etichette delle tacche	72
Modifica di testo	74
Visualizzazione delle etichette dei valori	75
Utilizzo dei modelli	76
Definizione delle opzioni del grafico.	81
7 <i>Utilizzo dell'output</i>	85
Utilizzo del Viewer	85
Utilizzo dell'Editor di tabelle pivot	86
Accesso alle definizioni dell'output	87
Tabelle pivot	87
Creazione e visualizzazione di strati	89
Modifica di tabelle.	90
Nascondere righe e colonne	91
Modifica dei formati di visualizzazione dei dati.	92

Modelli di tabelle	93
Utilizzo di formati predefiniti	93
Personalizzazione di stili TableLook	94
Modifica dei formati di tabella predefiniti	97
Personalizzazione delle impostazioni iniziali di visualizzazione	98
Visualizzazione delle etichette di valori e variabili	100
Utilizzo dell'output in altre applicazioni	102
Incollare i risultati come tabelle Word	102
Incollare i risultati come testo	103
Esportazione dei risultati in file di Microsoft Word, PowerPoint ed Excel	105
Esportazione dei risultati in formato PDF	112
Esportazione dei risultati in formato HTML	115
8 Utilizzo della sintassi	116
Incollare la sintassi	116
Modificare la sintassi	117
Apertura ed esecuzione di un file di sintassi	119
Informazioni sul riquadro degli errori	120
Utilizzo dei punti di interruzione	120
9 Modifica dei valori dei dati	123
Creazione di una variabile categoriale da una variabile di scala	123
Calcolo di nuove variabili	129
Utilizzo delle funzioni nelle espressioni	131
Utilizzo delle espressioni logiche	133
Utilizzo di date e orari	135
Calcolo del periodo di tempo tra due date	136
Aggiunta di una durata a una data	139
10 Ordinamento e selezione di dati	143
Ordinamento di dati	143
Distinzione per gruppi	144
Ordinamento dei casi per la distinzione per gruppi	145
Attivazione e disattivazione della distinzione per gruppi	146

Selezione di sottoinsiemi di casi	146
Per selezionare sottoinsiemi di casi in base a un'espressione logica	147
Per selezionare un campione casuale di dati	148
Selezione di un intervallo di tempo o di casi	149
Trattamento dei casi non selezionati	150
Criteri di selezione dei casi	150

Appendici

<i>A File di esempio</i>	152
---------------------------------	------------

<i>B Note legali</i>	162
-----------------------------	------------

<i>Indice</i>	165
----------------------	------------

Introduzione

Nella presente guida viene fornita una vasta gamma di esercitazioni che ha lo scopo di illustrare come eseguire utili analisi dei dati. È possibile eseguire le esercitazioni sia seguendo l'ordine predefinito sia passando direttamente agli argomenti sui quali si desiderano maggiori informazioni.

In questo capitolo vengono descritte le funzioni di base e viene esemplificata una sessione tipica. Verrà recuperato un file di dati di IBM® SPSS® Statistics precedentemente definito e in seguito verranno generati un riepilogo statistico semplice e un grafico.

Ulteriori istruzioni su molti argomenti accennati in questo capitolo si trovano nei capitoli successivi. Le istruzioni che seguono rappresentano un valido aiuto per la comprensione delle esercitazioni successive.

File di esempio

La maggior parte degli esempi illustrati utilizzano il file di dati *demo.sav*. Si tratta di un file di dati contenente un sondaggio fittizio realizzato su un campione di un migliaio di persone con informazioni di base di tipo demografico e consumistico.

Se si sta utilizzando la versione per studenti, il file *demo.sav* è un campione rappresentativo del file di dati originale, ridotto per soddisfare il limite di 1.500 casi. I risultati che si otterranno utilizzando tale file di dati saranno diversi da quelli mostrati qui.

Il file di esempio installato con il prodotto si trova nella sottodirectory *Samples* della directory di installazione. La sottodirectory *Samples* contiene cartelle separate per ciascuna delle seguenti lingue: Inglese, Francese, Tedesco, Italiano, Giapponese, Coreano, Polacco, Russo, Cinese semplificato, Spagnolo e Cinese tradizionale.

Non tutti i file di esempio sono disponibili in tutte le lingue. Se un file di esempio non è disponibile in una lingua, la cartella di tale lingua contiene una versione inglese del file.

Apertura di un file di dati

Per aprire un file di dati:

- ▶ Dai menu, scegliere:
File > Apri > Dati...

È anche possibile utilizzare il pulsante Apri file nella barra degli strumenti.

Figura 1-1

Pulsante Apri File della barra degli strumenti



Viene visualizzata una finestra di dialogo per l'apertura dei file.

Per impostazione predefinita, vengono visualizzati i file di dati in formato IBM® SPSS® Statistics (con estensione *.sav*).

In questo esempio viene utilizzato il file *demo.sav*.

Figura 1-2

File *demo.sav* nell'Editor dei dati

	età	statciv	indirizz	reddito	fascredd	au
1	55	Stato civile	12	72,00	3,00	
2	56	0	29	153,00	4,00	
3	28	1	9	28,00	2,00	
4	24	1	4	26,00	2,00	
5	25	0	2	23,00	1,00	
6	45	1	9	76,00	4,00	
7	42	0	19	40,00	2,00	
8	35	0	15	57,00	3,00	

Il file di dati viene visualizzato nell'Editor dei dati. Nell'Editor dei dati, se si posiziona il cursore del mouse sul nome di una variabile (le intestazioni di colonna), verrà visualizzata un'etichetta di variabile più descrittiva, se presente.

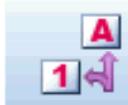
Per impostazione predefinita, vengono visualizzati i valori effettivi dei dati. Per visualizzare le etichette:

- Dai menu, scegliere:
Visualizza > Etichette di valore

È anche possibile utilizzare il pulsante Etichette dei valori della barra degli strumenti.

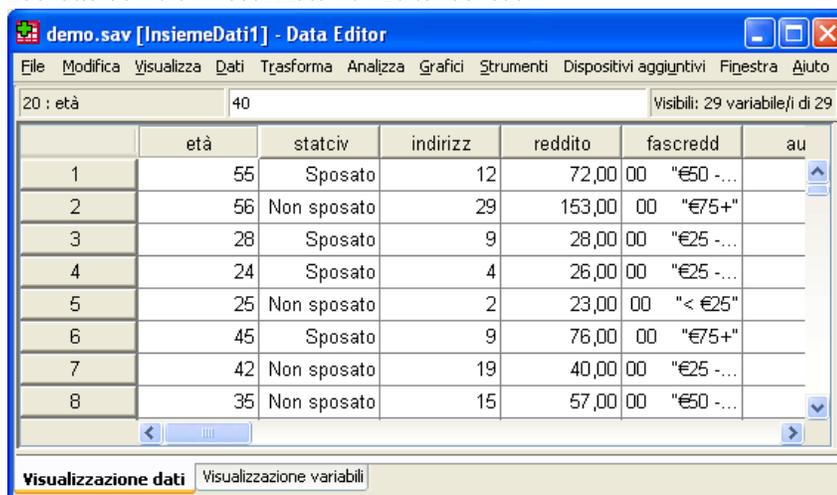
Figura 1-3

Pulsante Etichette dei valori



A questo punto vengono visualizzate le etichette dei valori descrittive per semplificare l'interpretazione delle risposte.

Figura 1-4
Etichette dei valori visualizzate nell'Editor dei dati



The screenshot shows the SPSS Data Editor window for a file named 'demo.sav [InsiemeDati1]'. The window title bar includes standard OS controls. The menu bar contains: File, Modifica, Visualizza, Dati, Trasforma, Analizza, Grafici, Strumenti, Dispositivi aggiuntivi, Finestra, Aiuto. The main area displays a data table with 8 rows and 7 columns. The columns are labeled: età, statciv, indirizz, reddito, fascredd, and au. The 'reddito' column contains numerical values with two decimal places. The 'fascredd' column contains categorical labels representing income brackets. The status bar at the bottom indicates 'Visualizzazione dati' and 'Visualizzazione variabili'.

	età	staciv	indirizz	reddito	fascredd	au
1	55	Sposato	12	72,00 00	"€50 - ..."	
2	56	Non sposato	29	153,00 00	"€75+"	
3	28	Sposato	9	28,00 00	"€25 - ..."	
4	24	Sposato	4	26,00 00	"€25 - ..."	
5	25	Non sposato	2	23,00 00	"< €25"	
6	45	Sposato	9	76,00 00	"€75+"	
7	42	Non sposato	19	40,00 00	"€25 - ..."	
8	35	Non sposato	15	57,00 00	"€50 - ..."	

Esecuzione di un'analisi

Se si dispone di moduli aggiuntivi, il menu Analizza contiene un elenco di categorie per l'analisi statistica e la preparazione di report.

Il primo esempio si riferisce a una tabella di frequenza (tabella di conteggi). Questo esempio richiede il modulo Statistics Base.

- Dai menu, scegliere:
Analizza > Statistiche descrittive > Frequenze...

Si visualizza la finestra di dialogo Frequenze.

Figura 1-5
Finestra di dialogo Frequenze

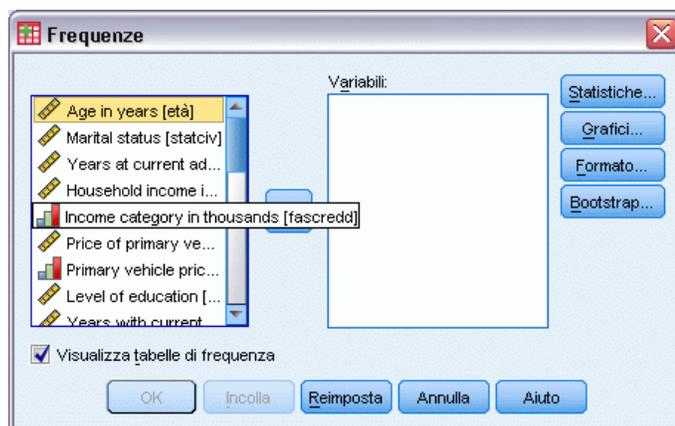


L'icona accanto a ogni variabile consente di ottenere le informazioni sul tipo di dati e sul livello di misurazione.

	Numerico	Stringa	Data	Ora
Scala (continuo)		n/d		
Ordinale				
Nominale				

- Fare clic sulla variabile *Categoria di reddito in migliaia [catredd]*.

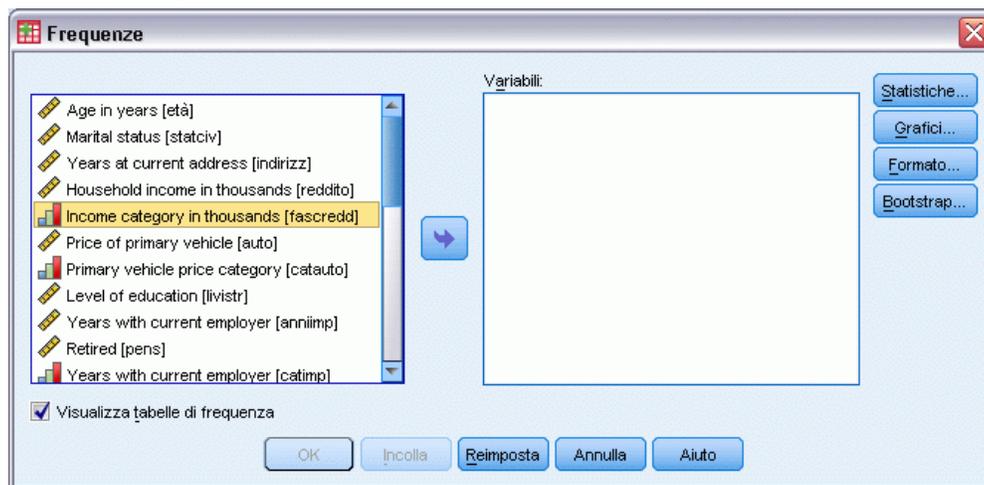
Figura 1-6
Nomi ed etichette delle variabili nella finestra di dialogo Frequenze



Se l'etichetta e/o il nome della variabile appare troncato nell'elenco, la dicitura completa viene visualizzata quando il cursore si sofferma sull'etichetta/sul nome. Il nome di variabile *catredd* viene visualizzato tra parentesi quadre dopo l'etichetta di variabile descrittiva. *Categoria reddito in migliaia* è l'etichetta della variabile. Se non fosse presente alcuna etichetta di variabile, nella casella di riepilogo verrebbe visualizzato solo il nome della variabile.

È possibile ridimensionare le finestre di dialogo proprio come per le finestre, selezionando e trascinando i bordi o gli angoli esterni. Ad esempio, se la finestra di dialogo viene allargata, anche gli elenchi delle variabili risulteranno più ampi.

Figura 1-7
Finestra di dialogo ridimensionata

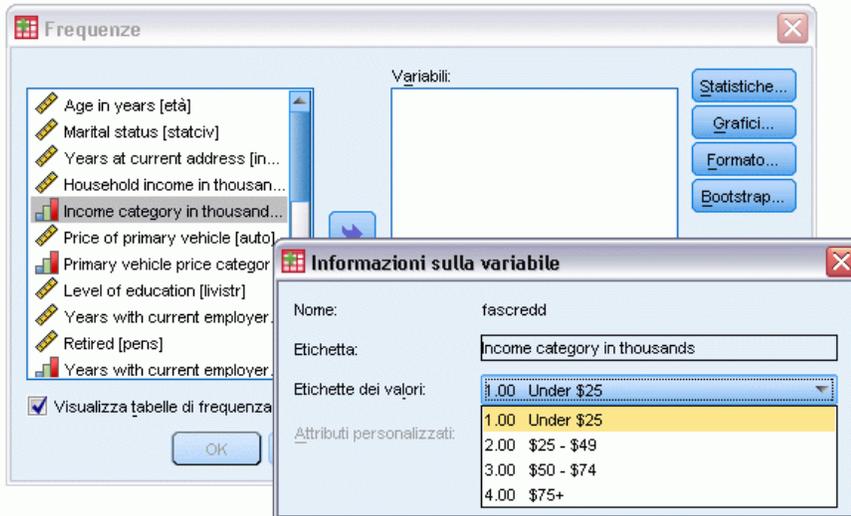


Nella finestra di dialogo, scegliere le variabili che si desidera analizzare dall'elenco di origine a sinistra e trascinarle nell'elenco Variabili a destra. Il pulsante OK che esegue l'analisi viene disabilitato finché almeno una variabile non viene inserita nell'elenco Variabili.

In molte finestre di dialogo, è possibile ottenere ulteriori informazioni facendo clic sul nome della variabile nell'elenco con il pulsante destro del mouse.

- ▶ fare clic con il pulsante destro del mouse su *Categoria di reddito in migliaia [catredd]* e scegliere informazioni sulle variabili.
- ▶ Fare clic sulla freccia rivolta verso il basso nell'elenco a discesa Etichette dei valori.

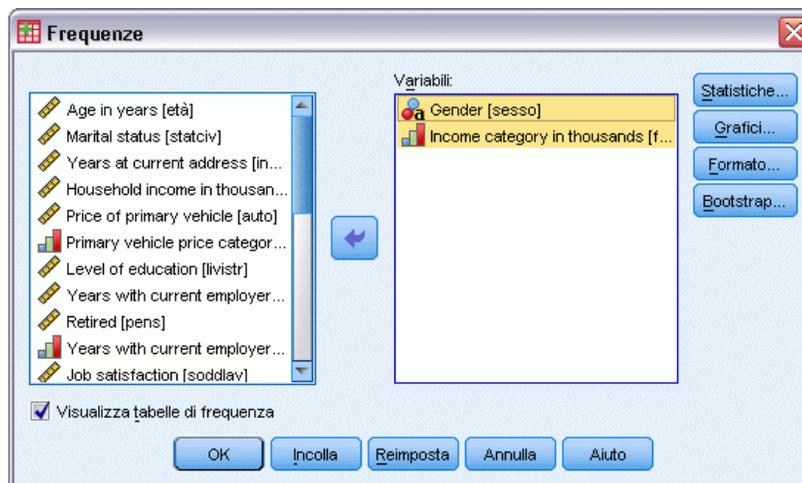
Figura 1-8
Etichette definite per la variabile reddito



Vengono visualizzate tutte le etichette dei valori definite per la variabile.

- ▶ Fare clic su *Sesso [sesso]* nell'elenco delle variabili sorgente e trascinare la variabile nell'elenco Variabili di destinazione.
- ▶ Quindi fare clic su *Categoria di reddito in migliaia [catredd]* nell'elenco sorgente e trascinarla nell'elenco di destinazione.

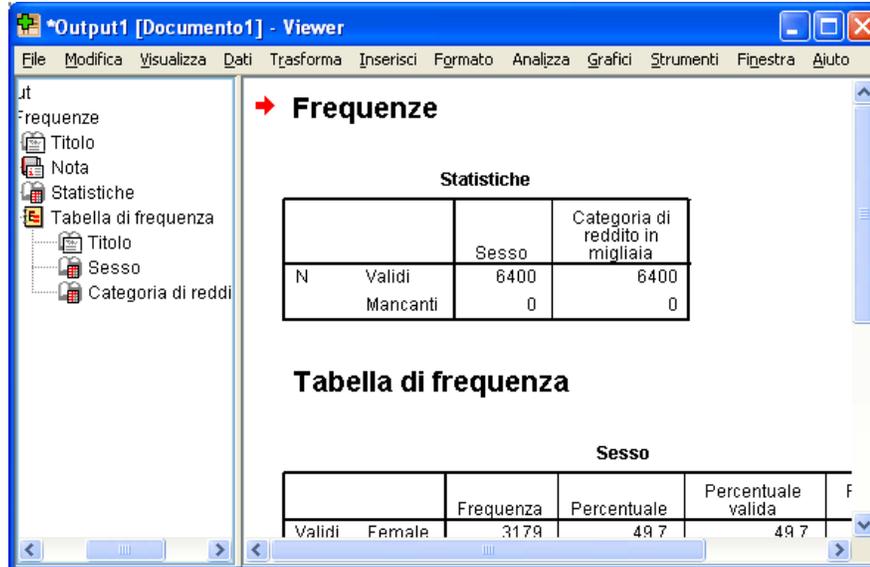
Figura 1-9
Variabili selezionate per l'analisi



- ▶ Fare clic su OK per eseguire la procedura.

Visualizzazione di risultati

Figura 1-10
Finestra Viewer

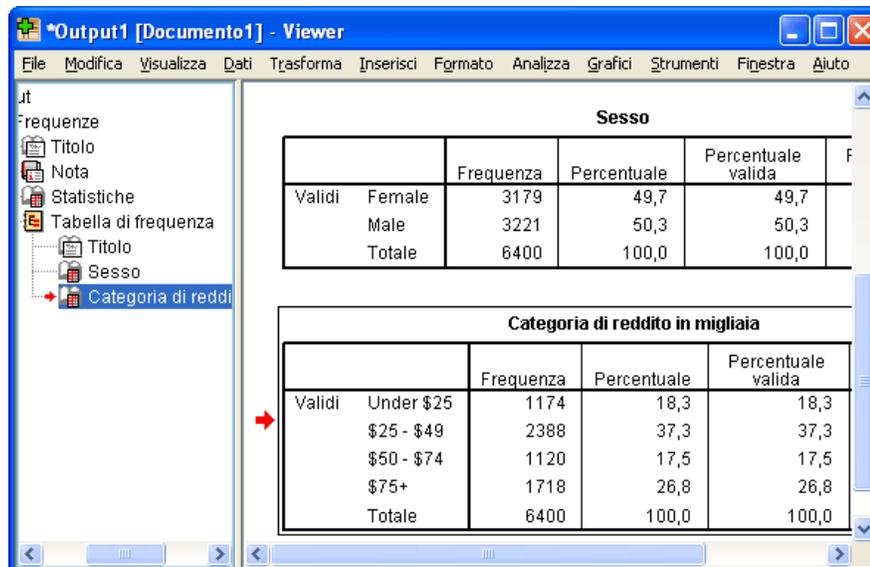


I risultati vengono visualizzati nel Viewer.

Per aprire qualsiasi elemento del Viewer, selezionarlo dalla vista riassuntiva.

- Fare clic su *Categoria di reddito in migliaia* [catredd].

Figura 1-11
Tabella di frequenza per le categorie di reddito



Viene visualizzata la tabella di frequenza relativa alle categorie di reddito. La tabella di frequenza mostra il numero e la percentuale di persone in ogni categoria di reddito.

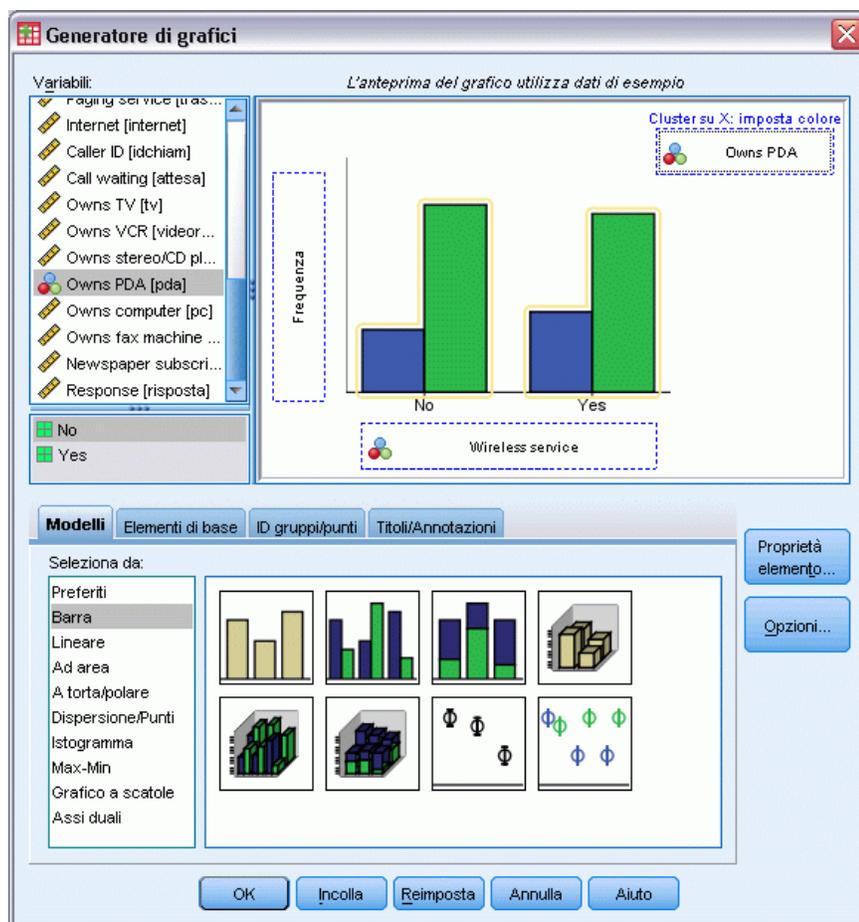
Creazione di grafici

Sebbene alcune procedure statistiche possano creare grafici, è anche possibile utilizzare il menu Grafici per crearli.

Ad esempio, è possibile creare un grafico che mostra la relazione esistente tra titolarità di servizi telefonici senza fili e PDA (Personal Digital Assistant).

- ▶ Dai menu, scegliere:
Grafici > Generatore di grafici...
- ▶ Fare clic sulla scheda Modelli (se non è selezionata).
- ▶ Fare clic su A barre (se non è selezionata).
- ▶ Trascinare l'icona A barre raggruppato nel disegno, ovvero nell'area estesa situata sopra i Modelli.

Figura 1-12
Finestra di dialogo Generatore di grafici

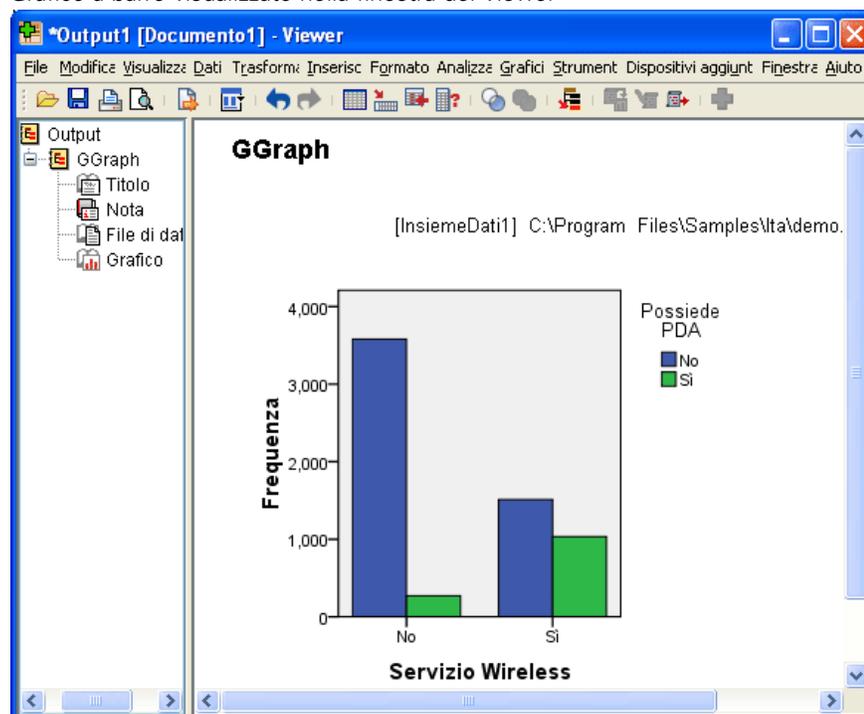


- ▶ Scorrere verso il basso l'elenco Variabili, fare clic con il pulsante destro del mouse su *Telefono cellulare [cellul]*, quindi selezionare Nominale come rispettivo livello di misurazione.

- ▶ Trascinare la variabile *Telefono cellulare [cellul]* nell'asse *x*.
- ▶ Fare clic su *Possiede PDA [pda]* e selezionare *Nominale* come rispettivo livello di misurazione.
- ▶ Trascinare la variabile *Possiede PDA [pda]* nell'area di rilascio di raggruppamento situata nell'angolo destro superiore del disegno.
- ▶ E fare clic su *OK* per creare il grafico.

Figura 1-13

Grafico a barre visualizzato nella finestra del Viewer



Il grafico a barre viene visualizzato nel Viewer. Il grafico mostra che chi possiede un servizio di telefonia senza fili è più propenso ad avere un PDA rispetto a chi non lo possiede.

È possibile modificare grafici e tabelle facendo doppio clic su di essi nel riquadro del contenuto della finestra Viewer, quindi è possibile copiare e incollare i risultati in altre applicazioni. Questi argomenti verranno trattati in seguito.

Lettura dei dati

È possibile inserire i dati direttamente oppure importarli da sorgenti diverse. In questo capitolo vengono descritte le procedure per la lettura dei dati in file dati di IBM® SPSS® Statistics, in fogli di calcolo quali Microsoft Excel, in applicazioni database quali Microsoft Access e nei file di testo.

Struttura di base dei file di dati di IBM SPSS Statistics

Figura 2-1
Editor dei dati

	età	statciv	indirizz	reddito	fascredd
1	55	1	12	72,00	3,00
2	56	0	29	153,00	4,00
3	28	1	9	28,00	2,00
4	24	1	4	26,00	2,00
5	25	0	2	23,00	1,00
6	45	1	9	76,00	4,00
7	42	0	19	40,00	2,00
8	35	0	15	57,00	3,00

I file di dati in formato IBM® SPSS® Statistics sono organizzati per casi (righe) e per variabili (colonne). In questo file di dati, i casi rappresentano i singoli intervistati di un sondaggio, mentre le variabili rappresentano le risposte fornite a ogni quesito del sondaggio.

Lettura dei file di dati di IBM SPSS Statistics

I file di dati di IBM® SPSS® Statistics hanno estensione *.sav* e contengono i dati salvati. Per aprire il file di esempio *demo.sav*, installato con il prodotto:

- ▶ Dai menu, scegliere:
File > Apri > Dati...
- ▶ Spostarsi e aprire il file *demo.sav*. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice A a pag. 152.](#)

I dati vengono visualizzati nell'Editor dei dati.

Figura 2-2

File di dati aperto

	età	statciv	indirizz	reddito	fascredd
1	55	1	12	72,00	3,00
2	56	0	29	153,00	4,00
3	28	1	9	28,00	2,00
4	24	1	4	26,00	2,00
5	25	0	2	23,00	1,00
6	45	1	9	76,00	4,00
7	42	0	19	40,00	2,00
8	35	0	15	57,00	3,00

Lettura dei dati di un foglio elettronico

Anziché digitare direttamente i dati nell'Editor dei dati, è possibile leggere i dati da applicazioni quali Microsoft Excel. È inoltre possibile leggere le intestazioni delle colonne come nomi di variabili.

- ▶ Dai menu, scegliere:
File > Apri > Dati...
- ▶ Selezionare Excel (*.xls) come tipo di file da visualizzare.
- ▶ Aprire il file *demo.xls*. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice A a pag. 152.](#)

Viene visualizzata la finestra di dialogo Apertura di database Excel, nella quale è possibile specificare se nel foglio di calcolo sono inclusi o meno i nomi delle variabili e inoltre le celle che si desidera importare. In Excel 95 o versione successiva, è inoltre possibile specificare i fogli di lavoro che si desidera importare.

Figura 2-3
Finestra di dialogo Apertura di database Excel



- Assicurarsi che l'opzione *Leggi i nomi delle variabili dalla prima riga di dati* sia selezionata. Questa opzione consente di leggere le intestazioni delle colonne come nomi delle variabili.

Se le intestazioni delle colonne non sono conformi alle regole per la denominazione delle variabili IBM® SPSS® Statistics, vengono convertite in nomi di variabili validi e le intestazioni di colonna originali vengono salvate come etichette delle variabili. Per importare solo una parte di un foglio di calcolo, specificare l'intervallo di celle che si desidera importare nella casella di testo *Intervallo*.

- Fare clic su *OK* per leggere il file di Excel.

I dati vengono visualizzati nell'Editor dei dati, con le intestazioni delle colonne utilizzate come nomi di variabili. Poiché i nomi delle variabili non possono contenere degli spazi, sono stati rimossi gli spazi dalle intestazioni delle colonne originali. Per esempio, *Stato civile* nel file Excel, diventa la variabile *statociv*. L'intestazione di colonna originale viene mantenuta come etichetta della variabile.

Figura 2-4
Dati di Excel importati

	Age	Maritalstatus	Address	Income	IncomeCategory
1	55	1	12	72,00	3,00
2	56	0	29	153,00	4,00
3	28	1	9	28,00	2,00
4	24	1	4	26,00	2,00
5	25	1	2	23,00	1,00
6	45	0	9	76,00	4,00

Visualizzazione dati Visualizzazione variabili

Il processore è pronto

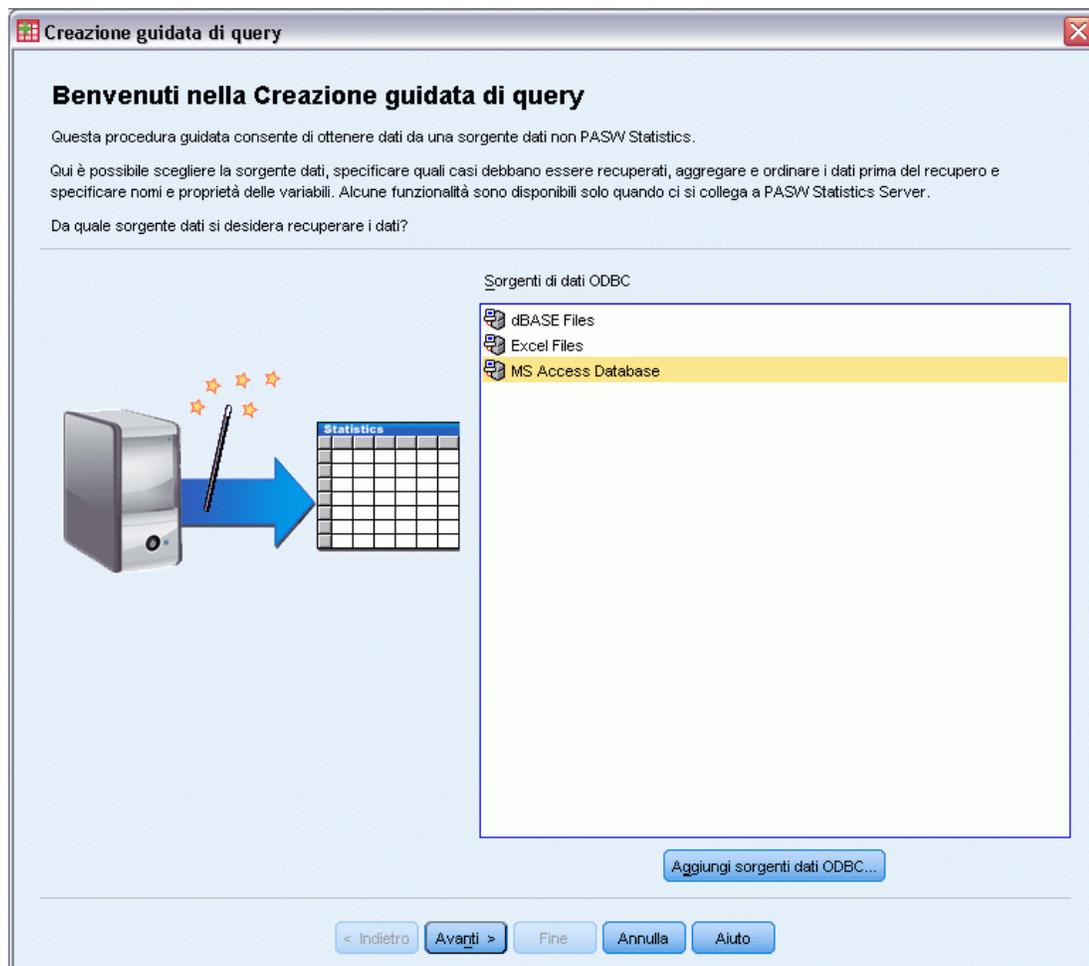
Letture dei dati di un database

Creazione guidata di query semplifica l'importazione dei dati da un database. I database che utilizzano driver ODBC (Open Database Connectivity) possono essere letti direttamente dopo l'installazione dei driver. Nel CD di installazione sono disponibili i driver ODBC per numerosi formati di database. Per ottenere driver aggiuntivi, è possibile rivolgersi a fornitori di terze parti. In questo esempio viene utilizzata una delle applicazioni di database più comuni, ovvero Microsoft Access.

Nota: Questo esempio è specifico per Microsoft Windows e richiede un driver ODBC per Access. La procedura è simile a quella di altre piattaforme, ma potrebbe richiedere un driver ODBC di terze parti per Access.

- Dai menu, scegliere:
File > Apri database > Nuova query...

Figura 2-5
Finestra di dialogo di benvenuto di Creazione guidata di query

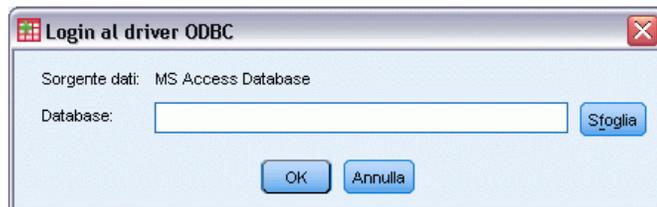


- Selezionare Database di Microsoft Access nell'elenco delle sorgenti dati e fare clic su Avanti.

Nota: A seconda dell'installazione, è possibile che sul lato sinistro della procedura guidata (solo sistemi operativi Windows) venga visualizzato un elenco di sorgenti dati OLEDB. In questo esempio, viene però utilizzato l'elenco della sorgente dati ODBC, visualizzato a destra.

Figura 2-6

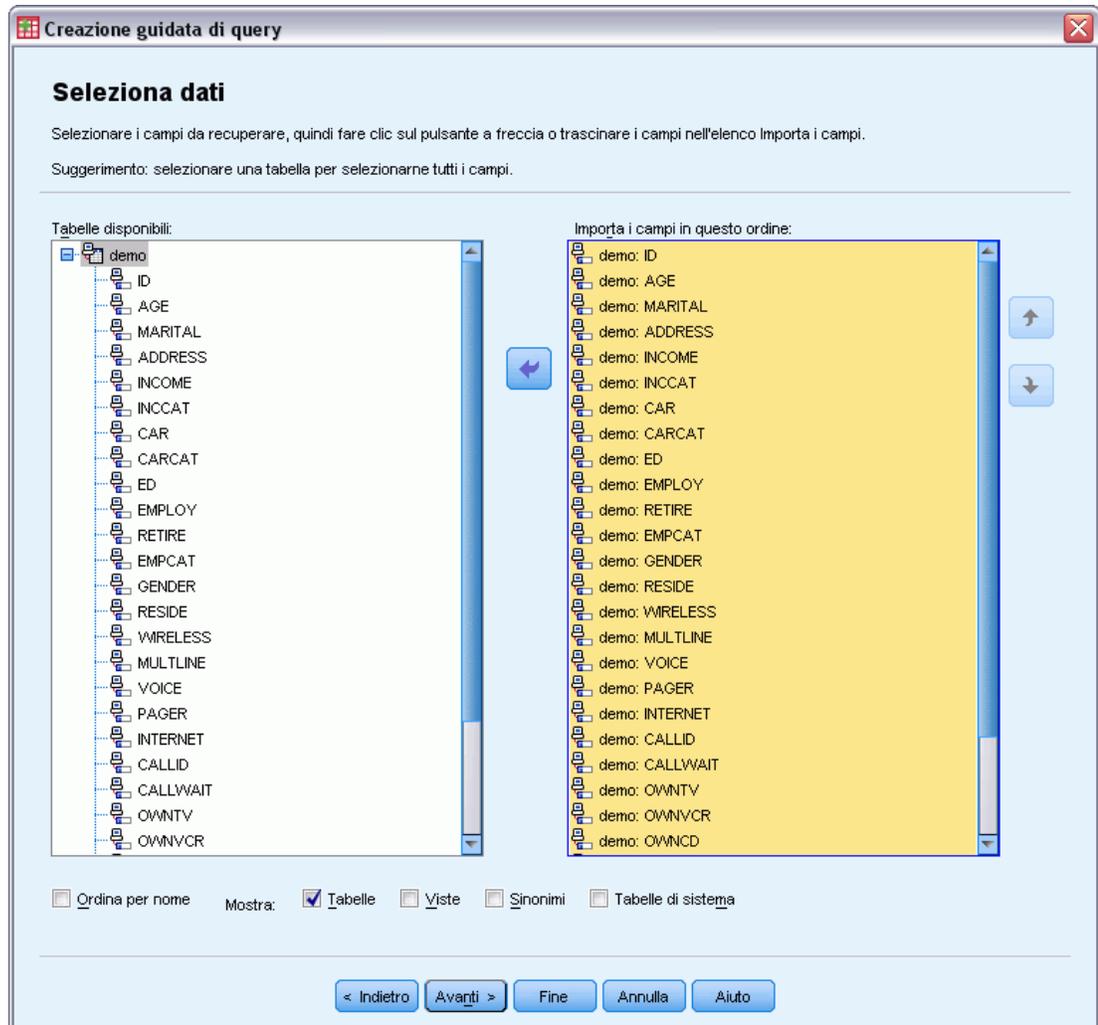
Finestra di dialogo Login al driver ODBC



- ▶ Fare clic su Sfoglia per spostarsi sul file del database di Access che si desidera aprire.
- ▶ Aprire il file *demo.mdb*. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice A a pag. 152.](#)
- ▶ Fare clic su OK nella finestra di dialogo di login.

Nel passo successivo è possibile specificare le tabelle e le variabili che si desidera importare.

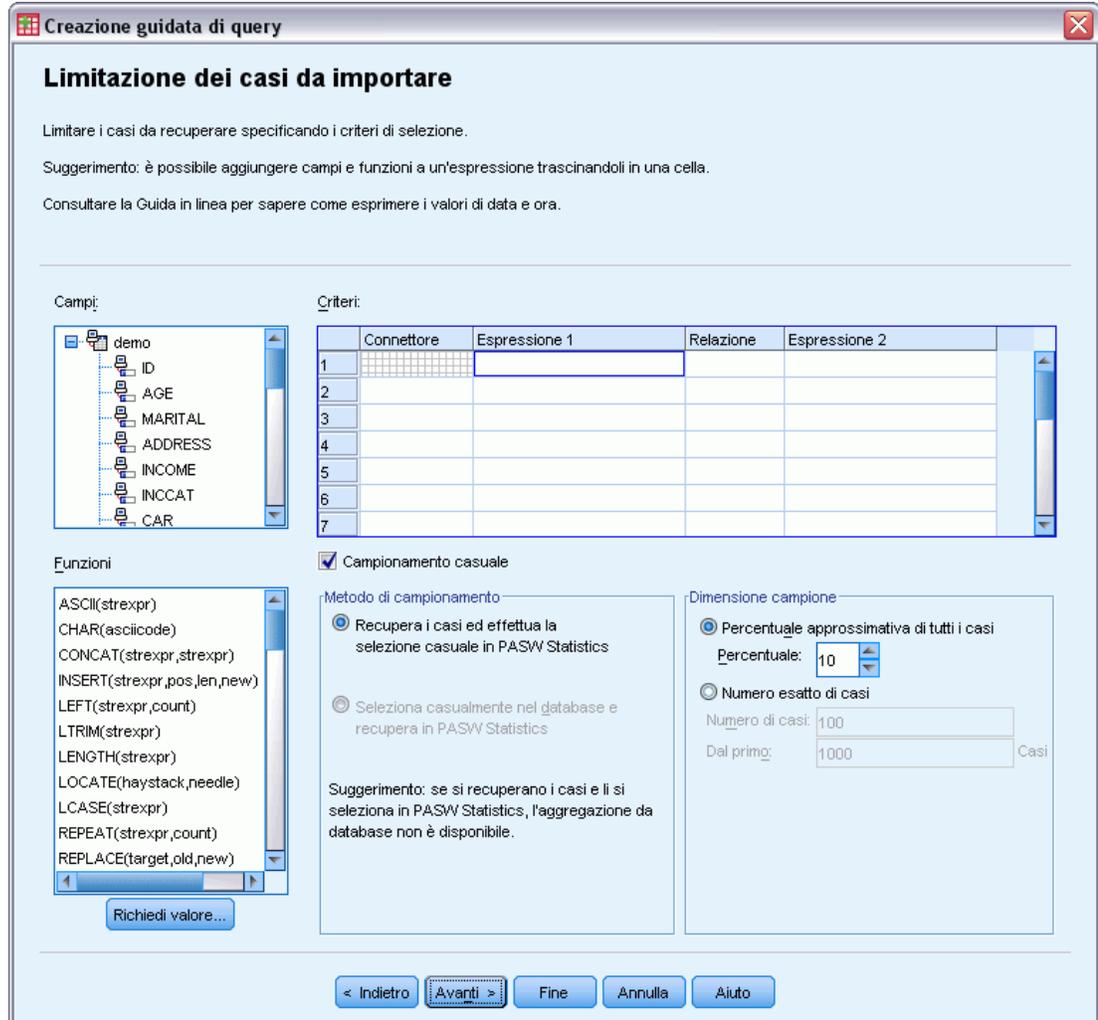
Figura 2-7
Passo Seleziona dati



- ▶ Trascinare l'intera tabella della demo nell'elenco Importa i campi in questo ordine.
- ▶ Fare clic su Avanti.

Nel passo successivo è possibile selezionare i record (casi) da importare.

Figura 2-8
Passo Limitazione dei casi da importare



Se non si desidera importare tutti i casi, è possibile importare un sottoinsieme di casi, ad esempio i maschi con età superiore a 30 anni, oppure un campione casuale di casi dalla sorgente dati. Per le sorgenti dati di grandi dimensioni, è possibile ridurre i tempi di elaborazione limitando il numero dei casi a un campione piccolo e rappresentativo.

- Fare clic su Avanti per continuare.

I nomi dei campi vengono utilizzati per creare nomi di variabili. Se necessario, i nomi vengono convertiti in nomi di variabili validi. I nomi dei campi originali vengono conservati come etichette delle variabili. È possibile, inoltre, modificare i nomi delle variabili prima di importare il database.

Figura 2-9
Passo Definisci variabili

Creazione guidata di query

Definizione delle variabili

Modificare i nomi e le proprietà delle variabili.

Se non si specificano i nomi, questi verranno generati automaticamente.

Per convertire una variabile stringa in variabile numerica utilizzando i valori originali come etichette dei valori, selezionare Ricodifica in numerica.

	Nome della variabile	Tipo di dati	Ricodifica in Numerica
demo: ID	ID	Numerico	
demo: AGE	AGE	Numerico	
demo: MARITAL	MARITAL	Numerico	
demo: ADDRESS	ADDRESS	Numerico	
demo: INCOME	INCOME	Numerico	
demo: INCCAT	INCCAT	Numerico	
demo: CAR	CAR	Numerico	
demo: CARCAT	CARCAT	Numerico	
demo: ED	ED	Numerico	
demo: EMPLOY	EMPLOY	Numerico	
demo: RETIRE	RETIRE	Numerico	
demo: EMPCAT	EMPCAT	Numerico	
demo: GENDER	GENDER	Stringa	<input checked="" type="checkbox"/>
demo: RESIDE	RESIDE	Numerico	
demo: WIRELESS	WIRELESS	Numerico	
demo: MULTLINE	MULTLINE	Numerico	
demo: VOICE	VOICE	Numerico	
demo: PAGER	PAGER	Numerico	
demo: INTERNET	INTERNET	Numerico	
demo: CALLID	CALLID	Numerico	
demo: CALLWAIT	CALLWAIT	Numerico	
demo: OWNTV	OWNTV	Numerico	
demo: OWNVCR	OWNVCR	Numerico	
demo: OWNCD	OWNCD	Numerico	
demo: OWNPDA	OWNPDA	Numerico	
demo: OWNPC	OWNPC	Numerico	

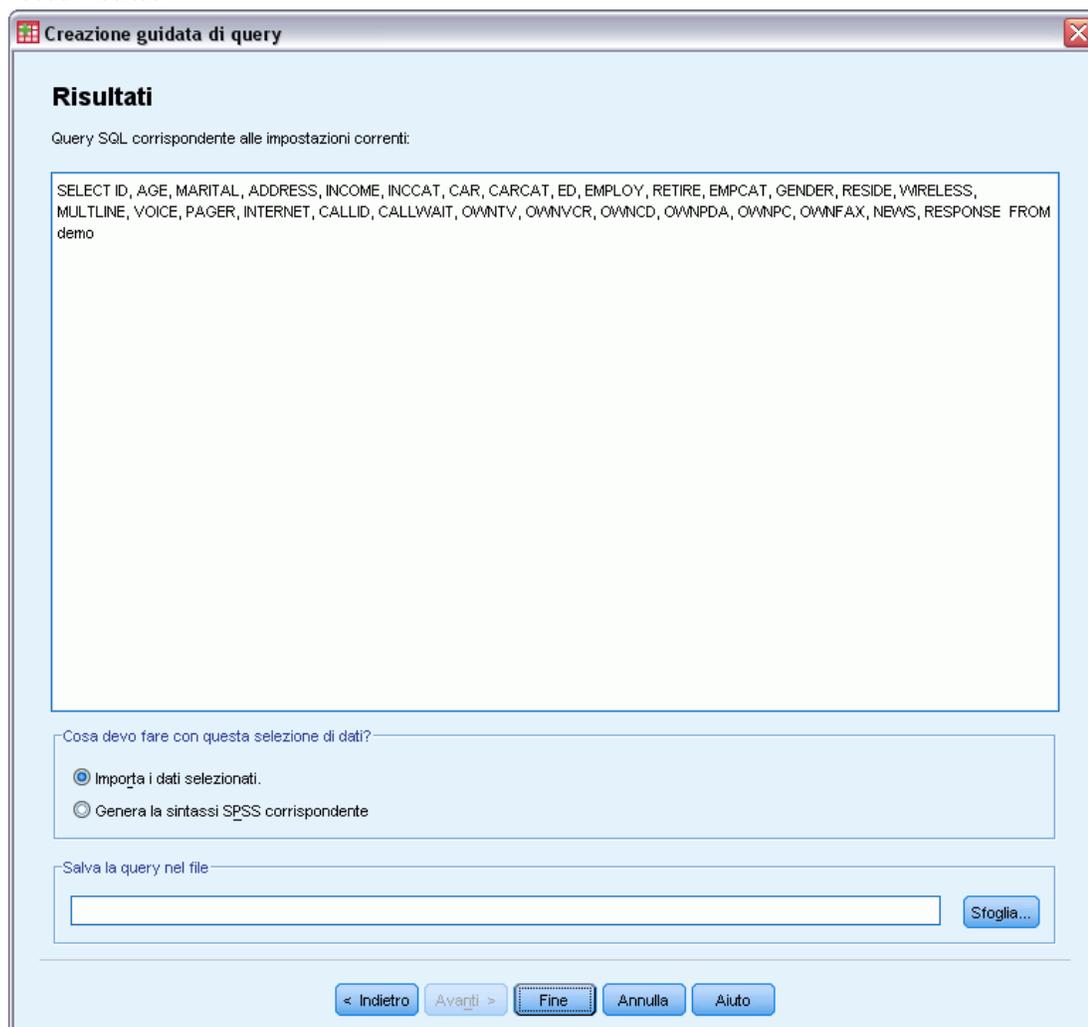
Larghezza per i campi stringa di larghezza variabile: Minimizza le larghezze delle stringhe sulla base dei valori osservati

< Indietro **Avanti >** Fine Annulla Aiuto

- Fare clic sulla cella Ricodifica in numerica nel campo Sesso. Questa opzione converte le variabili stringa in variabili intero e mantiene il valore originale dell'etichetta per la nuova variabile.
- Fare clic su Avanti per continuare.

L'istruzione SQL creata in base alle selezioni in Autocomposizione dati viene visualizzata nel passo Risultati e può essere eseguita immediatamente oppure può essere salvata in un file per utilizzarla in seguito.

Figura 2-10
Passo Risultati



- Fare clic su Fine per importare i dati.

Tutti i dati del database di Access selezionati per l'importazione sono ora disponibili nell'Editor dei dati.

Figura 2-11
Dati importati da un database di Access

	Age	Maritalstatus	Address	Income	IncomeCategory
1	55	1	12	72,00	3,00
2	56	0	29	153,00	4,00
3	28	1	9	28,00	2,00
4	24	1	4	26,00	2,00
5	25	1	2	23,00	1,00
6	45	0	9	76,00	4,00

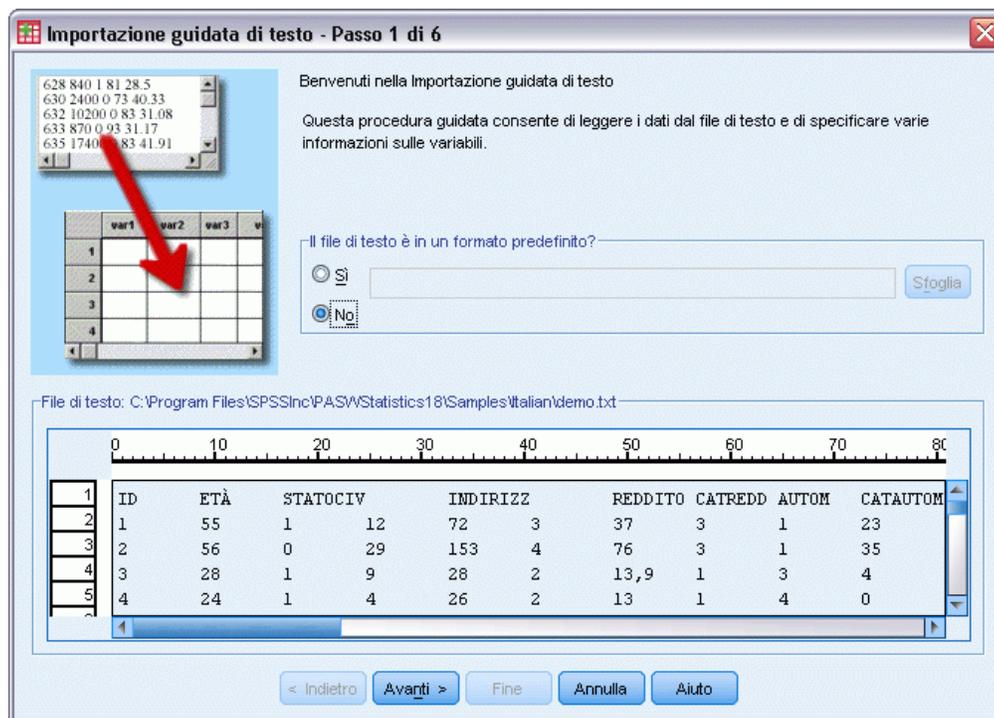
Lettura dei dati di un file di testo

I file di testo sono un'altra sorgente dati di uso comune. Molti fogli di calcolo e database sono in grado di salvare i contenuti in numerosi formati di file di testo. Nei file delimitati da virgole o da tabulazioni, viene fatto riferimento a righe di dati che utilizzano le virgole o le tabulazioni per indicare ciascuna variabile. In questo esempio, i dati sono delimitati da tabulazioni.

- ▶ Dai menu, scegliere:
File > Leggi dati testo...
- ▶ Selezionare Testo (*.xls) come tipo di file da visualizzare.
- ▶ Aprire il file *demo.txt*. Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice A a pag. 152.

Importazione guidata di testo consente di definire la modalità di interpretazione del file di testo specificato.

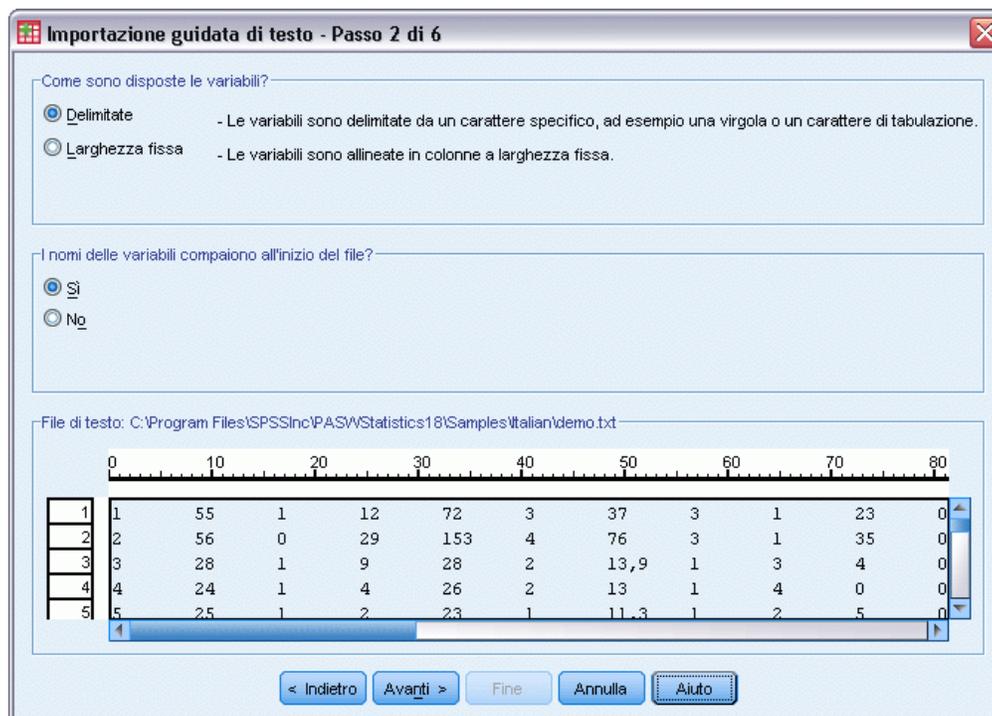
Figura 2-12
Importazione guidata di testo - Passo 1 di 6



- ▶ Nel Passo 1, è possibile scegliere un formato predefinito oppure crearne uno nuovo. Selezionare No per specificare che verrà creato un nuovo formato.
- ▶ Fare clic su Avanti per continuare.

Come indicato in precedenza, in questo file viene utilizzato un formato delimitato da tabulazioni. I nomi delle variabili vengono inoltre definiti nella riga iniziale del file.

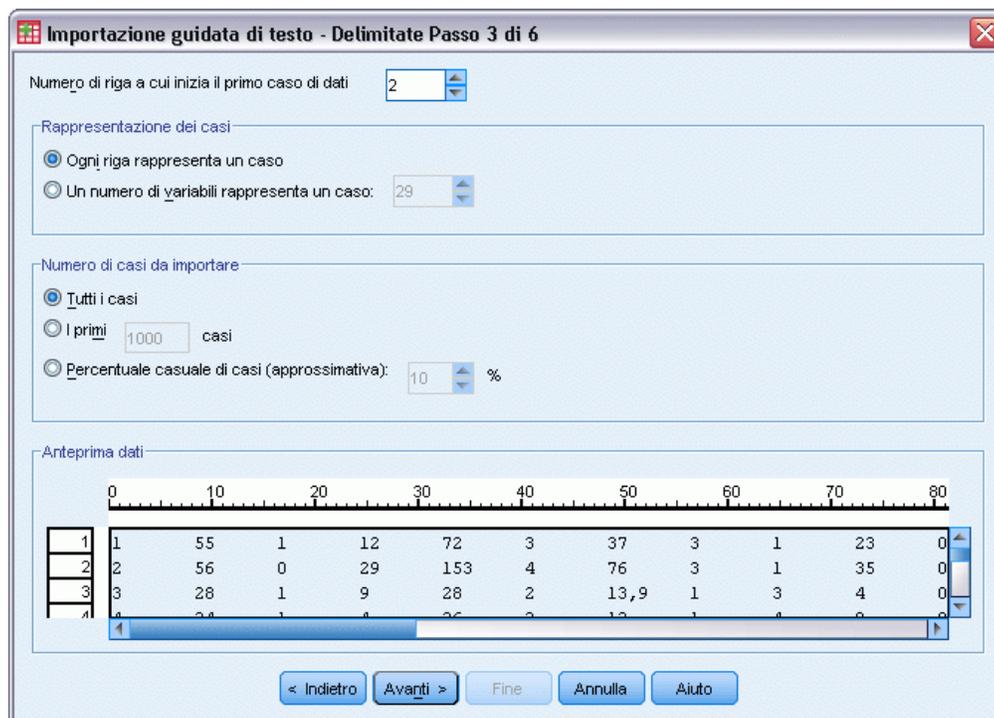
Figura 2-13
Importazione guidata di testo - Passo 2 di 6



- ▶ Selezionare Delimitate per specificare che i dati utilizzano una struttura di formattazione delimitata.
- ▶ Selezionare Sì per specificare che i nomi delle variabili vanno letti dalla parte iniziale del file.
- ▶ Fare clic su Avanti per continuare.

- Digitare 2 nella parte superiore della finestra di dialogo successiva, per specificare che la prima riga di dati inizia nella seconda riga del file di testo.

Figura 2-14

Importazione guidata di testo - Passo 3 di 6

- Mantenere i valori predefiniti per la parte rimanente di questa finestra di dialogo e fare clic su Avanti per continuare.

In Anteprima dati (passo 4) è possibile verificare in modo rapido che i dati vengano letti correttamente.

Figura 2-15
Importazione guidata di testo - Passo 4 di 6

Delimitatori tra le variabili

Tabulatore Spazio
 Virgola Punto e virgola
 Altro:

Qual è il qualificatore di testo?

Nessuno
 Virgolette doppie
 Virgolette singole
 Altro:

Anteprima dati

ID	ETÀ	STATOCIV	INDIRIZZ	REDDITO	CATREDD	AUTOM	CATAUTOM	K
1	55	1	12	72	3	37	3	1
2	56	0	29	153	4	76	3	1
3	28	1	9	28	2	13,9	1	3
4	24	1	4	26	2	13	1	4
5	25	1	2	23	1	11,3	1	2
6	45	0	9	76	4	37,3	3	3
7	44	1	17	144	4	72,1	3	2
8	46	1	20	75	4	37,1	3	1
9	41	0	10	56	2	12	1	1

< Indietro Avanti > Fine... Annulla Aiuto

- ▶ Selezionare Tabulazione e deselegionare le altre opzioni.
- ▶ Fare clic su Avanti per continuare.

È possibile che i nomi delle variabili vengano troncati per adattarli ai requisiti di formattazione e in questa finestra di dialogo viene offerta la possibilità di modificarli in base alle specifiche esigenze.

Figura 2-16
Importazione guidata di testo - Passo 5 di 6

Nome variabile: ID Nome originale: ID

Formato dati: Numerica

ID	ETÀ	STATOCIV	INDIRIZZ	REDDITO	CATREDD	AUTOM	CATAUTOM	IS
1	55	1	12	72	3	37	3	1
2	56	0	29	153	4	76	3	1
3	28	1	9	28	2	13,9	1	3
4	24	1	4	26	2	13	1	4
5	25	1	2	23	1	11,3	1	2

< Indietro Avanti > Fine Annulla Aiuto

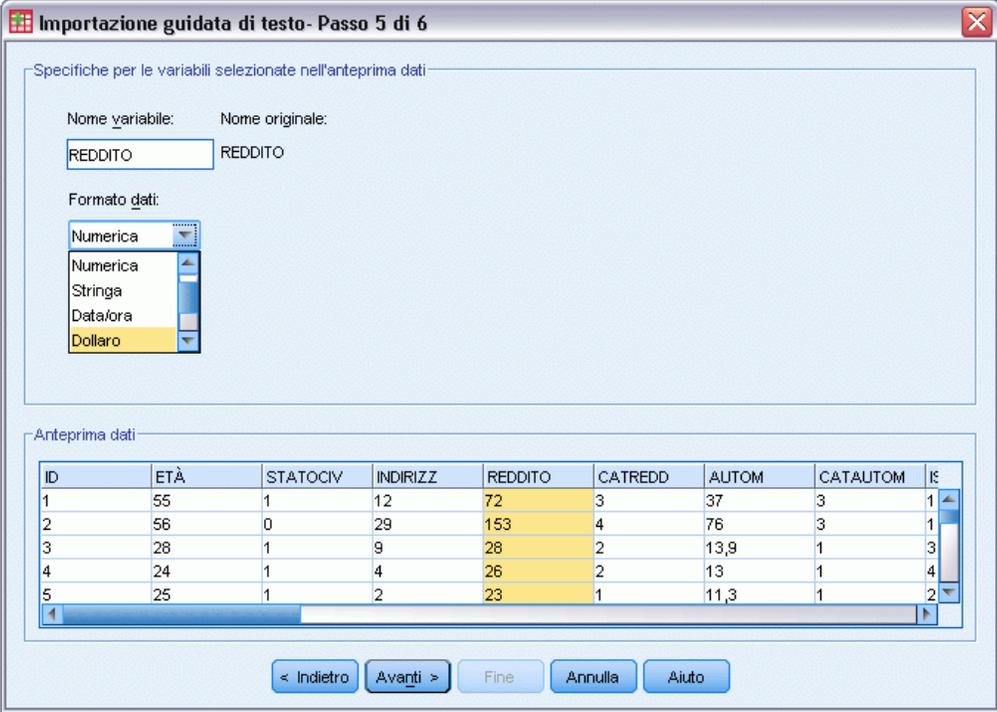
In questa finestra è inoltre possibile definire i tipi di dati, ad esempio in genere si presume che la variabile reddito contenga un importo in Euro specifico.

Per modificare un tipo di dati:

- In Anteprima dati selezionare la variabile che si desidera modificare, in questo caso *Reddito*.

- Selezionare Euro nell'elenco a discesa Formato dei dati.

Figura 2-17
Modificare il tipo di dati



Importazione guidata di testo - Passo 5 di 6

Specifiche per le variabili selezionate nell'anteprima dati

Nome variabile: Nome originale: REDDITO

Formato dati:

- Numerica
- Stringa
- Data/ora
- Dollaro

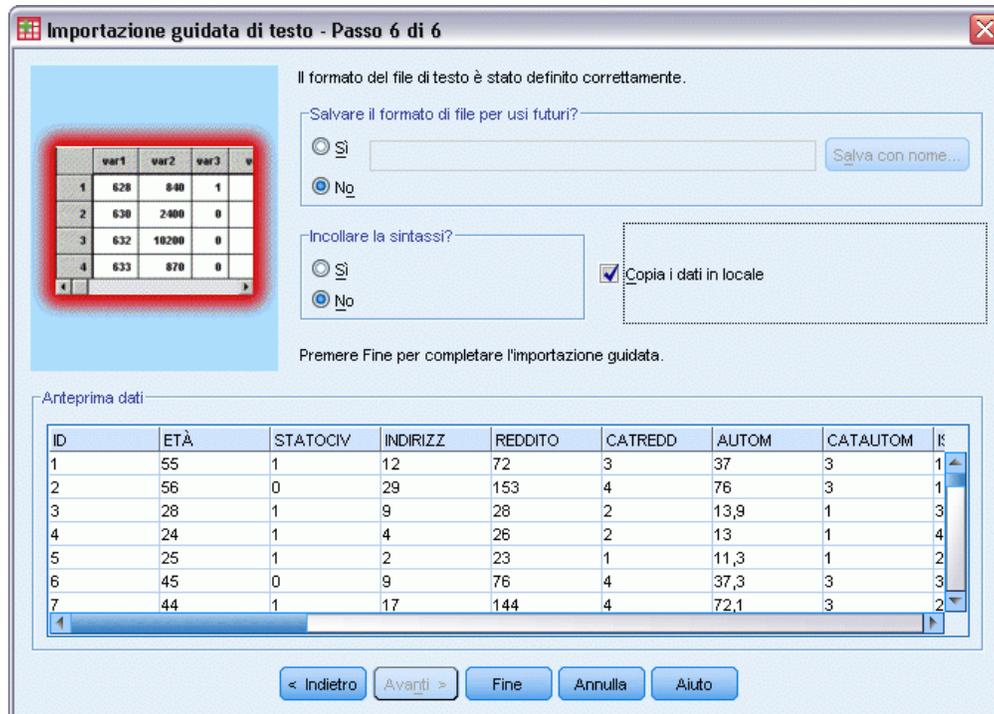
Anteprima dati

ID	ETÀ	STATOCIV	INDIRIZZ	REDDITO	CATREDD	AUTOM	CATAUTOM	IS
1	55	1	12	72	3	37	3	1
2	56	0	29	153	4	76	3	1
3	28	1	9	28	2	13,9	1	3
4	24	1	4	26	2	13	1	4
5	25	1	2	23	1	11,3	1	2

< Indietro Avanti > Fine Annulla Aiuto

- Fare clic su Avanti per continuare.

Figura 2-18
Importazione guidata di testo - Passo 6 di 6



- Lasciare inalterate le selezioni predefinite nella finestra di dialogo e fare clic su Fine per importare i dati.

Utilizzo dell'Editor dei dati

Nell'Editor dei dati viene visualizzato il contenuto del file di dati attivo e le informazioni sono rappresentate da variabili e casi.

- Nella Visualizzazione dati, le colonne rappresentano le variabili e le righe i casi (osservazioni).
- Nella Visualizzazione variabili, ogni riga rappresenta una variabile e ogni colonna rappresenta un attributo associato alla variabile specifica.

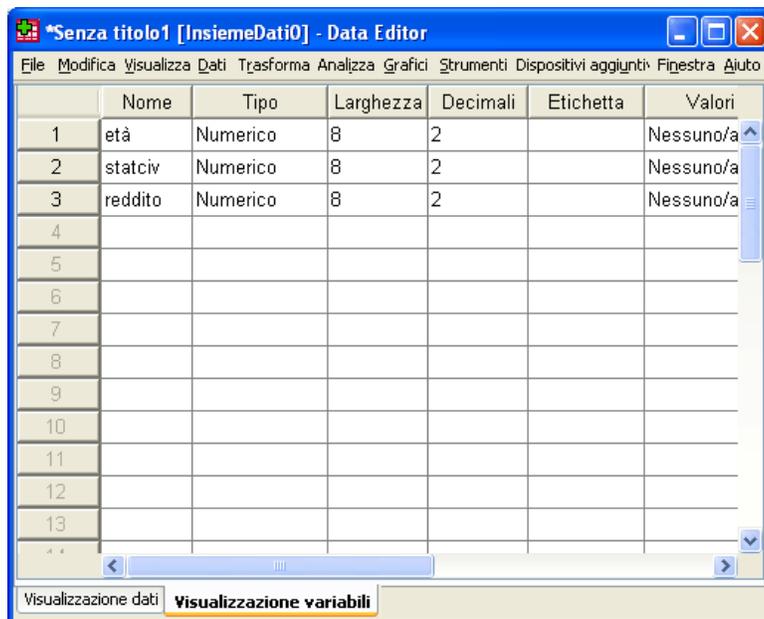
Le variabili consentono di rappresentare i diversi tipi di dati che sono stati compilati. Un'analogia frequente è quella del sondaggio, in cui la risposta a ognuna delle domande equivale a una variabile. Esistono molti tipi diversi di variabili, ad esempio le variabili numeriche, le variabili stringa, le variabili di valuta e di dati.

Inserimento di dati numerici

È possibile digitare dati nell'Editor dei dati e ciò può risultare utile per i file di dati di piccole dimensioni o per apportare modifiche di lieve entità ai file di dati più grandi.

- ▶ Fare clic sulla scheda Visualizzazione variabili nella parte inferiore della finestra dell'Editor dei dati.
È necessario definire le variabili che si desidera utilizzare. In questo caso sono richieste solo tre variabili: *età*, *statociv* e *reddito*.

Figura 3-1
Nomi di variabili in Visualizzazione variabili



- ▶ Nella prima riga della prima colonna digitare età.
- ▶ Nella seconda riga digitare statciv.
- ▶ Nella terza riga digitare reddito.

Alle nuove variabili viene automaticamente assegnato un tipo di dati numerico.

Se non si specificano i nomi delle variabili, vengono creati automaticamente nomi univoci, che non sono tuttavia descrittivi e non sono comunque consigliabili per i file di dati di grandi dimensioni.

- ▶ Fare clic sulla scheda Visualizzazione dati per continuare a inserire dati.

I nomi inseriti nella Visualizzazione variabili rappresentano le intestazioni delle prime tre colonne della Visualizzazione dati.

Iniziare a inserire i dati nella prima riga, a partire dalla prima colonna.

Figura 3-2

Valori inseriti in Visualizzazione dati

	age	marital	income	var	var
1	55,00	1,00	72000,0		
2	53,00	0,00	153000,0		
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					

- ▶ Nella colonna *età* digitare 55.
- ▶ Nella colonna *statociv* digitare 1.
- ▶ Nella colonna *reddito* digitare 72000.
- ▶ Per inserire i dati relativi al soggetto successivo, spostare il cursore sulla seconda riga della prima colonna.
- ▶ Nella colonna *età* digitare 53.
- ▶ Nella colonna *statociv* digitare 0.
- ▶ Nella colonna *reddito* digitare 153000.

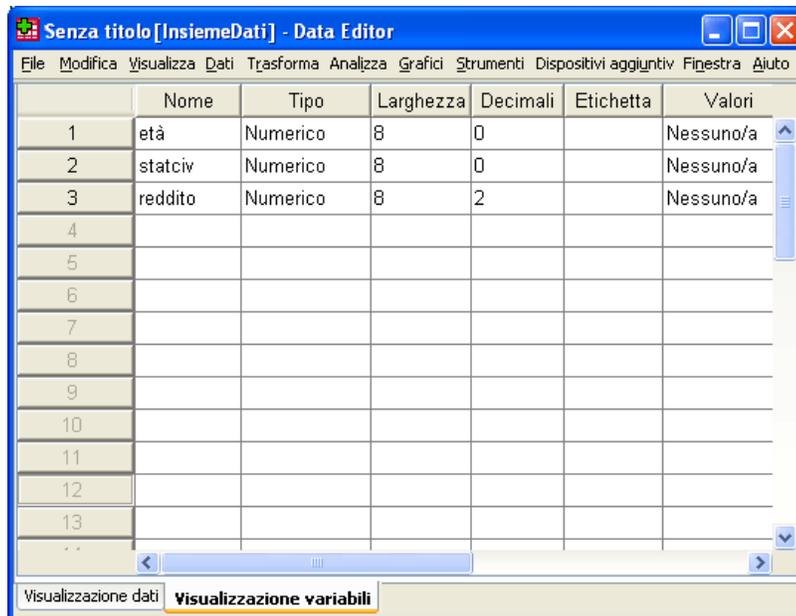
Nelle colonne *età* e *statociv* sono visualizzati i valori decimali, anche se i valori vengono considerati come interi. Per nascondere i decimali delle variabili:

- ▶ Fare clic sulla scheda Visualizzazione variabili nella parte inferiore della finestra dell'Editor dei dati.
- ▶ Nella colonna *Decimali* della riga *età* digitare 0 per nascondere il valore decimale.

- ▶ Nella colonna *Decimali* della riga *statociv* digitare 0 per nascondere il valore decimale.

Figura 3-3

Proprietà del valore decimale aggiornata per età e statociv



	Nome	Tipo	Larghezza	Decimali	Etichetta	Valori
1	età	Numerico	8	0		Nessuno/a
2	statociv	Numerico	8	0		Nessuno/a
3	reddito	Numerico	8	2		Nessuno/a
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
...						

Inserimento di dati stringa

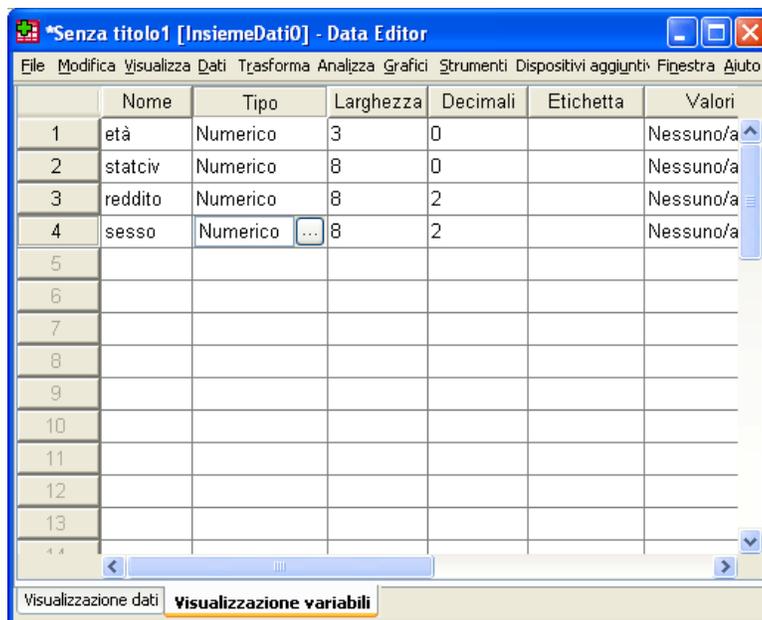
Nell'Editor dei dati è inoltre possibile inserire dati non numerici, ad esempio stringhe di testo.

- ▶ Fare clic sulla scheda Visualizzazione variabili nella parte inferiore della finestra dell'Editor dei dati.
- ▶ Nella prima cella della prima riga vuota digitare sesso come nome di variabile.
- ▶ Fare clic sulla cella *Tipo* accanto alla voce desiderata.

- Fare clic sul pulsante a destra della cella *Tipo* per aprire la finestra di dialogo Tipo di variabile.

Figura 3-4

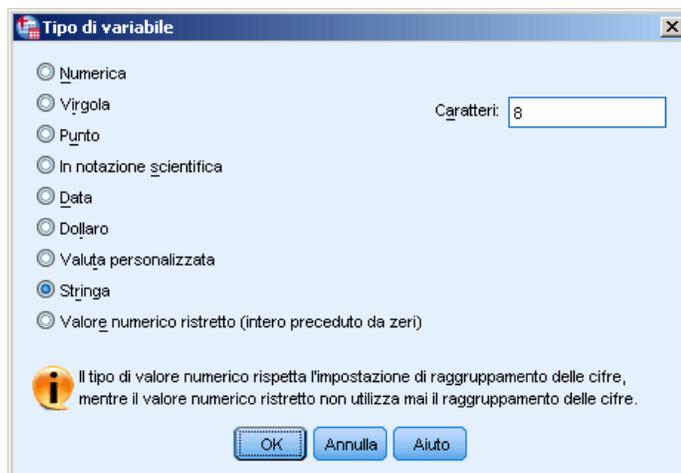
Pulsante visibile nella cella *Tipo* per sesso



- Selezionare Stringa per specificare il tipo di variabile.
- Fare clic su OK per salvare la selezione e tornare all'Editor dei dati.

Figura 3-5

Finestra di dialogo Tipo di variabile



Definizione di dati

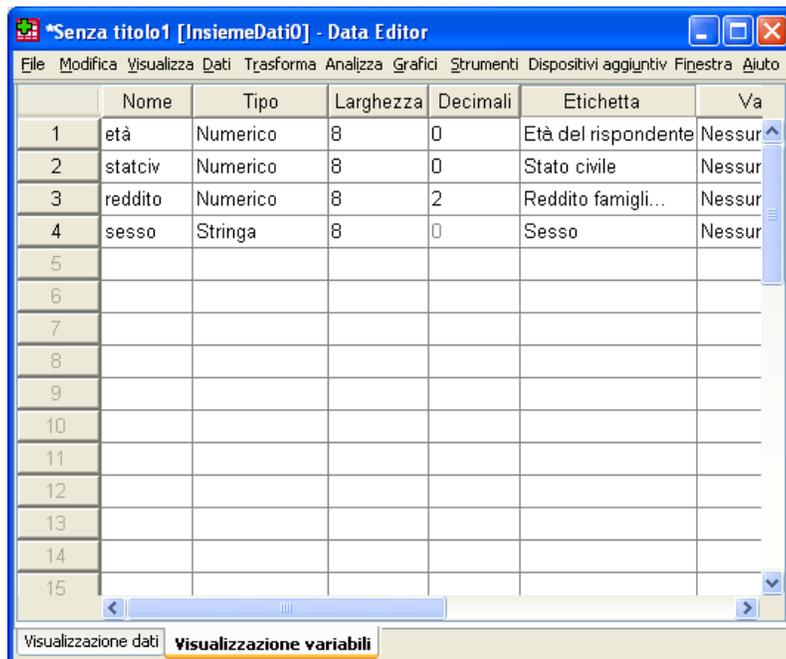
Oltre a definire i tipi di dati, è possibile definire anche le etichette descrittive dei valori e delle variabili per i nomi delle variabili e i valori dei dati, che vengono utilizzate nei report statistici e nei grafici.

Aggiunta delle etichette delle variabili

Le etichette forniscono una descrizione delle variabili, che spesso è rappresentata da una versione estesa del nome della variabile. La dimensione massima delle etichette è 255 byte. Tali etichette consentono di identificare le diverse variabili nell'output.

- ▶ Fare clic sulla scheda Visualizzazione variabili nella parte inferiore della finestra dell'Editor dei dati.
- ▶ Nella colonna *Etichetta* della riga *età* digitare Età del rispondente.
- ▶ Nella colonna *Etichetta* della riga *statociv* digitare Stato civile.
- ▶ Nella colonna *Etichetta* della riga *reddito* digitare Reddito familiare.
- ▶ Nella colonna *Etichetta* della riga *Sesso* digitare Sesso.

Figura 3-6
Etichette di variabili inserite in Visualizzazione variabili



	Nome	Tipo	Larghezza	Decimali	Etichetta	Va
1	età	Numerico	8	0	Età del rispondente	Nessur
2	statociv	Numerico	8	0	Stato civile	Nessur
3	reddito	Numerico	8	2	Reddito famigli...	Nessur
4	sesso	Stringa	8	0	Sesso	Nessur
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						

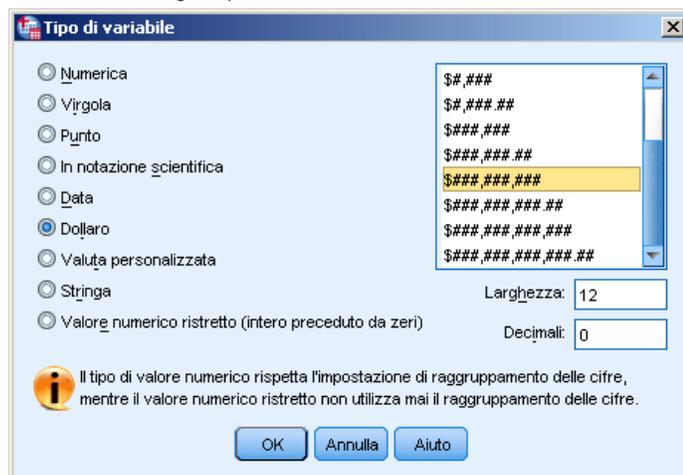
Visualizzazione dati **Visualizzazione variabili**

Modifica del formato e del tipo di variabile

Nella colonna *Tipo* viene visualizzato il tipo di dati corrente per ogni variabile. I tipi di dati più frequenti sono i dati numerici e le stringhe, ma sono supportati molti altri formati. Nel file dati corrente, la variabile *reddito* è definita come un tipo numerico.

- ▶ Fare clic sulla cella *Tipo* della riga *reddito*, quindi fare clic sul pulsante a destra della cella per aprire la finestra di dialogo Tipo di variabile.
- ▶ Selezionare Dollaro.

Figura 3-7
Finestra di dialogo Tipo di variabile



Vengono visualizzate le opzioni di formattazione per il tipo di dati correntemente selezionato.

- ▶ Per il formato della valuta relativo a questo esempio, selezionare `$###,###,###`.
- ▶ Fare clic su OK per salvare le modifiche.

Aggiunta di etichette dei valori per variabili numeriche

Le etichette dei valori consentono di associare i valori delle variabili all'etichetta di una stringa. In questo esempio, sono disponibili due valori validi per la variabile *statociv*. Il valore 0 significa che il soggetto è celibe o nubile, mentre il valore 1 significa che è coniugato o coniugata.

- ▶ Fare clic sulla cella *Valori* della riga *statociv*, quindi fare clic sul pulsante a destra della cella per aprire la finestra di dialogo Etichette dei valori.

Il **valore** corrisponde al valore numerico effettivo.

L'**etichetta di valore** rappresenta l'etichetta di stringa applicata al valore numerico specificato.

- ▶ Digitare 0 nel campo Valore.
- ▶ Digitare Celibe/Nubile nel campo Etichetta.

- Fare clic su Aggiungi per aggiungere l'etichetta all'elenco.

Figura 3-8

Finestra di dialogo Etichette dei valori.



- Digitare 1 nel campo Valore, quindi digitare Coniugato/a nel campo Etichetta.
- Fare clic su Aggiungi e quindi su OK per salvare le modifiche e tornare all'Editor dei dati.

Le etichette possono essere visualizzate anche nella Visualizzazione dati, in modo da rendere più leggibili i dati.

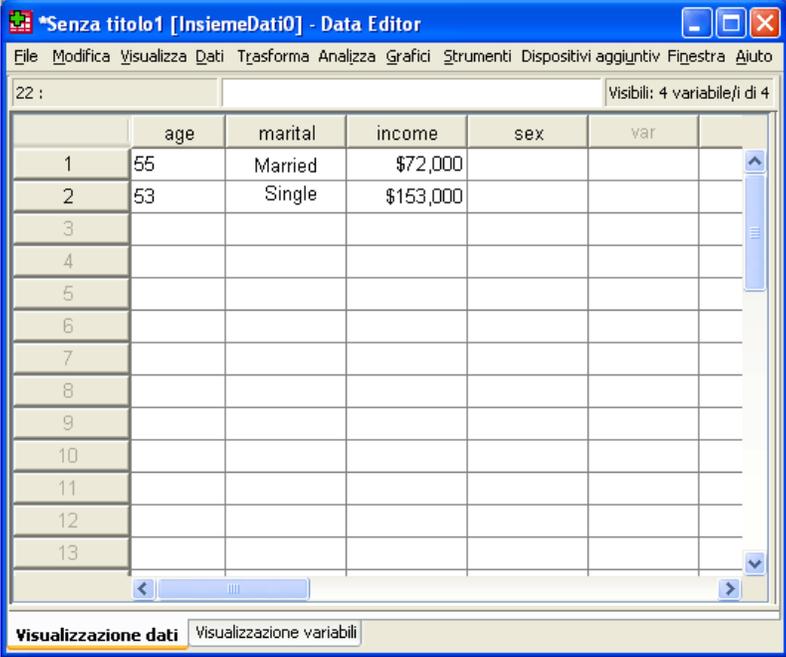
- Fare clic sulla scheda Visualizzazione dati nella parte inferiore della finestra dell'Editor dei dati.
- Dai menu, scegliere:
Visualizza > Etichette di valore

Adesso le etichette vengono visualizzate in un elenco quando si inseriscono valori nell'Editor dei dati. Con questa impostazione le opzioni valide hanno una forma più descrittiva.

Se la voce del menu Etichette dei valori è già attiva (la rispettiva casella di controllo è contrassegnata), selezionando di nuovo Etichette dei valori si *disattiva* la visualizzazione delle etichette dei valori.

Figura 3-9

Etichette dei valori visualizzate nella Visualizzazione dei dati



	age	marital	income	sex	var
1	55	Married	\$72,000		
2	53	Single	\$153,000		
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					

Aggiunta di etichette dei valori per variabili stringa

Le etichette di valore possono essere richieste anche per le variabili stringa. Per l'identificazione del sesso del soggetto, ad esempio, potrebbero venire utilizzate singole lettere, quali *M* o *F*. Per specificare che *M* sta per *Maschio* e *F* sta per *Femmina*, è possibile utilizzare etichette dei valori.

- ▶ Fare clic sulla scheda Visualizzazione variabili nella parte inferiore della finestra dell'Editor dei dati.
- ▶ Fare clic sulla cella *Valori* della riga *sesso*, quindi fare clic sul pulsante a destra della cella per aprire la finestra di dialogo Etichette dei valori.
- ▶ Digitare F nel campo Valore, quindi digitare Femmina nel campo Etichetta.

- Fare clic su Aggiungi per aggiungere l'etichetta al file di dati.

Figura 3-10

Finestra di dialogo Etichette dei valori.



- Digitare M nel campo Valore, quindi digitare Maschio nel campo Etichetta.
- Fare clic su Aggiungi e quindi su OK per salvare le modifiche e tornare all'Editor dei dati.

Poiché per i valori stringa è attiva la distinzione tra lettere maiuscole e minuscole, è necessario mantenere una certa coerenza. La *m* minuscola e la *M* maiuscola sono considerate lettere diverse.

Utilizzo di etichette dei valori definite per l'inserimento dei dati

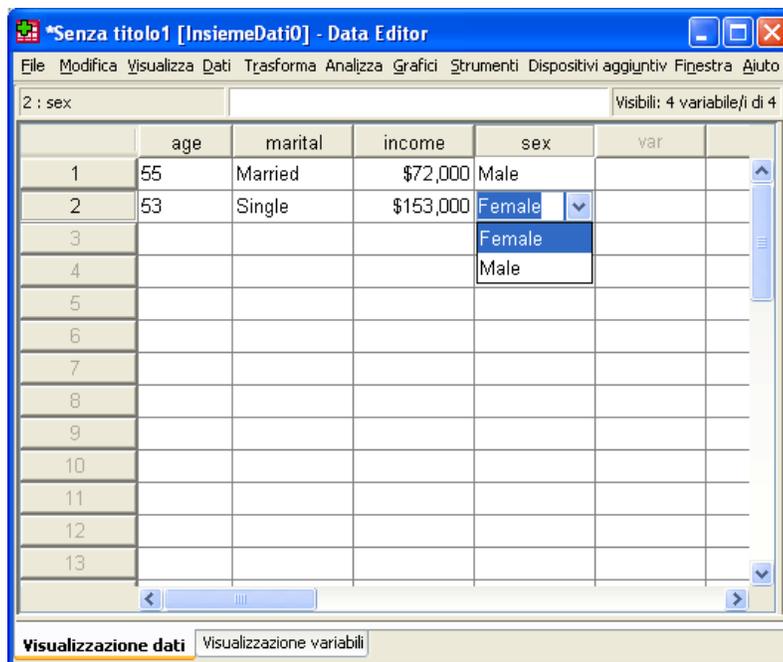
Per l'inserimento dei dati è possibile utilizzare le etichette dei valori.

- Fare clic sulla scheda Visualizzazione dati nella parte inferiore della finestra dell'Editor dei dati.
- Nella prima riga, selezionare la cella relativa a *sesso*.
- Fare clic sul pulsante a destra della cella, quindi selezionare Maschio dall'elenco a discesa.
- Nella seconda riga, selezionare la cella relativa a *sesso*.

- Fare clic sul pulsante a destra della cella, quindi selezionare Femmina dall'elenco a discesa.

Figura 3-11

Utilizzo delle etichette di variabile per selezionare i valori



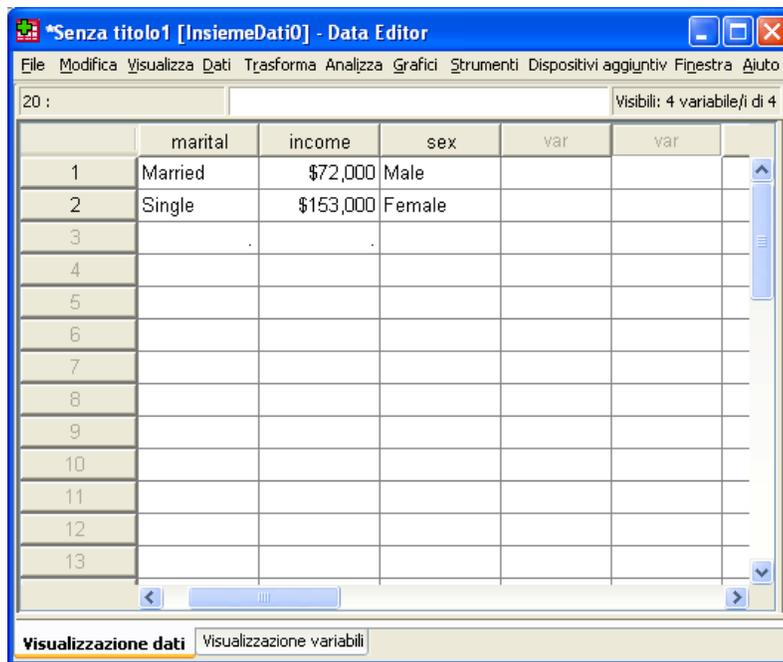
Vengono visualizzati solo i valori definiti e pertanto si può essere certi che i dati vengono inseriti in uno dei formati previsti.

Gestione dei dati mancanti

I dati mancanti o non validi sono in genere molto frequenti e pertanto non è possibile ignorarli. È possibile che i rispondenti si rifiutino di rispondere ad alcune domande del sondaggio, non sappiano cosa rispondere oppure rispondano in una forma non prevista. Se non si filtrano o si identificano questi dati, i risultati dell'analisi potrebbero rivelarsi inesatti.

Per quanto riguarda i dati numerici, i campi vuoti o i campi che contengono voci non valide vengono convertiti in dati mancanti di sistema, identificati da un punto.

Figura 3-12
Valori mancanti visualizzati come punti



The screenshot shows the SPSS Data Editor window titled '*Senza titolo1 [InsiemeDati0] - Data Editor'. The menu bar includes File, Modifica, Visualizza, Dati, Trasforma, Analizza, Grafici, Strumenti, Dispositivi aggiuntivi, Finestra, and Aiuto. The status bar at the top indicates '20 :' and 'Visibili: 4 variabile/i di 4'. The main table has columns for 'marital', 'income', 'sex', and two empty 'var' columns. The data rows are numbered 1 through 13. Row 1 shows 'Married', '\$72,000', and 'Male'. Row 2 shows 'Single', '\$153,000', and 'Female'. Rows 3 through 13 show missing values for all variables, represented by a single dot in the first column.

	marital	income	sex	var	var
1	Married	\$72,000	Male		
2	Single	\$153,000	Female		
3	.	.	.		
4	.	.	.		
5	.	.	.		
6	.	.	.		
7	.	.	.		
8	.	.	.		
9	.	.	.		
10	.	.	.		
11	.	.	.		
12	.	.	.		
13	.	.	.		

At the bottom, there are two tabs: 'Visualizzazione dati' (selected) and 'Visualizzazione variabili'.

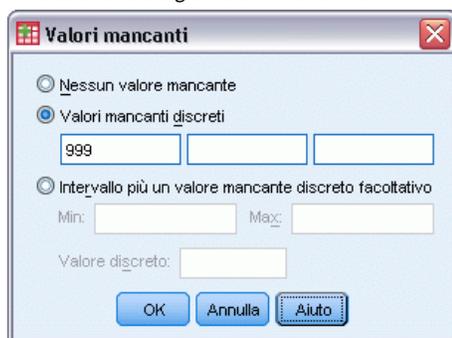
Il motivo per cui un valore risulta mancante può rivelarsi importante ai fini dell'analisi. È possibile, ad esempio, che sia importante distinguere i rispondenti che hanno rifiutato di rispondere a una domanda da quelli che invece non hanno fornito una risposta perché hanno ritenuto che la domanda non li riguardasse.

Valori mancanti per una variabile numerica

- ▶ Fare clic sulla scheda Visualizzazione variabili nella parte inferiore della finestra dell'Editor dei dati.
- ▶ Fare clic sulla cella *Mancanti* della riga *età*, quindi fare clic sul pulsante a destra della cella per aprire la finestra di dialogo Valori mancanti.

In questa finestra di dialogo è possibile specificare fino a tre valori mancanti distinti oppure un intervallo di valori e un valore discreto aggiuntivo.

Figura 3-13
Finestra di dialogo Valori mancanti

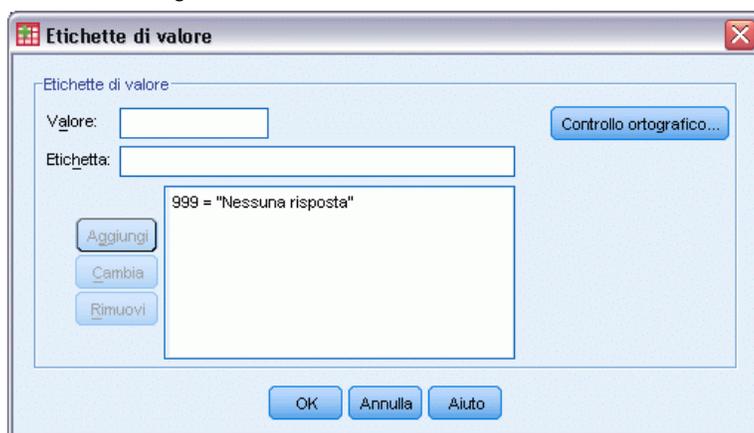


- ▶ Selezionare Valori mancanti discreti.
- ▶ Digitare 999 nella prima casella di testo e lasciare vuote le altre due.
- ▶ Fare clic su OK per confermare le modifiche e tornare all'Editor dei dati.

Dopo aver aggiunto il valore mancante, è possibile applicarvi un'etichetta.

- ▶ Fare clic sulla cella *Valori* della riga *età*, quindi fare clic sul pulsante a destra della cella per aprire la finestra di dialogo Etichette dei valori.
- ▶ Digitare 999 nel campo Valore.
- ▶ Digitare Nessuna risposta nel campo Etichetta.

Figura 3-14
Finestra di dialogo Etichette dei valori.



- ▶ Fare clic su Aggiungi per aggiungere l'etichetta al file di dati.
- ▶ Fare clic su OK per confermare le modifiche e tornare all'Editor dei dati.

Valori mancanti per una variabile stringa

I valori mancanti per le variabili stringa vengono gestiti allo stesso modo di quelli per le variabili numeriche. Tuttavia, a differenza delle variabili numeriche, i campi vuoti nelle variabili stringa non vengono identificati come valori mancanti di sistema, ma vengono interpretati come stringhe vuote.

- ▶ Fare clic sulla scheda Visualizzazione variabili nella parte inferiore della finestra dell'Editor dei dati.
- ▶ Fare clic sulla cella *Mancanti* della riga *sex*, quindi fare clic sul pulsante a destra della cella per aprire la finestra di dialogo Valori mancanti.
- ▶ Selezionare Valori mancanti discreti.
- ▶ Digitare NR nella prima casella di testo.

Per i valori mancanti delle variabili stringa è attiva la distinzione tra maiuscole e minuscole e pertanto il valore *nr* non viene identificato come un valore mancante.

- ▶ Fare clic su OK per confermare le modifiche e tornare all'Editor dei dati.

È ora possibile aggiungere un'etichetta per il valore mancante.

- ▶ Fare clic sulla cella *Valori* della riga *sex*, quindi fare clic sul pulsante a destra della cella per aprire la finestra di dialogo Etichette dei valori.
- ▶ Digitare NR nel campo Valore.
- ▶ Digitare Nessuna risposta nel campo Etichetta.

Figura 3-15
Finestra di dialogo Etichette dei valori.



- ▶ Fare clic su Aggiungi per aggiungere l'etichetta al progetto.
- ▶ Fare clic su OK per confermare le modifiche e tornare all'Editor dei dati.

Copiare e incollare gli attributi delle variabili

Dopo avere specificato gli attributi di una variabile, è possibile copiarli e applicarli ad altre variabili.

- Nella Visualizzazione variabili digitare *età*matr nella prima cella della prima riga vuota.

Figura 3-16

Variabile *età*matr nella Visualizzazione variabili

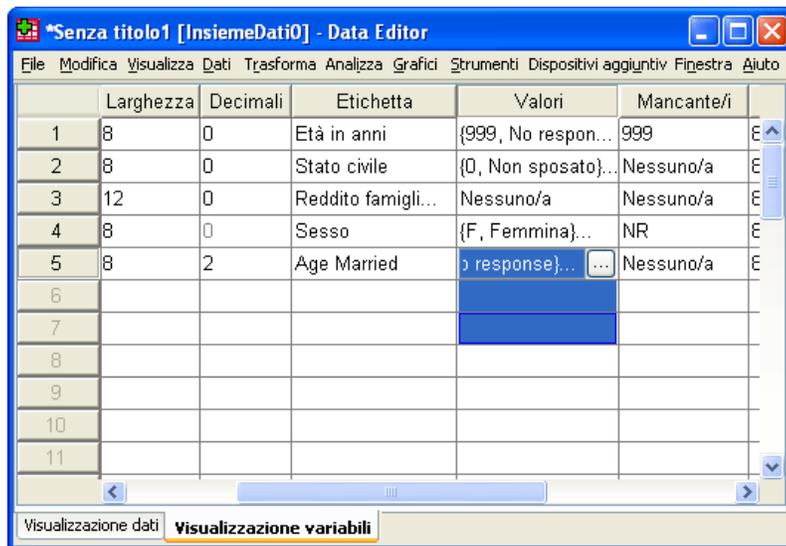
	Nome	Tipo	Lunghezza	Decimali	Etichetta
1	età	Numerica	8	0	Età dell'intervistato
2	statociv	Numerica	8	0	Stato civile
3	reddito	Dollaro	12	2	Reddito familiare
4	sesso	Stringa	8	0	Sesso
5	etàmatr	Numerica	8	2	Età al matrimonio
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					

- Nella colonna *Etichetta* digitare Età al matrimonio.
- Fare clic sulla cella *Valori* della riga *età*.
- Dai menu, scegliere:
Modifica > Copia
- Fare clic sulla cella *Valori* della riga *età*matr.
- Dai menu, scegliere:
Modifica > Incolla

I valori definiti dalla variabile *età* vengono applicati alla variabile *età*matr.

Per applicare l'attributo a più variabili, è sufficiente selezionare più celle di destinazione facendo clic e trascinando la colonna verso il basso.

Figura 3-17
Selezione di più celle



The screenshot shows the SPSS Data Editor window titled '*Senza titolo1 [InsiemeDati0] - Data Editor'. The window has a menu bar with 'File', 'Modifica', 'Visualizza', 'Dati', 'Trasforma', 'Analizza', 'Grafici', 'Strumenti', 'Dispositivi aggiuntivi', 'Finestra', and 'Aiuto'. The main area contains a table with the following columns: 'Larghezza', 'Decimali', 'Etichetta', 'Valori', and 'Mancante/i'. The rows are numbered 1 through 11. Rows 1-5 are visible, and rows 6-11 are selected, indicated by a blue background. The 'Valori' column for row 5 contains a dropdown menu with 'No response' selected.

	Larghezza	Decimali	Etichetta	Valori	Mancante/i
1	8	0	Età in anni	{999, No respon...	999
2	8	0	Stato civile	{0, Non sposato}...	Nessuno/a
3	12	0	Reddito famigli...	Nessuno/a	Nessuno/a
4	8	0	Sesso	{F, Femmina}...	NR
5	8	2	Age Married	No response... [dropdown]	Nessuno/a
6					
7					
8					
9					
10					
11					

Quando si incolla l'attributo, questo viene applicato a tutte le celle selezionate.

Se si incollano i valori in righe vuote, vengono automaticamente create nuove variabili.

Per copiare tutti gli attributi da una variabile all'altra:

- Fare clic sul numero di riga nella riga *stataciv*.

Figura 3-18
Riga selezionata

	Nome	Tipo	Larghezza	Decimali	Etichetta	
1	età	Numerico	8	0	Età in anni	{999
2	stataciv	Numerico	8	0	Stato civile	{0, S
3	reddito	Dollaro	12	0	Reddito famigli...	Nes
4	sezzo	Stringa	8	0	Sesso	{F, F
5	agewed	Numerico	8	2	Age Married	{999
6	VAR00001	Numerico	8	2		{999
7	VAR00002	Numerico	8	2		{999
8	VAR00003	Numerico	8	2		{999
9						
10						
11						
12						
13						
..						

Visualizzazione dati **Visualizzazione variabili**

- Dai menu, scegliere:
Modifica > Copia
- Fare clic sul numero della prima riga vuota.
- Dai menu, scegliere:
Modifica > Incolla

Tutti gli attributi della variabile *stato* vengono applicati alla nuova variabile.

Figura 3-19

Tutti i valori incollati in una riga

	Nome	Tipo	Larghezza	Decimali	Etichetta	
1	età	Numerico	8	0	Età in anni	{999
2	stato	Numerico	8	0	Stato civile	{0, S
3	reddito	Dollaro	12	0	Reddito famigli...	Nes
4	sesto	Stringa	8	0	Sesso	{F, F
5	agewed	Numerico	8	2	Age Married	{999
6	VAR00001	Numerico	8	2		{999
7	VAR00002	Numerico	8	2		{999
8	VAR00003	Numerico	8	2		{999
9	VAR00004	Numerico	8	0	Stato civile	{0, S
10						
11						
12						
13						

Definizione delle proprietà delle variabili categoriali

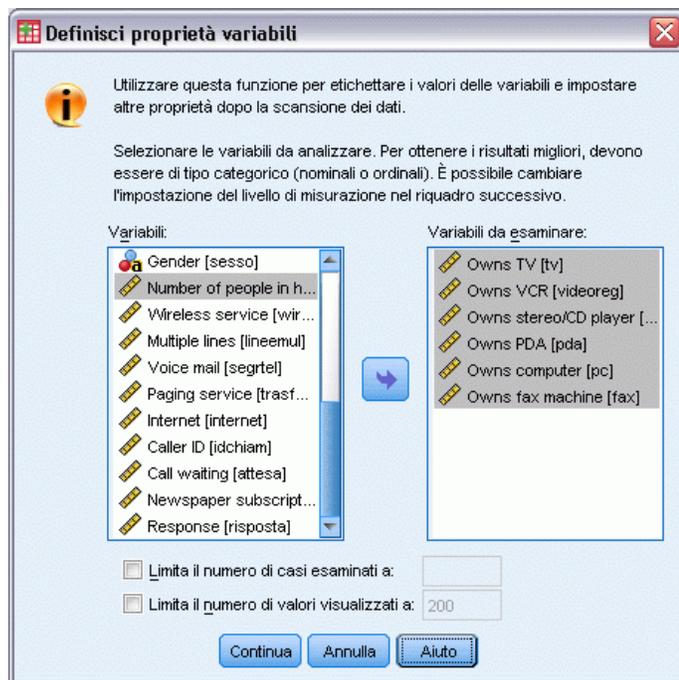
Per i dati (ordinali, nominali) categoriali, la finestra Definisci proprietà variabili consente di definire le etichette dei valori e altre proprietà delle variabili. Il processo Definisci proprietà variabili:

- Esamina i valori effettivi dei dati ed elenca tutti i valori univoci dei dati per ogni variabile selezionata.
- Identifica i valori senza etichetta e fornisce una funzione per l'aggiunta automatica dell'etichetta.
- Consente di copiare etichette di valori definite da un'altra variabile nella variabile selezionata o dalla variabile selezionata nelle variabili aggiuntive.

In questo esempio viene utilizzato il file di dati *demo.sav*. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice A a pag. 152.](#) Questo file di dati contiene già etichette dei valori definite, pertanto si deve immettere un valore per il quale non siano state definite etichette di valore.

- ▶ Nella Visualizzazione dati dell'Editor dei dati, fare clic sulla prima cella di dati relativa alla variabile *computer* (è possibile che sia necessario scorrere verso destra) e immettere 99.
- ▶ Dai menu, scegliere:
Dati > Definisci proprietà variabili...

Figura 3-20
Finestra di dialogo iniziale *Definisci proprietà variabili*



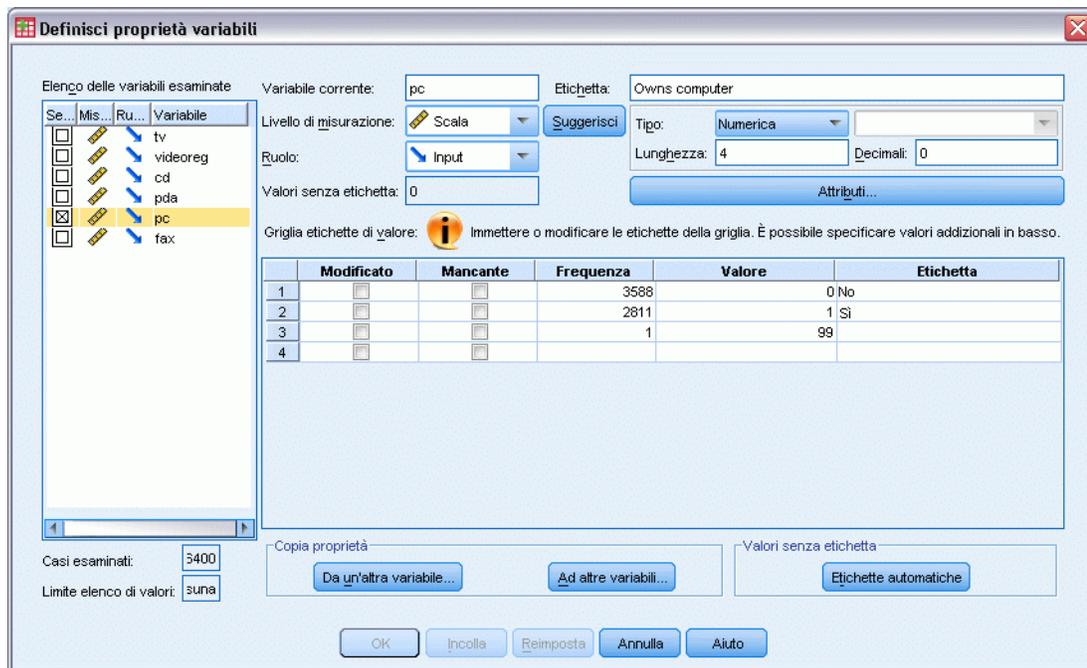
Nella finestra di dialogo iniziale della procedura *Definisci proprietà variabili*, selezionare le variabili ordinali o nominali di cui si desiderano definire le etichette di valore e/o altre proprietà.

- Trascinare e rilasciare da *Possiede TV [owntv]* a *Possiede fax [ownfax]* nell'elenco *Variabili da esaminare*.

Notare che le icone del livello di misurazione relative a tutte le variabili selezionate indicano che si tratta di variabili di scala, non di variabili categoriali. Tutte le variabili selezionate in questo esempio sono in effetti variabili categoriali che utilizzano i valori numerici 0 e 1 per indicare rispettivamente *No* e *Sì* e una delle proprietà delle variabili che verrà modificata in *Definisci proprietà variabili* è il livello di misurazione.

- Fare clic su *Continua*.

Figura 3-21
Definisci proprietà variabili, finestra di dialogo principale



- In Elenco variabili esaminate, selezionare *computer*.

Il livello di misurazione corrente relativo alla variabile selezionata è scala. È possibile modificare il livello di misurazione selezionando un livello dall'elenco a discesa oppure è possibile accettare il livello di misurazione suggerito in Definisci proprietà variabili.

- Fare clic su Suggerisci.

Si visualizza la finestra di dialogo Suggerisci il livello di misurazione.

Figura 3-22
Finestra di dialogo *Suggerisci il livello di misurazione*

Definisci proprietà variabili - Suggerire il livello di misurazione

Variabile: pc

Livello di misurazione corrente: Scala

Livello consigliato: Ordinale o Nominale

Ordinale Nominale

Spiegazione: pc ha solo valori bassi, non negativi e di numeri interi e non ha un numero alto di valori diversi.

i In un livello nominale i valori delle variabili non hanno una classifica naturale, ad esempio i nomi delle nazioni.

In un livello ordinale i valori delle variabili hanno un ordine naturale ma le differenze tra i valori non sono significative, ad esempio l'importanza di una carica politica codificata come "bassa", "media" e "alta".

In un livello di scala le differenze tra i valori delle variabili possono essere confrontate, ad esempio le età in anni.

Il livello di misurazione suggerito si basa su regole empiriche e non va a sostituire il giudizio dell'utente.

SPSS Statistics utilizza in alcuni casi il livello di misurazione per determinare se

Continua Annulla Aiuto

Poiché la variabile non contiene molti valori diversi e tutti i casi esaminati contengono valori interi, è probabile che il livello di misurazione corretto sia ordinale o nominale.

- Selezionare Ordinale, quindi fare clic su Continua.

Ora il livello di misurazione relativo alla variabile selezionata è ordinale.

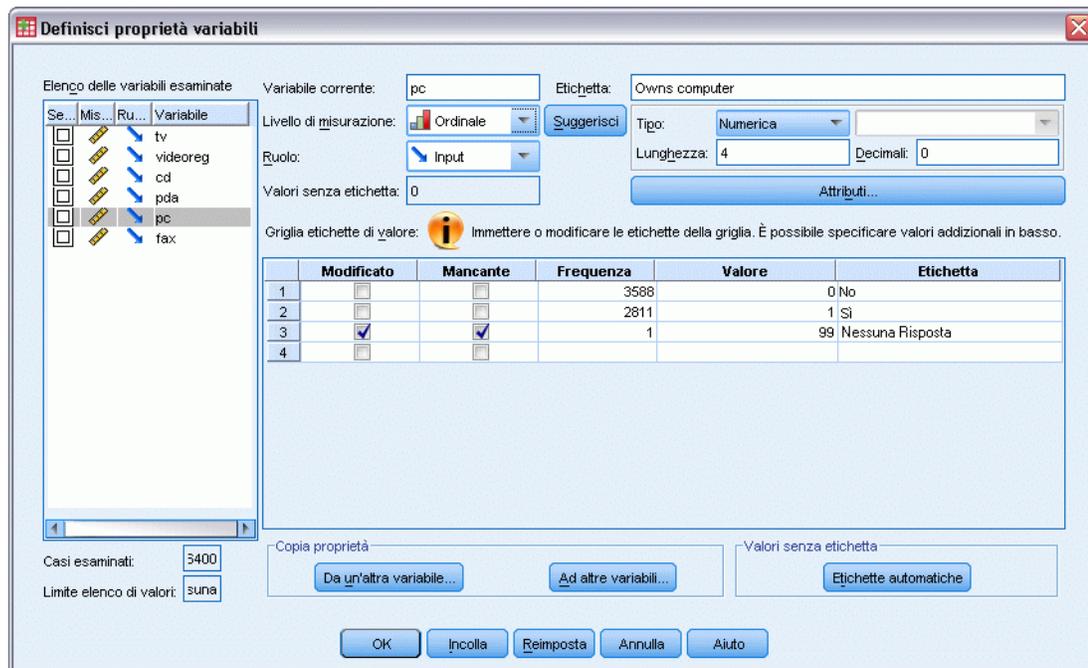
La griglia Etichette dei valori visualizza tutti i valori dei dati univoci relativi alla variabile selezionata, qualsiasi etichetta di valore definita per tali valori e il numero di volte (conteggio) in cui appare ogni valore nei casi esaminati.

Il valore specificato in Visualizzazione dati, 99, viene visualizzato nella griglia. Il conteggio è 1 perché è stato modificato il valore di un solo caso e la colonna *Etichetta* è vuota perché non è ancora stata definita un'etichetta di valore per 99. Il segno di spunta X nella prima colonna dell'Elenco variabili esaminate indica che la variabile selezionata contiene almeno un valore osservato senza un'etichetta di valore definita.

- Nella colonna *Etichetta* relativa al valore 99, specificare Nessuna risposta.
- Selezionare la casella nella colonna *Mancante* per il valore 99 per identificarlo come valore **mancante definito dall'utente**.

I dati specificati come valori mancanti definiti dall'utente vengono evidenziati in modo da poter essere elaborati diversamente ed esclusi dalla maggior parte dei calcoli.

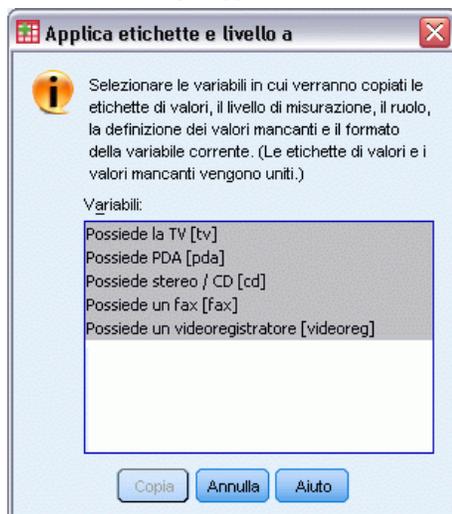
Figura 3-23
Nuove proprietà definite per la variabile *computer*



Prima di terminare l'operazione di modifica delle proprietà della variabile *computer*, applicare lo stesso livello di misurazione, le stesse etichette di valore e definizioni dei valori mancanti alle altre variabili nell'elenco.

- Nell'area Copia proprietà, fare clic su Ad altre variabili.

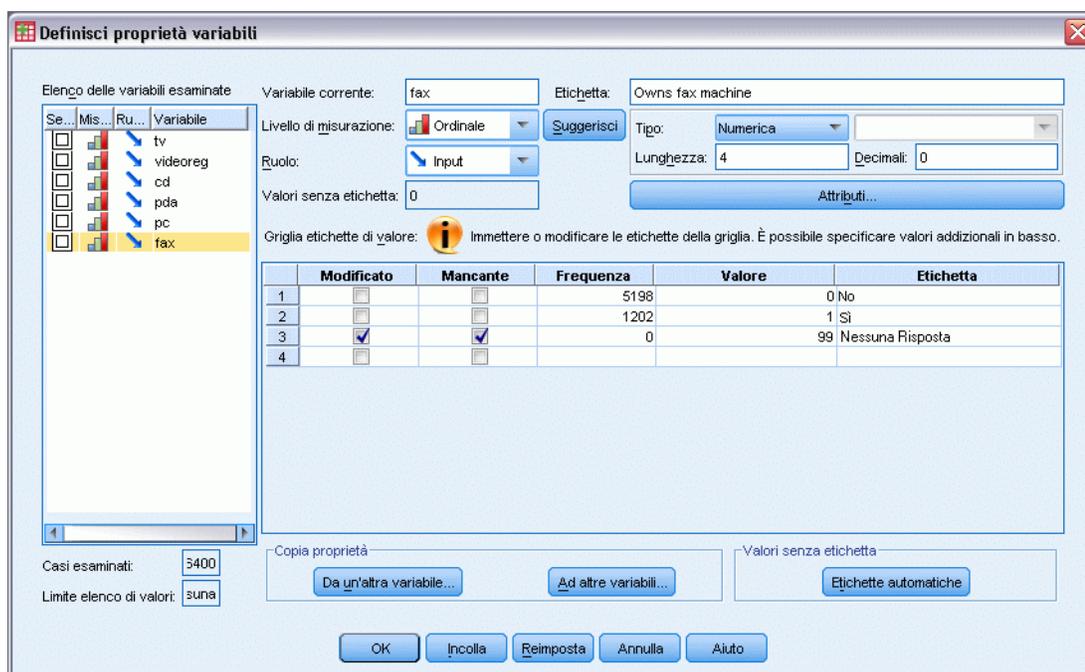
Figura 3-24
Finestra di dialogo *Applica etichette e livello*



- Nella finestra di dialogo *Applica etichette e livello*, selezionare tutte le variabili nell'elenco, quindi fare clic su *Copia*.

Se a questo punto si seleziona qualsiasi altra variabile nell'Elenco variabili esaminate della finestra di dialogo *Definisci proprietà variabili*, tutte le variabili visualizzate saranno ordinali, 99 verrà specificato come valore mancante definito dall'utente e *Nessuna risposta* come etichetta di valore.

Figura 3-25
Nuove proprietà definite per la variabile *fax*



- ▶ Fare clic su OK per salvare tutte le proprietà definite per le variabili.

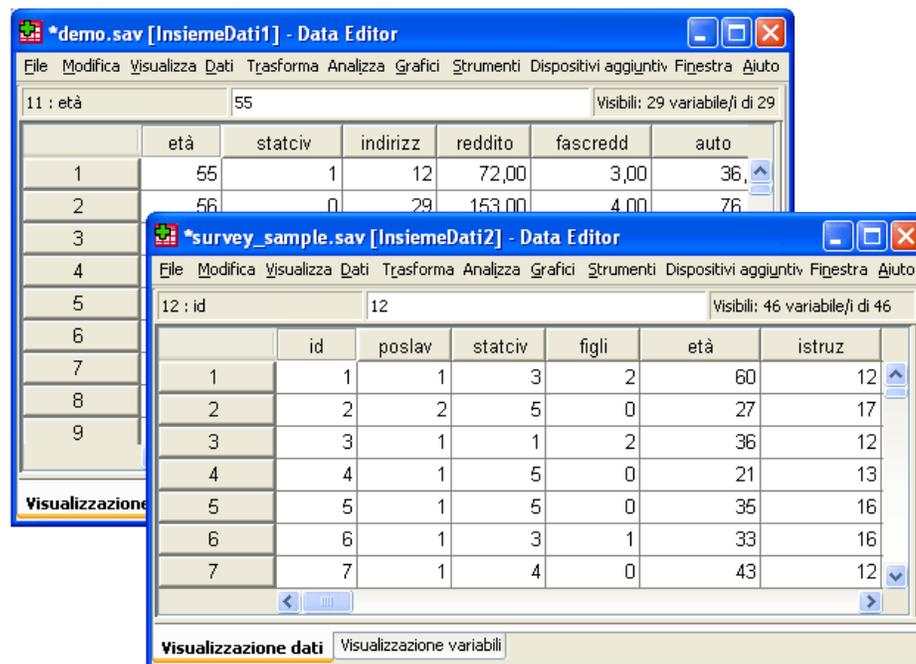
Utilizzo di più sorgenti dati

A partire dalla versione 14.0, è consentito tenere aperte più sorgenti dati contemporaneamente. Ciò permette agevolmente di:

- Passare da una sorgente dati a un'altra.
- Confrontare i contenuti di più sorgenti dati.
- Copiare e incollare i dati tra le sorgenti dati.
- Creare più sottoinsiemi di casi e/o di variabili per l'analisi.
- Unire più sorgenti dati di formati dati diversi (ad esempio fogli di calcolo, database e dati di testo) senza dover salvare prima ciascuna sorgente dati.

Gestione di più sorgenti di dati

Figura 4-1
Due sorgenti di dati aperte contemporaneamente



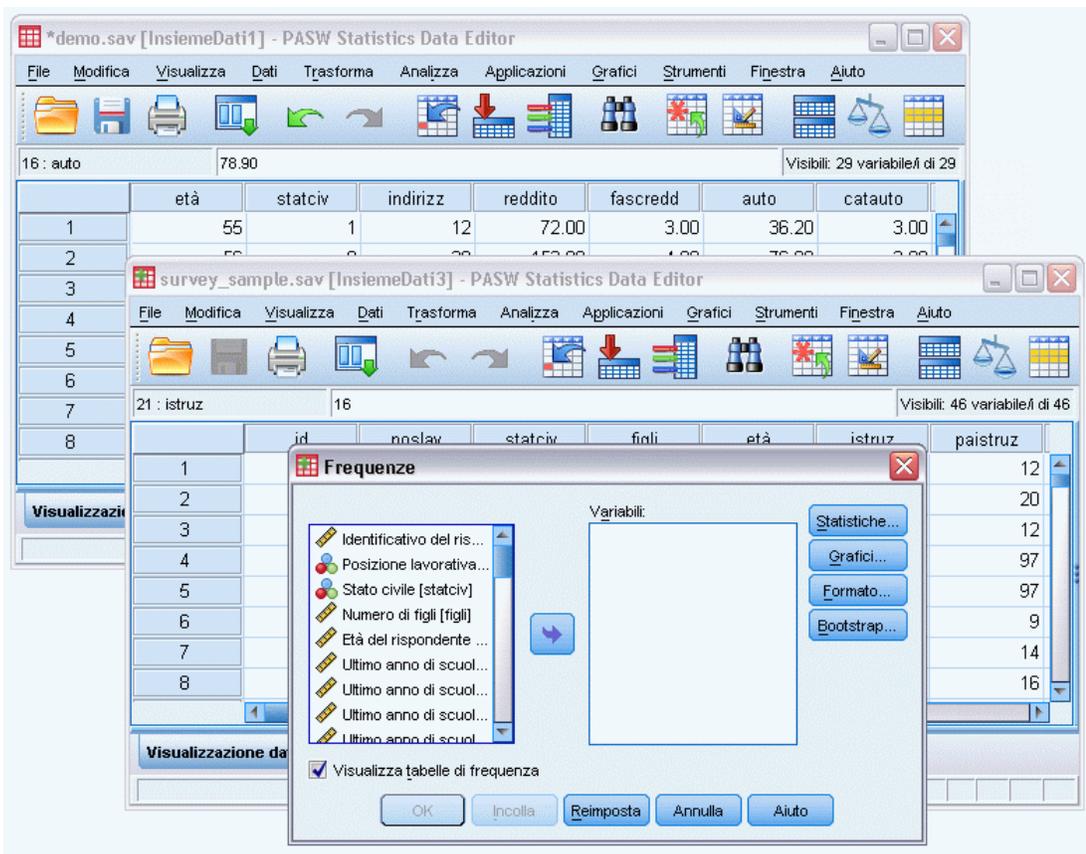
Per impostazione predefinita, le sorgenti dati aperte vengono visualizzate in una nuova finestra dell'Editor dei dati.

- Tutte le sorgenti di dati già aperte rimangono tali e possono essere usate.
- La sorgente di dati aperte diventa automaticamente il **insieme di dati attivo**.

- Per modificare l'insieme di dati attivo, è sufficiente fare clic in qualsiasi punto della finestra dell'Editor dei dati della sorgente di dati da usare oppure selezionare la finestra Editor dei dati per la sorgente di dati dal menu Finestra.
- Solo le variabili dell'insieme di dati attivo possono essere usate per l'analisi.

Figura 4-2

Elenco di variabili contenente le variabili dell'insieme di dati attivo



- L'insieme di dati attivo non può essere modificato se è aperta una finestra di dialogo che ha accesso ai dati (ciò include anche le finestre di dialogo che visualizzano gli elenchi di variabili).
- È sempre necessario aprire almeno una finestra dell'Editor dei dati per sessione. Quando si chiude l'ultima finestra aperta dell'Editor dei dati, IBM® SPSS® Statistics la chiude automaticamente chiedendo all'utente se desidera salvare le modifiche.

Utilizzo di più insiemi di dati nella sintassi dei comandi

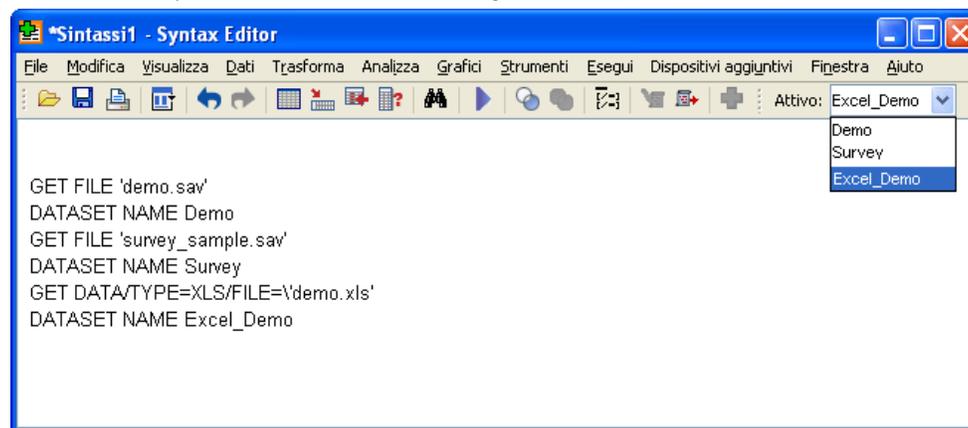
Se si utilizza la sintassi dei comandi per aprire le sorgenti dati (ad esempio, GET FILE, GET DATA), è necessario utilizzare il comando DATASET NAME per denominare esplicitamente ciascun insieme di dati per poter tenere contemporaneamente aperte più sorgenti dati.

Quando si utilizza la sintassi dei comandi, il nome dell'insieme di dati attivo viene visualizzato sulla barra degli strumenti della finestra della sintassi. Tutte le azioni riportate di seguito consentono di modificare l'insieme di dati attivo:

- Utilizzare il comando `DATASET ACTIVATE`.
- Fare clic in un punto qualsiasi della finestra dell'Editor dei dati di un insieme di dati.
- Selezionare un nome di insieme di dati dalla barra degli strumenti nella finestra della sintassi.

Figura 4-3

Insiemi di dati aperti visualizzati sulla barra degli strumenti della finestra della sintassi



Copia e inserimento di informazioni in più insiemi di dati

È possibile copiare gli attributi delle definizioni dei dati e delle variabili da un insieme di dati a un altro eseguendo le stesse operazioni utilizzabili all'interno dei singoli file di dati.

- Se si copiano e si incollano le celle dati selezionati in Visualizzazione dei dati, vengono incollati solo i valori dei dati senza gli attributi delle definizioni della variabili.
- Se si copia e si incolla una variabile intera nella Visualizzazione dei dati selezionando il nome della variabile nella parte superiore della colonna, vengono incollati tutti i dati e tutti gli attributi delle definizioni delle variabili per la variabile in questione.
- Se si copiano e si incollano gli attributi delle definizioni delle variabili o le variabili intere nella Visualizzazione variabili, vengono incollati gli attributi selezionati (o la definizione completa della variabile) ma non i valori dati.

Ridenominazione degli insiemi di dati

Quando si apre una sorgente dati tramite i menu e le finestre di dialogo, a ciascuna sorgente viene automaticamente assegnato il nome di file dati *DataSetn*, dove *n* è il valore intero sequenziale.

Quando si apre una sorgente dati tramite la sintassi del comando, non viene assegnato alcun

nome di insieme di dati a meno che non se ne specifichi uno con il comando `DATASET NAME`. Per assegnare un nome di insieme di dati più descrittivo:

- ▶ Dai menu della finestra Editor dei dati relativo all'insieme di dati di cui si vuole modificare il nome selezionare:
File > Rinomina insieme dati...
- ▶ Immettere il nome di un nuovo insieme di dati conforme alle regole di denominazione delle variabili.

Soppressione di più insiemi di dati

Se si preferisce avere a disposizione un solo insieme di dati alla volta e si desidera sopprimere la funzione relativa a più insiemi di dati:

- ▶ Dai menu, scegliere:
Modifica > Opzioni...
- ▶ Fare clic sulla scheda Generale.

Selezionare Apri un solo insieme di dati alla volta.

Esame delle statistiche riassuntive per singole variabili

In questo capitolo vengono trattate alcune semplici misure riassuntive e viene spiegato come il livello di misurazione di una variabile influisce sui tipi di statistiche da utilizzare. Verrà utilizzato il file di dati *demo.sav*. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice A a pag. 152.](#)

Livello di misurazione

In base al livello di misurazione, sono disponibili diverse misure riassuntive per diversi tipi di dati:

Categoriali. Dati con un numero limitato di valori o categorie diversi (ad esempio, stato civile o sesso). Vengono denominati anche **dati qualitativi**. Le variabili categoriali possono essere variabili di dati stringa (alfanumerici) o variabili numeriche che utilizzano un codice numerico per rappresentare le categorie, ad esempio 0 = *Celibe/Nubile* e 1 = *Coniugato/a*. Esistono due tipi fondamentali di dati categoriali:

- **Nominale.** Dati categoriali in cui non esiste un ordine inerente alle categorie. Ad esempio, la categoria lavorativa *vendite* non è di un ordine più alto o più basso rispetto alla categoria lavorativa *marketing* o *ricerca*.
- **Ordinale.** Dati categoriali caratterizzati da un ordine significativo di categorie, ma senza una distanza misurabile tra categorie. Ad esempio, i valori possono essere di ordine *alto*, *medio* o *basso* ma non è possibile calcolare la “distanza” effettiva esistente tra i valori.

Scala. Dati misurati in base a una scala di intervallo o di rapporto, in cui i valori dei dati indicano sia l'ordine dei valori sia la distanza tra valori. Ad esempio, un salario di € 35.000 è più elevato rispetto a un salario di € 25.000 e la distanza tra i due è di € 10.000. Vengono denominati anche dati **quantitativi** o **continui**.

Misure riassuntive per dati categoriali

Per i dati categoriali, la misura riassuntiva più tipica è il numero o la percentuale di casi presenti in ogni categoria. **Moda** è la categoria con il maggior numero di casi. Per i dati ordinali, la **mediana** (il valore sopra e sotto il quale ricade metà dei casi) potrebbe essere una misura riassuntiva utile in presenza di un elevato numero di categorie.

La procedura Frequenze genera tabelle di frequenza che visualizzano sia il numero sia la percentuale di casi per ogni valore di variabile osservato.

- ▶ Dai menu, scegliere:
Analizza > Statistiche descrittive > Frequenze...

Nota: Questa funzione richiede il modulo Statistics Base.

- Selezionare *Possiede PDA [pda]* e *Possiede TV [tv]* e spostarli nell'elenco Variabili.

Figura 5-1
Variabili categoriali selezionate per l'analisi



- Fare clic su OK per eseguire la procedura.

Figura 5-2
Tabelle di frequenza

Tabella di frequenza

Possiede PDA

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulata
Validi	No	5093	79,6	79,6	79,6
	Sì	1307	20,4	20,4	100,0
Totale		6400	100,0	100,0	

Possiede la TV

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulata
Validi	No	63	1,0	1,0	1,0
	Sì	6337	99,0	99,0	100,0
Totale		6400	100,0	100,0	

Le tabelle di frequenza vengono visualizzate nel Viewer. Le tabelle di frequenza rivelano che solo il 20,4% delle persone possiede un PDA ma quasi tutti possiedono una TV (99,0%). Tali dati potrebbero non essere molto rivelatori, ma potrebbe essere interessante scoprire qualcosa'altro sul piccolo gruppo di persone che non possiede una TV.

Grafici per i dati categoriali

È possibile visualizzare le informazioni in una tabella di frequenza con un grafico a barre o a torta.

- Verrà di nuovo aperta la finestra di dialogo Frequenze. Ad ogni modo le due variabili dovrebbero essere selezionate.

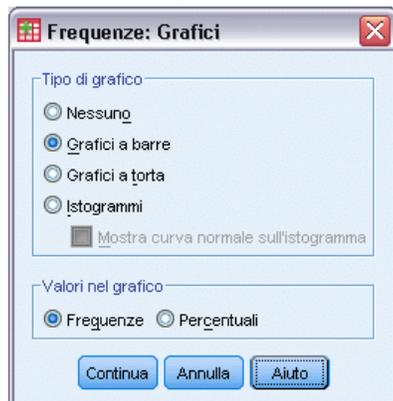
È possibile utilizzare il pulsante Richiama finestra nella barra degli strumenti per ritornare immediatamente alle procedure utilizzate recentemente.

Figura 5-3
Pulsante Richiama finestra



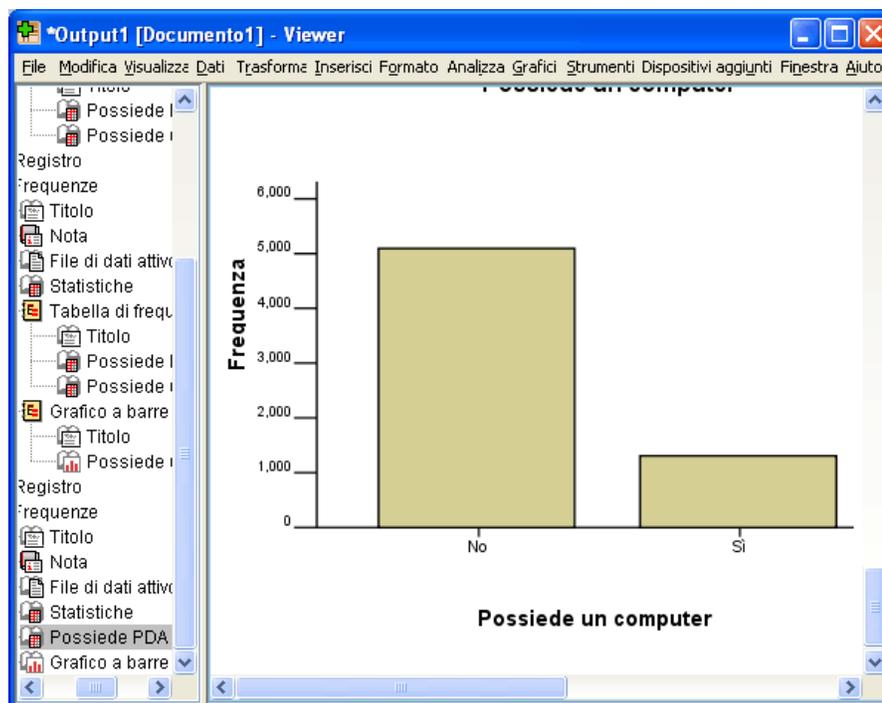
- Fare clic su Grafici.
- Selezionare Grafici a barre e quindi fare clic su Continua.

Figura 5-4
Finestra di dialogo Frequenze: Grafici.



- Fare clic su OK nella finestra di dialogo principale per eseguire la procedura.

Figura 5-5
Grafico a barre



Oltre alle tabelle di frequenza, le stesse informazioni vengono adesso visualizzate in formato di grafici a barre, in modo che risulti semplice constatare che la maggior parte delle persone non possiedono un PDA, ma quasi tutti possiedono una TV.

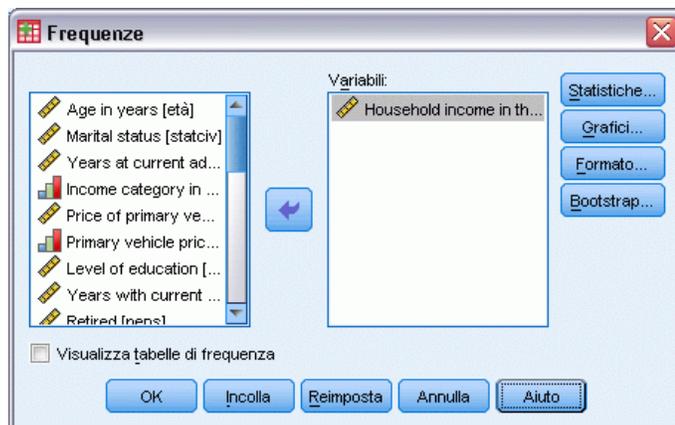
Misure riassuntive per variabili di scala

Sono disponibili molte misure riassuntive per variabili di scala, comprese le seguenti:

- **Misure di tendenza centrale.**Le misure di tendenza centrale più comuni sono la **media** (media aritmetica) e la **mediana** (valore sopra e sotto il quale ricade la metà dei casi).
 - **Misure di dispersione.**Le statistiche che misurano l'entità della variazione o della variabilità dei dati includono la deviazione standard, il valore minimo e massimo.
- ▶ Verrà di nuovo aperta la finestra di dialogo Frequenze.
 - ▶ Fare clic su Ripristina per annullare qualsiasi impostazione precedente.

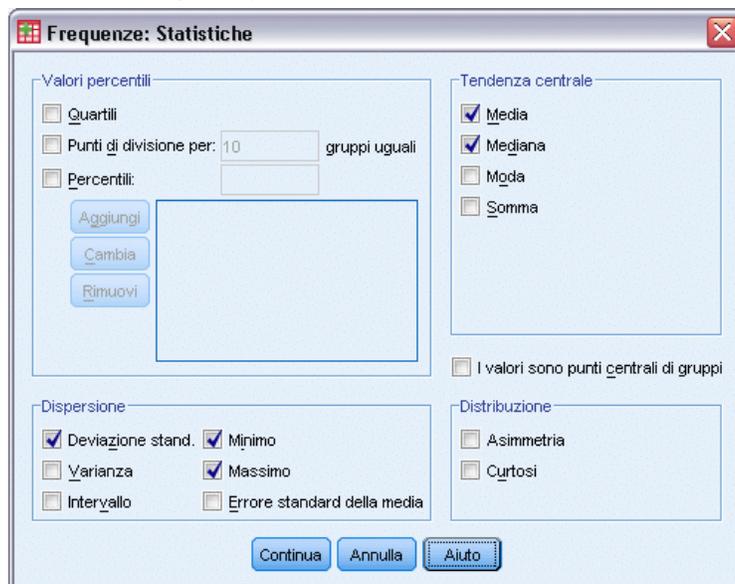
- Selezionare *Reddito familiare in migliaia [reddito]* e spostarlo nell'elenco Variabili.

Figura 5-6
Variabili di scala selezionate per l'analisi



- Fare clic su Statistiche.
- Selezionare Media, Mediana, Deviazione stand., Minimo e Massimo.

Figura 5-7
Finestra di dialogo Frequenze: Statistiche

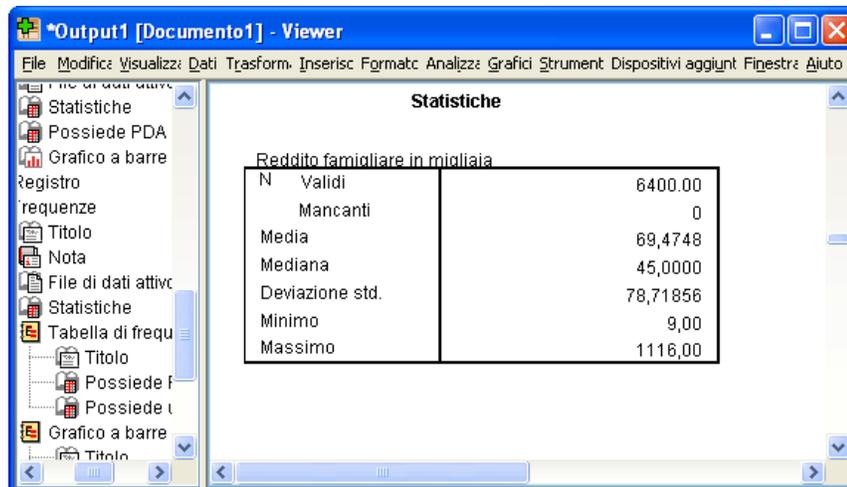


- Fare clic su Continua.
- Deselezionare Visualizza tabelle di frequenza nella finestra di dialogo principale. In genere le tabelle di frequenza non sono utili per le variabili di scala, poiché potrebbero esistere tanti valori distinti quanti sono i casi presenti nel file di dati.
- Fare clic su OK per eseguire la procedura.

La tabella Frequenze: Statistiche viene visualizzata nel Viewer.

Figura 5-8

Tabella Frequenze: Statistiche



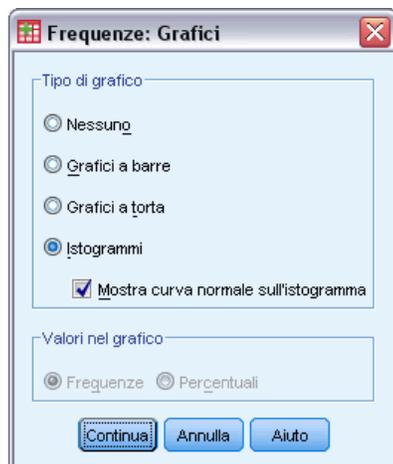
Nell'esempio vi è una grande differenza tra la media e la mediana. La media è quasi 25.000 volte più grande della mediana e ciò indica che i valori non sono distribuiti normalmente. È possibile verificare la distribuzione mediante un istogramma.

Istogrammi per variabili di scala

- ▶ Verrà di nuovo aperta la finestra di dialogo Frequenze.
- ▶ Fare clic su Grafici.
- ▶ Selezionare Istogramma e Con curva normale.

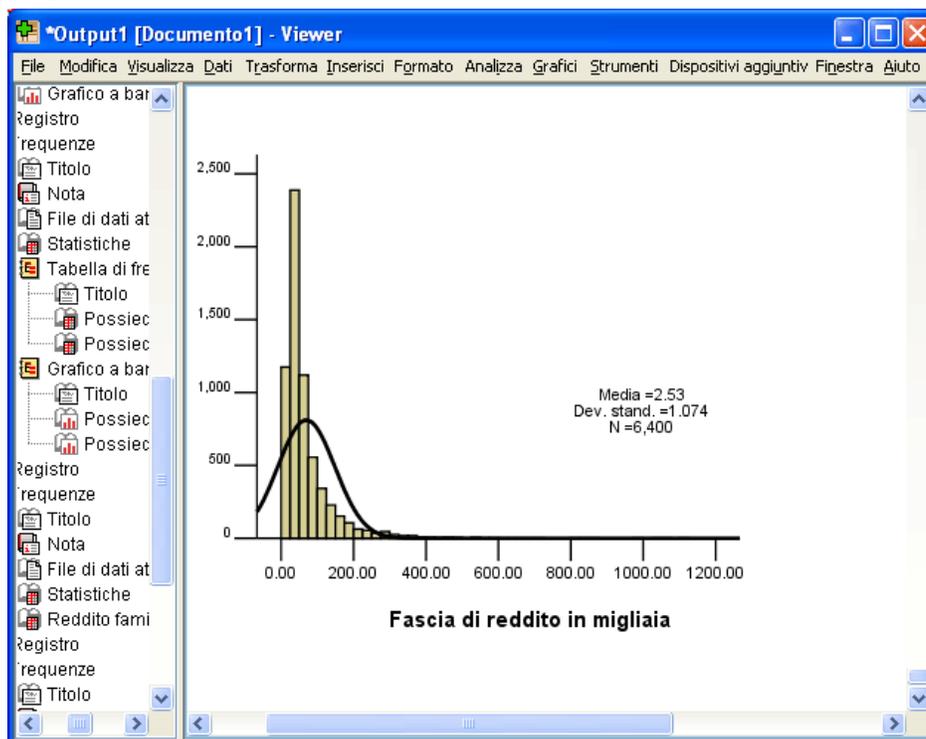
Figura 5-9

Finestra di dialogo Frequenze: Grafici.



- Fare clic su Continua, quindi fare clic su OK nella finestra di dialogo principale per eseguire la procedura.

Figura 5-10
Istogramma



La maggioranza dei casi vengono raggruppati al livello più basso della scala e molti di questi sono inferiori a 100.000. Tuttavia, esistono alcuni casi nella fascia dei 500.000 e oltre (ma il loro numero è troppo esiguo per essere visibili senza modificare l'istogramma). Tali valori alti per pochi casi soltanto hanno un effetto significativo sulla media, mentre hanno un effetto scarso o nullo sulla mediana, rendendo quest'ultima un indicatore di tendenza centrale migliore nell'esempio riportato.

Creazione e modifica di grafici

È possibile creare e modificare un'ampia varietà di tipi di grafici. In questo capitolo vengono creati e modificati grafici a barre. È possibile applicare i principi a qualsiasi tipo di grafico.

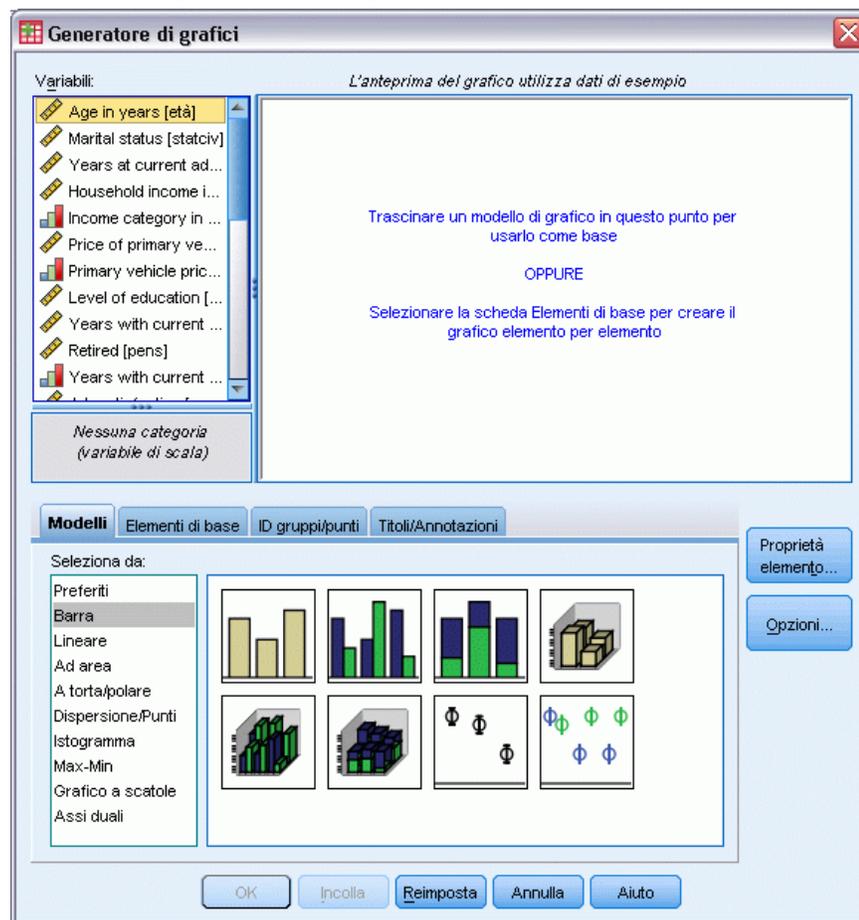
Nozioni base per la creazione di grafici

Per illustrare le nozioni base relative alla creazione di grafici, si procederà alla creazione di un grafico a barre relativo al reddito medio per diversi gradi di soddisfazione sul lavoro. In questo esempio viene utilizzato il file di dati *demo.sav*. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice A a pag. 152.](#)

- ▶ Dai menu, scegliere:
Grafici > Generatore di grafici...

Generatore di grafici è una finestra di dialogo interattiva che consente di prevedere l'aspetto futuro di un grafico durante la creazione dello stesso.

Figura 6-1
Finestra di dialogo Generatore di grafici



Utilizzo di Modelli del generatore di grafici

- Fare clic sulla scheda Modelli se non è selezionata.

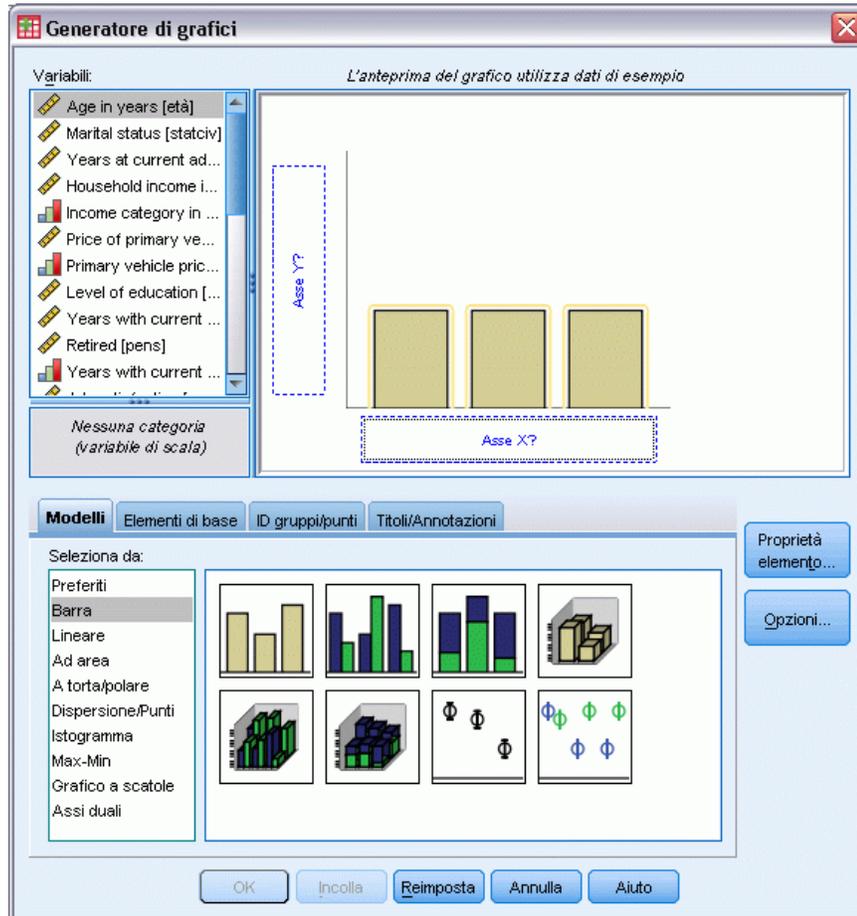
La scheda Modelli contiene vari grafici predefiniti, organizzati per tipo di grafico. Anche la scheda Elementi di base contiene gli elementi di base (quali gli assi e gli elementi del grafico) per la creazione di un grafico vuoto, ma rispetto alla scheda Modelli è più semplice da utilizzare.

- Fare clic su A barre se non è selezionata.

Le icone che rappresentano i grafici a barre disponibili in Modelli sono visualizzate nella finestra di dialogo. Le immagini devono fornire le informazioni sufficienti per identificare il tipo di grafico specifico. Se dovessero essere necessarie ulteriori informazioni, è inoltre possibile visualizzare una descrizione del grafico posizionando il cursore su un'icona.

- Trascinare l'icona del grafico a barre semplice nel disegno, ovvero nell'ampia area situata sopra i Modelli. Generatore di grafici visualizza sul disegno un'anteprima del grafico. Si noti che i dati utilizzati per tracciare il grafico non sono i dati effettivi, bensì dei dati esemplificativi.

Figura 6-2
Grafico a barre sul disegno di Generatore di grafici



Definizione delle variabili e delle statistiche

Sebbene vi sia un grafico nel disegno, esso non è completo perché è privo di variabili o di statistiche, necessarie per controllare l'altezza delle barre e per specificare quale categoria di variabili corrisponde a ogni barra. Senza variabili e statistiche, è impossibile ottenere un grafico. Per aggiungere le variabili, trascinarle dall'elenco Variabili situato sul lato sinistro del disegno.

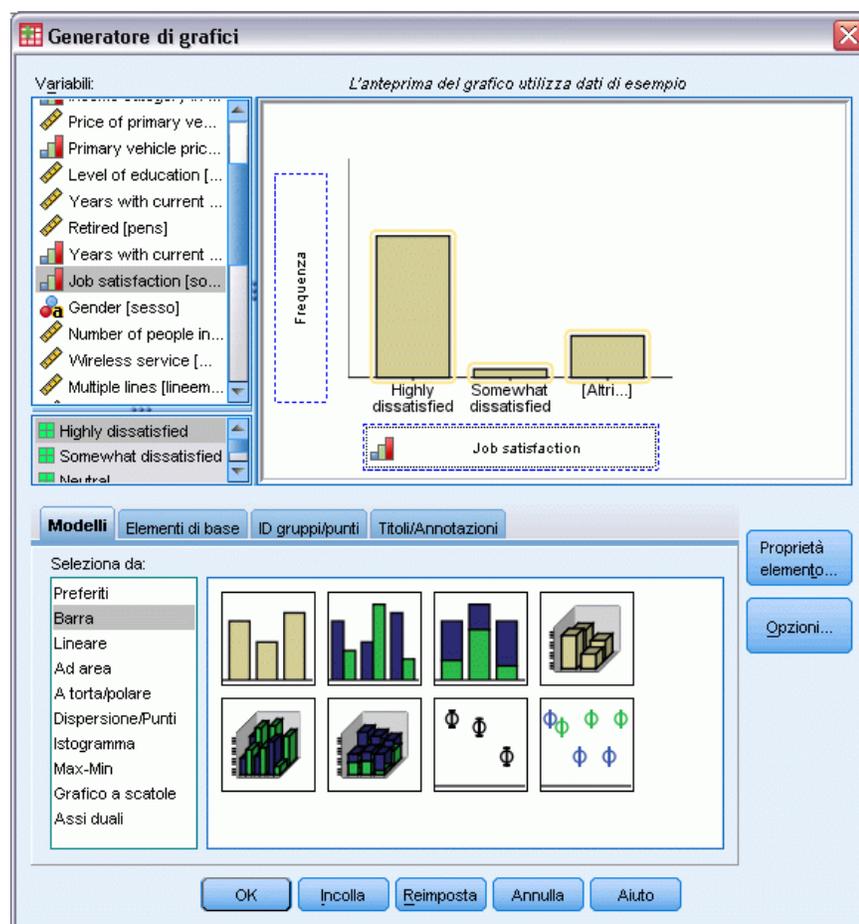
Nel Generatore di grafici è fondamentale il livello di misurazione delle variabili. Se si sta per utilizzare la variabile *Soddisfazione sul lavoro* sull'asse x , l'icona (che assomiglia a un righello) situata accanto alla variabile, indica che il livello di misurazione della variabile viene definito di scala. Per creare il grafico esatto, è necessario utilizzare un livello di misurazione categoriale. È

possibile modificare il livello di misurazione temporaneamente nel Generatore di grafici, invece di tornare indietro e modificarlo nella scheda Visualizzazione variabili.

- ▶ Fare clic con il pulsante destro del mouse su *Soddisfazione sul lavoro* nell'elenco Variabili e selezionare Ordinale. Il livello di misurazione Ordinale è appropriato perché le categorie in *Soddisfazione sul lavoro* possono essere classificate per livello di soddisfazione. Si noti che dopo aver modificato il livello di misurazione, l'icona cambia aspetto.
- ▶ Ora trascinare *Soddisfazione sul lavoro* dall'elenco Variabili nell'area di rilascio dell'asse x.

Figura 6-3

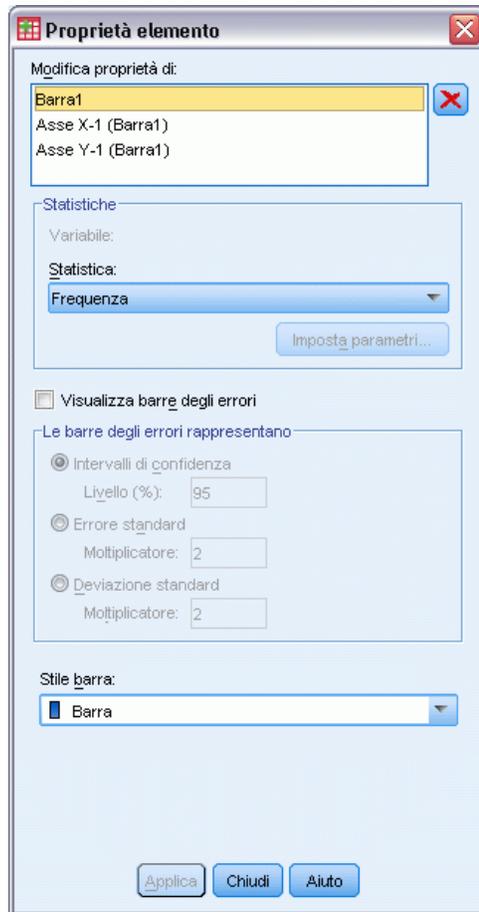
Soddisfazione sul lavoro nell'area di rilascio dell'asse x



L'area di rilascio dell'asse y è il default per la statistica *Conteggio*. Se si desidera utilizzare un'altra statistica, quale percentuale o media, è possibile farlo facilmente. Nell'esempio non si utilizzeranno queste statistiche, ma si ripasserà il processo qualora fosse necessario cambiare la statistica in un altro momento.

- ▶ Fare clic su *Proprietà elemento* per visualizzare la finestra corrispondente.

Figura 6-4
Finestra Proprietà elemento



La finestra Proprietà elemento consente di modificare le proprietà dei vari elementi del grafico. Tali elementi includono gli elementi del grafico (ad esempio le barre del grafico a barre) e gli assi del grafico. Selezionare uno degli elementi nell'elenco Modifica proprietà di per modificare le proprietà associate all'elemento. Si noti inoltre la *X* rossa situata a destra dell'elenco. Questo pulsante elimina un elemento del grafico dal disegno. Poiché A barra1 è selezionato, le proprietà illustrate vengono applicate agli elementi del grafico, in particolare alle barre.

L'elenco a discesa Statistiche mostra le statistiche specifiche disponibili. Di norma, per ogni tipo di grafico sono disponibili le stesse statistiche. Tenere presente che alcune statistiche richiedono che l'area di rilascio *y* contenga una variabile.

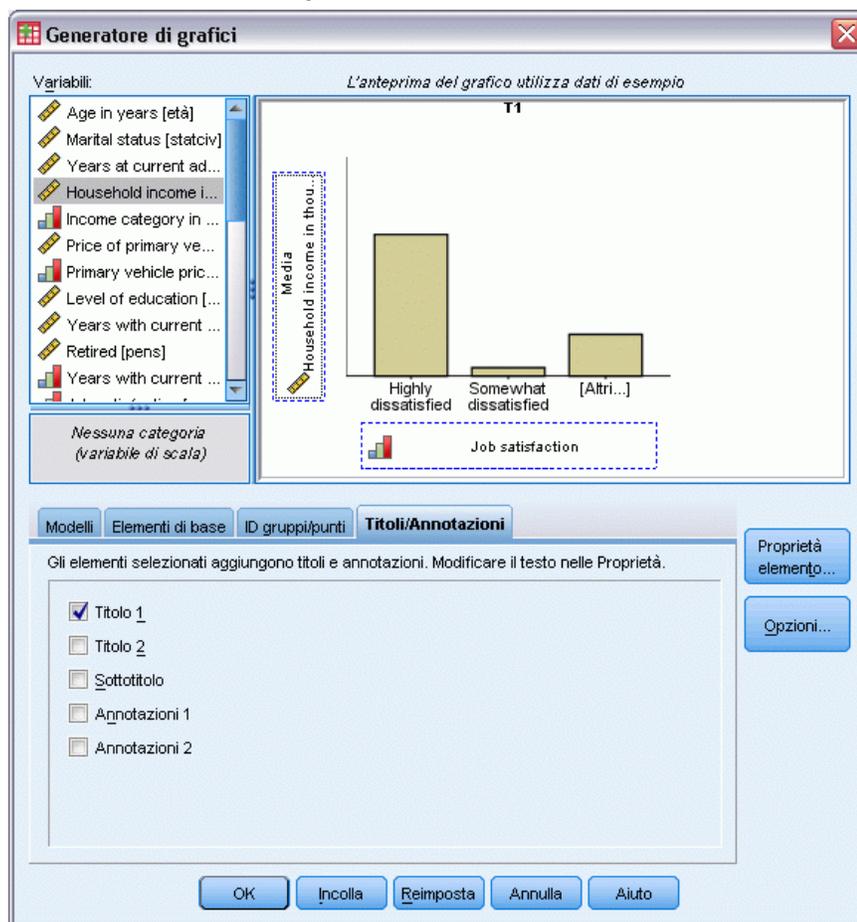
- Tornare alla finestra di dialogo Generatore di grafici e trascinare *Reddito familiare in migliaia* dall'elenco Variabili nell'area di rilascio dell'asse *y*. Poiché la variabile nell'asse *y* è scalare e la variabile nell'asse *x* è categoriale (ordinale è un tipo di livello di misurazione categoriale), l'area di rilascio dell'asse *y* è il default per la statistica *Media*. Tali variabili e statistiche sono quelle desiderate, quindi non è necessario modificare le proprietà degli elementi.

Aggiunta di testo

È inoltre possibile aggiungere titoli e annotazioni al grafico.

- ▶ Fare clic sulla scheda Titoli/Annotazioni.
- ▶ Selezionare Titolo 1.

Figura 6-5
Titolo 1 visualizzato sul disegno



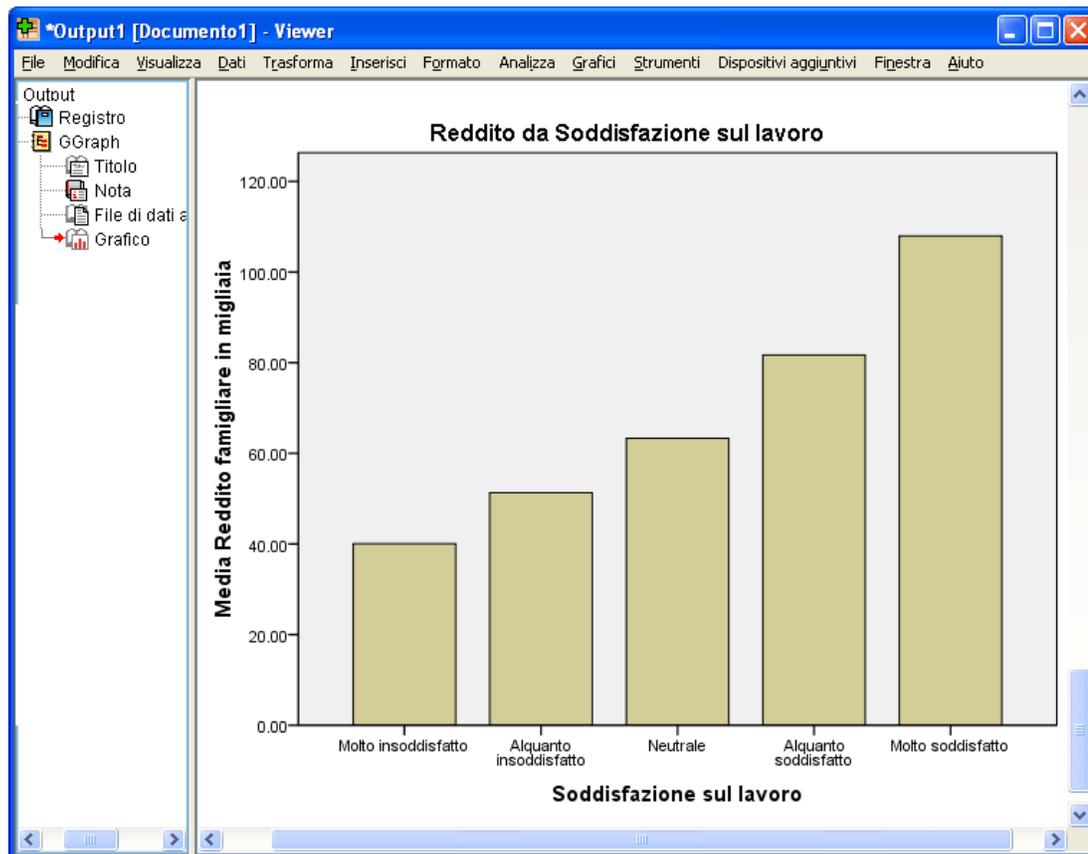
Sul disegno appare il titolo con l'etichetta T1.

- ▶ Nella finestra Proprietà elemento, selezionare Titolo 1 dall'elenco Modifica proprietà di.
- ▶ Nella casella di testo Contenuto, digitare Reddito per soddisfazione sul lavoro, che corrisponde al testo del titolo che sarà visualizzato.
- ▶ Fare clic su Applica per salvare il testo. Sebbene il testo non sia visualizzato in Generatore di grafici, quando si genera il grafico esso apparirà.

Per creare il grafico

- Fare clic su OK per creare il grafico a barre.

Figura 6-6
Grafico a barre



Il grafico a barre indica che il reddito dei rispondenti più soddisfatti del proprio lavoro tende a essere più elevato.

Nozioni base per la modifica di grafici

È possibile modificare i grafici in molti modi. Nel grafico a barre di esempio, verranno eseguite le operazioni seguenti:

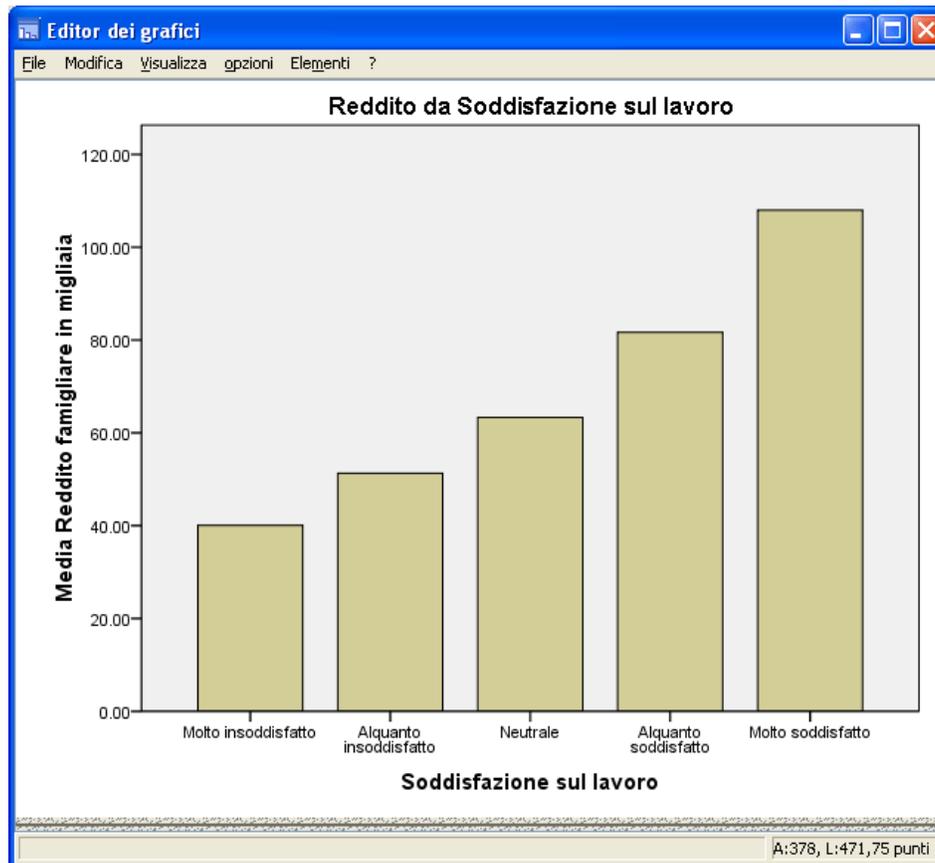
- Modificare i colori.
- Formattare i numeri in etichette delle tacche.
- Modificare il testo.
- Visualizzare le etichette dei valori.
- Utilizzare i modelli di grafici.

Per modificare il grafico, aprirlo nell'Editor dei grafici.

- Fare doppio clic sul grafico a barre per aprirlo nell'Editor dei grafici.

Figura 6-7

Grafico a barre nell'Editor dei grafici



Selezione degli elementi dei grafici

Per modificare un elemento di un grafico, è prima necessario selezionarlo.

- Fare clic su una delle barre. I rettangoli intorno alle barre indicano che sono selezionate.

Esistono regole generali per la selezione di elementi nei grafici semplici:

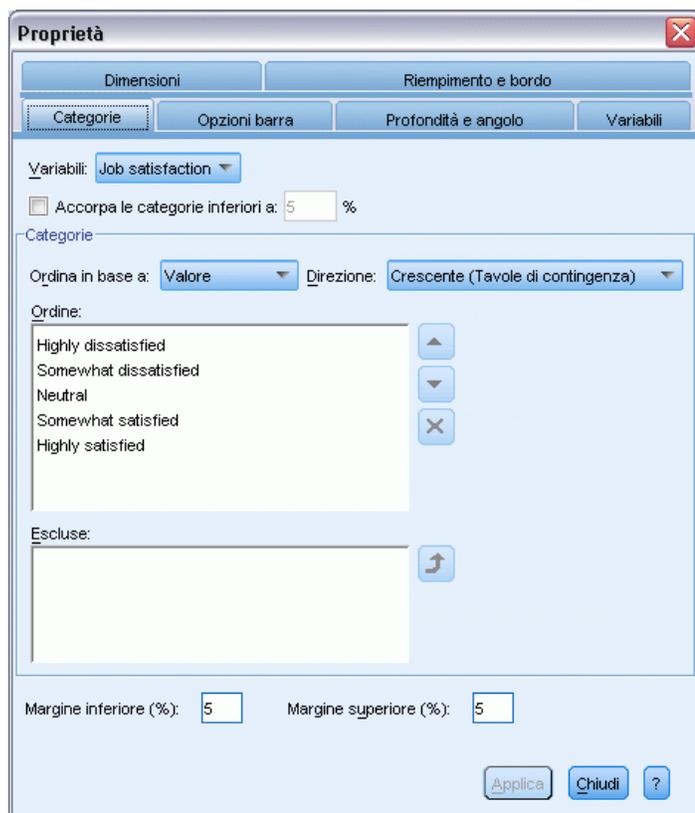
- Quando nessun elemento del grafico è selezionato, fare clic su un elemento qualsiasi per selezionarli tutti.
- Quando tutti gli elementi sono selezionati, fare clic su un elemento per selezionare solo quell'elemento. È possibile selezionare un diverso elemento facendo clic su di esso. Fare clic su un elemento tenendo premuto il tasto Ctrl per selezionare più elementi del grafico.
- Premere il tasto Esc per deselegionare tutti gli elementi.
- Fare clic su una delle barre per selezionare di nuovo tutte le barre.

Utilizzo della finestra Proprietà

- Dai menu dell'Editor dei grafici, scegliere:
Modifica > Proprietà

Viene aperta la finestra Proprietà, che mostra le schede relative alle barre selezionate. Tali schede variano in base agli elementi del grafico selezionati nell'Editor dei grafici. Ad esempio, se è stata selezionata una cornice di testo anziché una o più barre, nella finestra Proprietà si visualizzano schede differenti. Queste schede verranno utilizzate per eseguire la maggioranza delle operazioni di modifica del grafico.

Figura 6-8
Finestra Proprietà



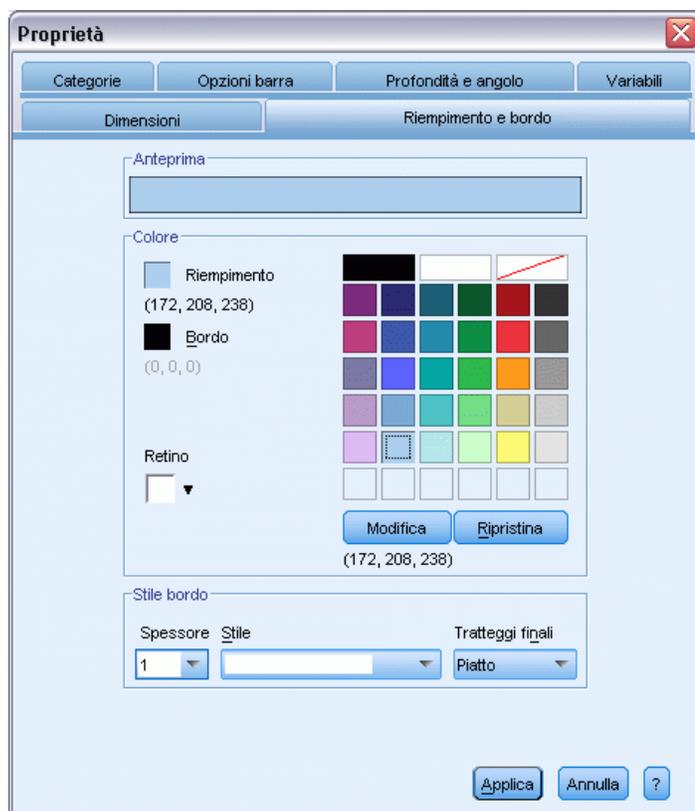
Modifica dei colori delle barre

Per prima cosa, verrà modificato il colore delle barre. L'impostazione degli attributi di colore degli elementi del grafico (esclusi linee e simboli) avviene tramite la scheda Riempimento & bordo.

- Fare clic sulla scheda Riempimento& bordo.

- Fare clic sul contrassegno accanto a Riempimento per indicare che si desidera modificare il colore di riempimento delle barre. I numeri sotto il contrassegno indicano le impostazioni relative a rosso, verde e blu per il colore corrente.
- Fare clic sul colore azzurro (il secondo da sinistra sulla seconda riga dal basso).

Figura 6-9
Scheda Riempimento & bordo

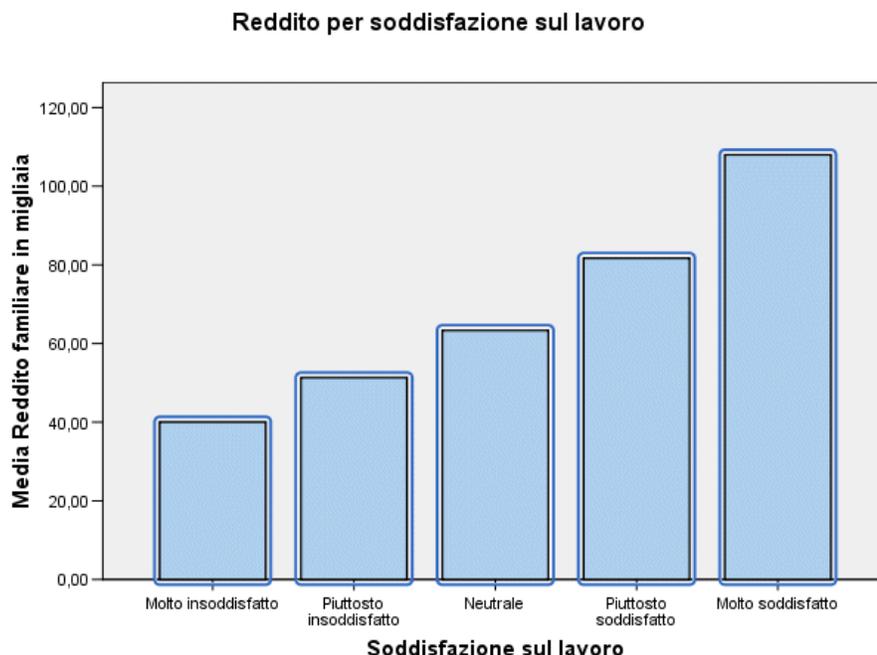


- Fare clic su Applica.

Le barre del grafico sono ora visualizzate in azzurro.

Figura 6-10

Grafico a barre modificato con barre di colore blu



Formattazione dei numeri nelle etichette delle tacche

Si noti che le cifre numeri sull'asse *y* sono in migliaia. Per rendere il grafico più gradevole e facile da interpretare, si modificherà il formato dei numeri nelle etichette delle tacche e quindi il titolo dell'asse di conseguenza.

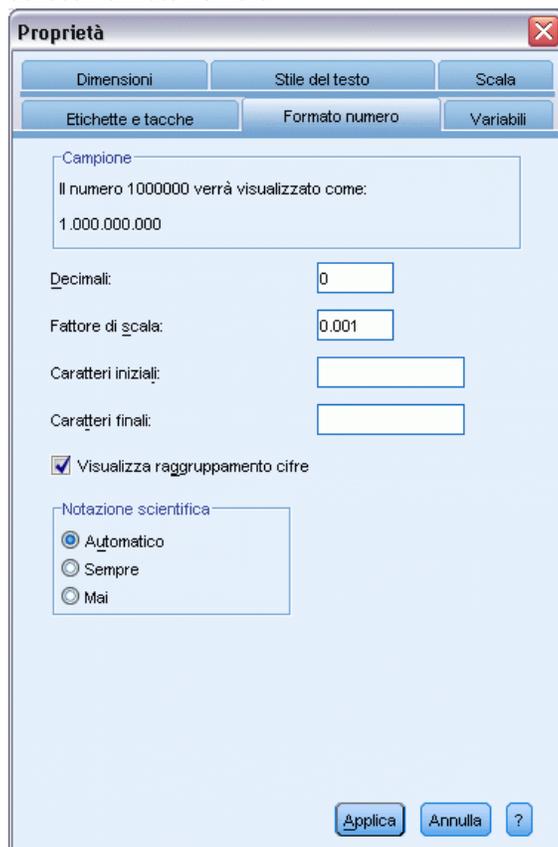
- ▶ Selezionare le etichette delle tacche dell'asse *y* facendo clic su una di esse.
- ▶ Per riaprire la finestra Proprietà (in caso sia stata chiusa in precedenza) dai menu scegliere: Modifica > Proprietà

Nota: Da questo punto si presupporrà che la finestra Proprietà sia aperta. Se la finestra Proprietà è stata chiusa, eseguire il passaggio precedente per riaprirla. Per riaprire la finestra è inoltre possibile utilizzare la scelta rapida da tastiera Ctrl+T.

- ▶ Fare clic sulla scheda Formato numero.
- ▶ Per evitare che sulle etichette delle tacche siano visualizzate cifre decimali, digitare 0 nella casella di testo Decimali.
- ▶ Digitare 0,001 nella casella di testo Fattore di scala. Il fattore di scala è il numero per il quale l'Editor dei grafici divide il numero visualizzato. Poiché 0,001 è una frazione, la divisione per tale numero *aumenterà* i numeri sulle etichette delle tacche di 1.000. Di conseguenza, i numeri non saranno più espressi in migliaia e saranno non in scala.

- Selezionare Visualizza raggruppamento cifre. Il raggruppamento cifre utilizza un carattere (specificato dalle impostazioni internazionali del computer in uso) come separatore delle migliaia all'interno del numero.

Figura 6-11
Scheda Formato numero

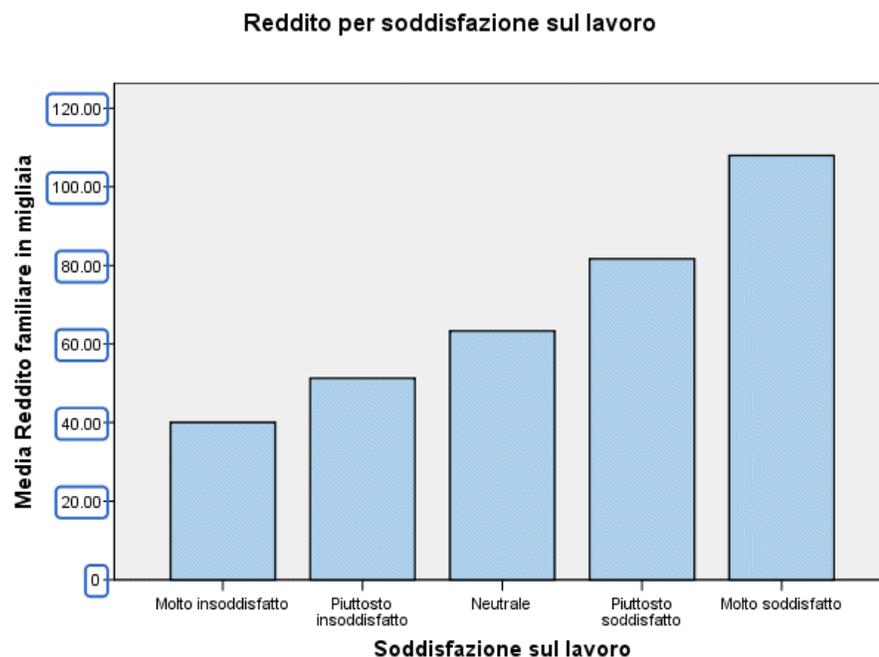


- Fare clic su Applica.

Le etichette delle tacche riflettono la nuova formattazione dei numeri: non sono presenti cifre decimali, i numeri non sono più in scala e sono presenti i separatori delle migliaia.

Figura 6-12

Grafico a barre con formattazione dei numeri modificata



Modifica di testo

Dopo la modifica del formato dei numeri per le etichette delle tacche, il titolo dell'asse risulta ora non corretto. La prossima operazione consisterà nel modificare il titolo dell'asse per riflettere il nuovo formato dei numeri.

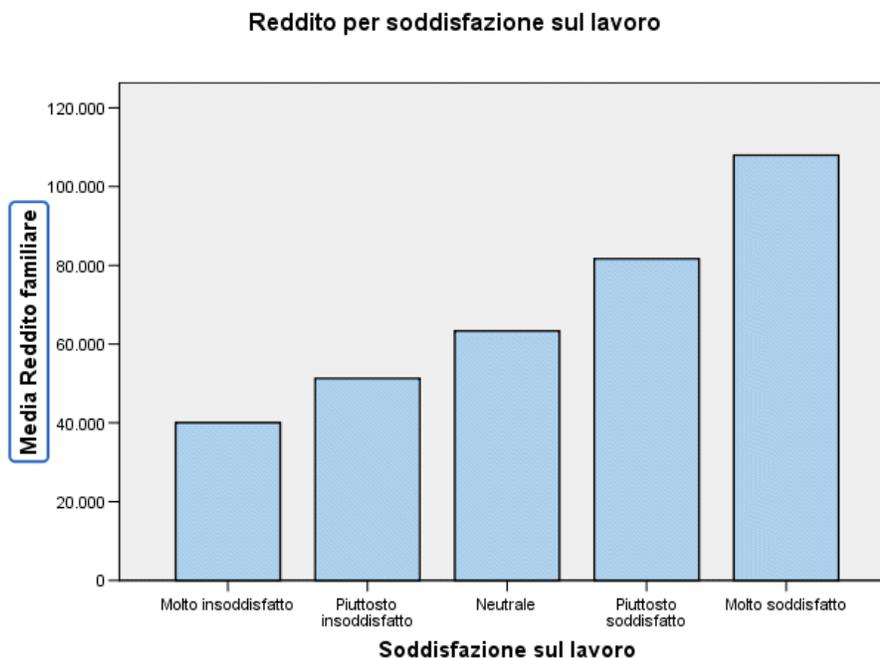
Nota: non è necessario aprire la finestra Proprietà per modificare il testo. È possibile modificare il testo direttamente nel grafico.

- ▶ Fare clic sul titolo dell'asse *y* per selezionarlo.
- ▶ Fare di nuovo clic sul titolo dell'asse per attivare la modalità di modifica. Quando è attiva la modalità di modifica, l'Editor dei grafici colloca in orizzontale il testo ruotato. Viene inoltre visualizzato un cursore lampeggiante a forma di linea rossa (non presente nell'esempio).

- ▶ Eliminare il testo seguente:
in migliaia
- ▶ Premere Invio per uscire dalla modalità di modifica e aggiornare il titolo dell'asse. Il titolo dell'asse ora descrive in modo preciso il contenuto delle etichette delle tacche.

Figura 6-13

Grafico a barre con titolo dell'asse y modificato

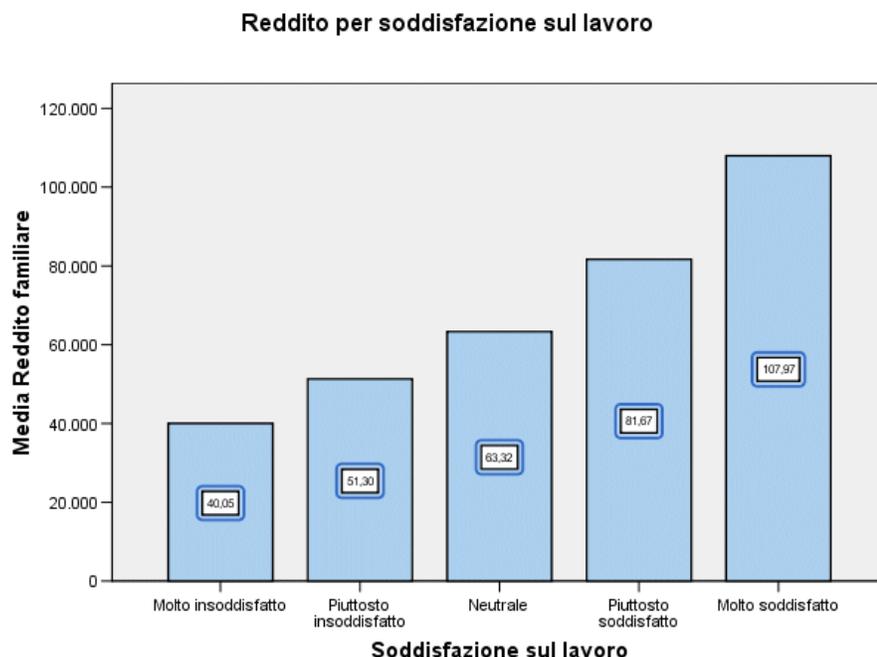


Visualizzazione delle etichette dei valori

Un'altra operazione comune è la visualizzazione dei valori esatti associati agli elementi del grafico (nell'esempio, le barre). Questi valori vengono visualizzati tramite le etichette dei dati.

- ▶ Dai menu dell'Editor dei grafici, scegliere:
Elementi > Mostra le etichette dei dati

Figura 6-14
Etichette dei valori visualizzate nel grafico a barre



In ogni barra del grafico è ora visualizzato l'esatto reddito familiare medio. Si noti che le unità sono in migliaia, quindi è possibile utilizzare di nuovo la scheda Formato numero per modificare il fattore di scala.

Utilizzo dei modelli

Se si effettuano una serie di modifiche di routine ai grafici, è possibile modificare un modello di grafico per ridurre il tempo necessario alla creazione e alla modifica di grafici. In un modello di grafico vengono salvati gli attributi di un grafico specifico. È quindi possibile applicare il modello al momento della creazione o della modifica di un grafico.

Il grafico corrente verrà salvato come modello e quindi applicato per creare un nuovo grafico.

- Dai menu, scegliere:
File > Salva modello...

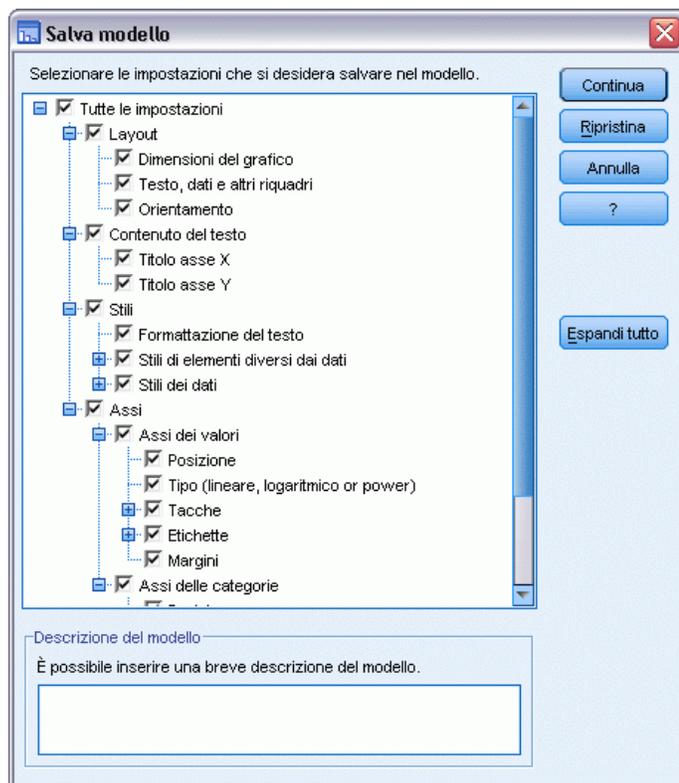
Nella finestra di dialogo Salva modello è possibile specificare gli attributi del grafico che si desidera includere nel modello.

Espandere uno o più elementi nella struttura ad albero per visualizzare gli attributi specifici che è possibile salvare con il grafico. Ad esempio, se si espande la parte della struttura relativa agli assi dei valori è possibile visualizzare tutti gli attributi delle etichette dei valori che saranno inclusi nel modello. Selezionare uno o più attributi per includerli nel modello.

- Selezionare Tutte le impostazioni per includere nel modello tutti gli attributi del grafico disponibili.

È inoltre possibile inserire una descrizione del modello. La descrizione sarà visibile al momento dell'applicazione del modello.

Figura 6-15
Finestra di dialogo Salva modello



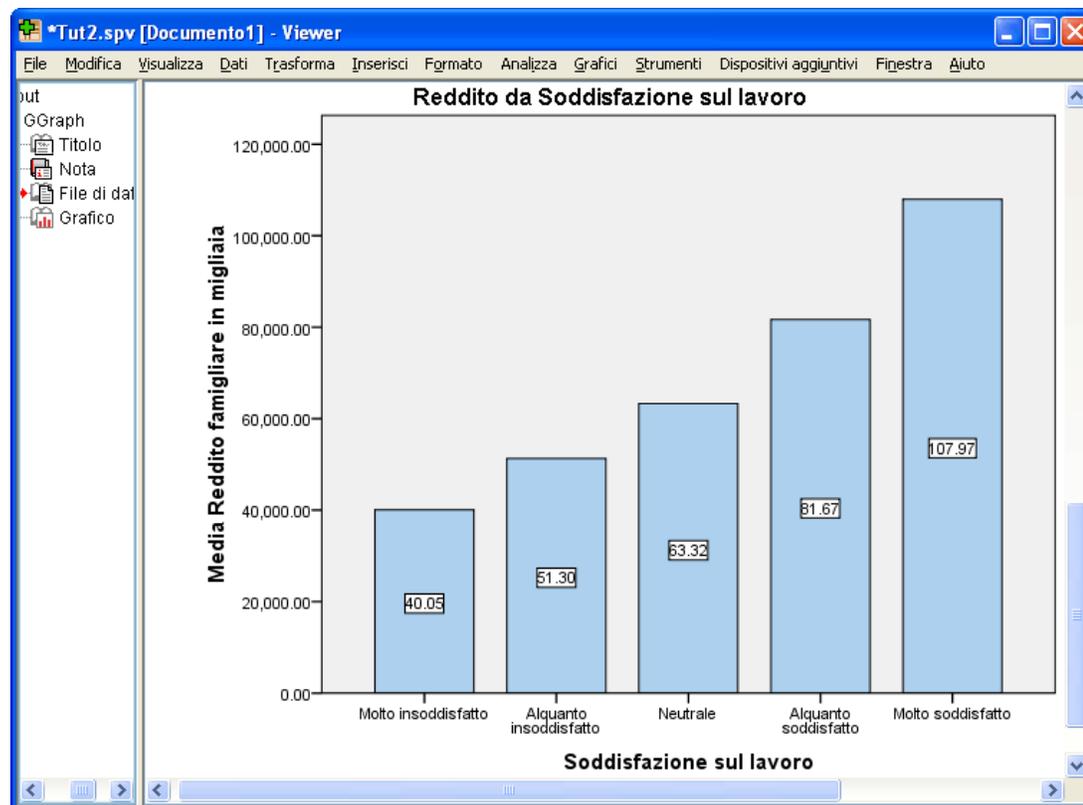
- ▶ Fare clic su Continua.
- ▶ Nella finestra di dialogo Salva modello specificare un percorso e un nome di file per il modello.
- ▶ Al termine dell'operazione, fare clic su Salva.

È possibile applicare il modello quando si crea un grafico o nell'Editor dei grafici. Nell'esempio seguente, l'applicazione del modello avverrà durante la creazione di un grafico.

- Chiudere l'Editor dei grafici. Il grafico a barre aggiornato viene visualizzato nel Viewer.

Figura 6-16

Grafico a barre aggiornato nel Viewer



- Dai menu del Viewer scegliere:
Grafici > Generatore di grafici...

Nella finestra di dialogo Generatore di grafici sono visualizzate automaticamente le variabili inserite al momento della creazione del grafico originale. Tuttavia, in questo caso si creerà un grafico leggermente diverso per verificare in che modo l'applicazione di un modello definisce la formattazione di un grafico.

- Rimuovere *Soddifazione sul lavoro* dall'asse x trascinandola dall'area di rilascio di nuovo nell'elenco Variabili. È inoltre possibile fare clic sull'area di rilascio e premere Elimina.
- Fare clic con il pulsante destro del mouse su *Livello di istruzione* nell'elenco Variabili e selezionare Ordinale.
- Trascinare *Livello di istruzione* dall'elenco Variabili nell'area di rilascio dell'asse x.

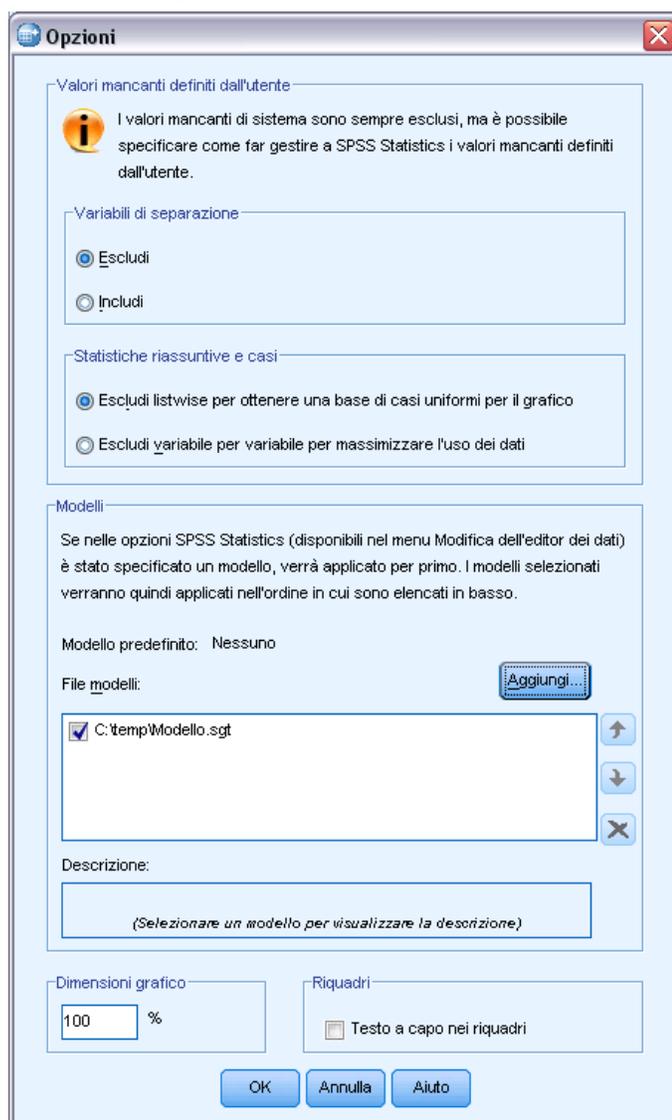
Poiché adesso il titolo non è esatto, è necessario eliminarlo.

- Nella scheda Titoli/Annotazioni deselezionare Titolo 1.

Ora si specifica il modello da applicare al grafico nuovo.

- ▶ Fare clic su Opzioni.
- ▶ Nel gruppo Modelli della finestra di dialogo Opzioni, fare clic su Aggiungi.
- ▶ Nella finestra di dialogo Trova file del modello, individuare il file del modello salvato in precedenza utilizzando la finestra di dialogo Salva modello.
- ▶ Selezionare il file e fare clic su Apri.

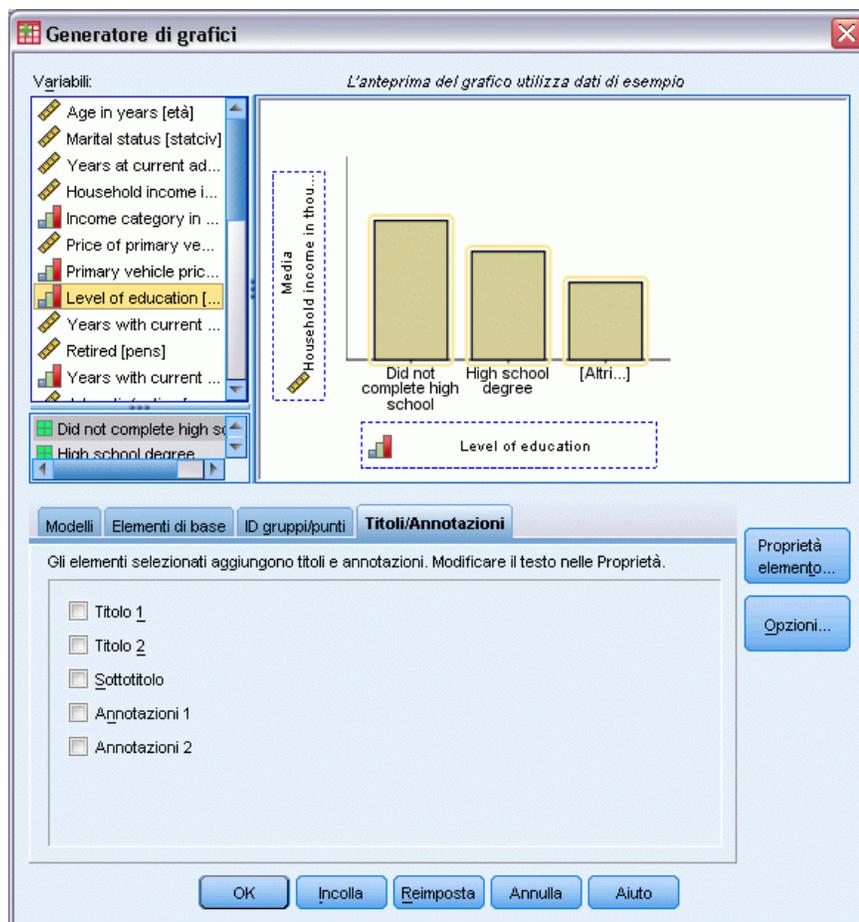
Figura 6-17
Finestra di dialogo Opzioni con modello



Nella finestra di dialogo Opzioni è visualizzato il percorso del file del modello selezionato.

- ▶ Fare clic su OK per chiudere la finestra di dialogo Opzioni.

Figura 6-18
Generatore di grafici con aree di rilascio completate

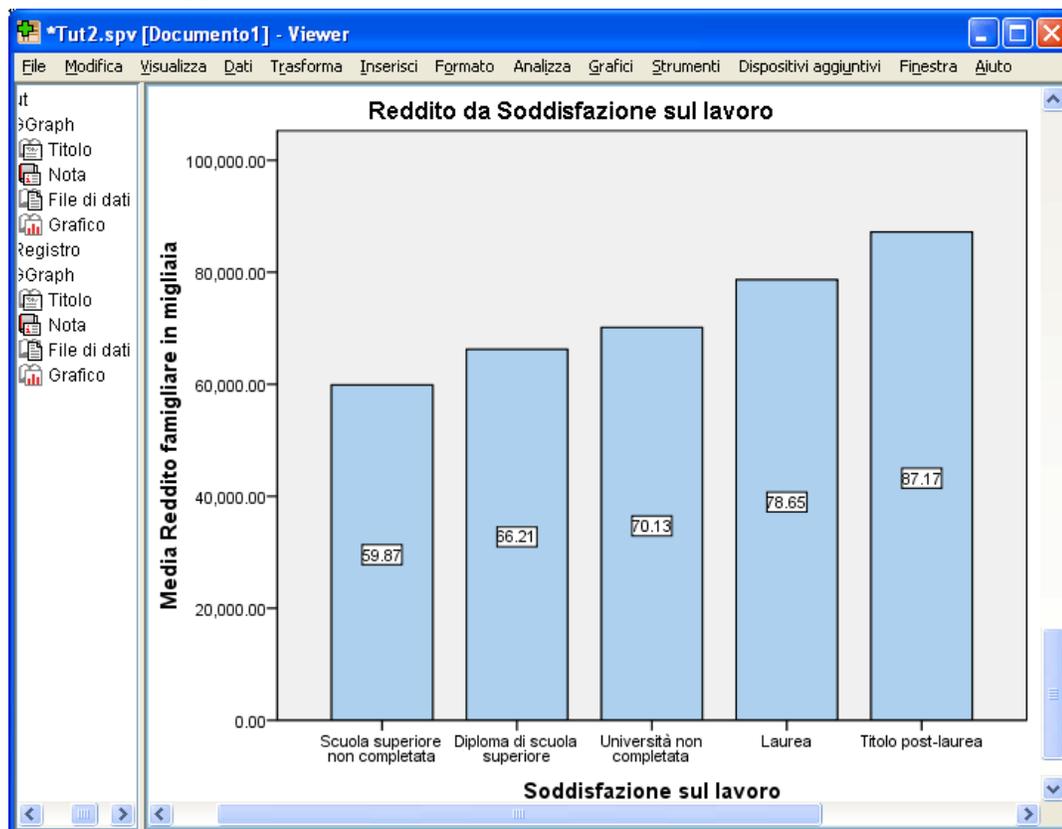


- Nella finestra di dialogo Generatore di grafici, fare clic su OK per creare il grafico e applicare il modello.

La formattazione del nuovo grafico corrisponde a quella del grafico creato e salvato in precedenza. Sebbene le variabili sull'asse x siano diverse, il grafico per il resto è analogo all'altro. Si noti che nel modello è stato mantenuto il titolo del grafico precedente anche se in Generatori di grafici lo si aveva eliminato.

È possibile applicare dei modelli dopo la creazione di un grafico utilizzando l'Editor dei grafici (dal menu File scegliere Applica modello grafico).

Figura 6-19
Grafico a barre aggiornato nel Viewer



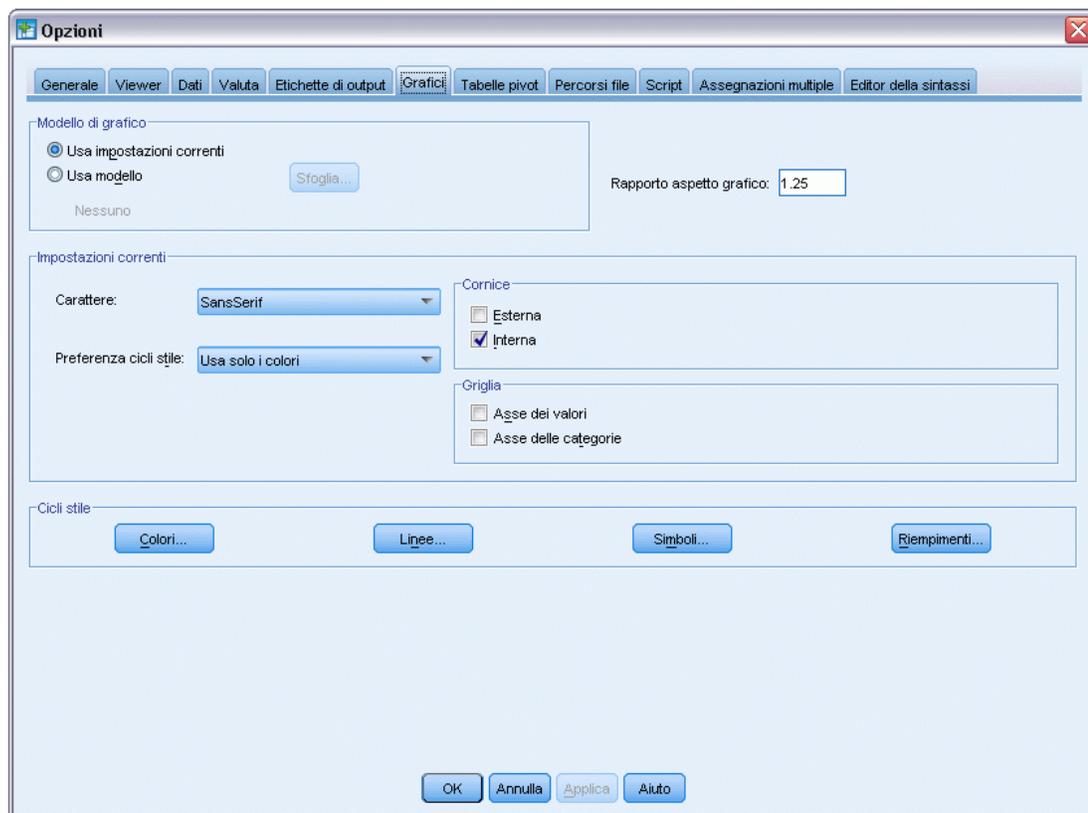
Definizione delle opzioni del grafico

Oltre a utilizzare i modelli per formattare i grafici, è possibile utilizzare le opzioni per controllare i vari aspetti delle modalità di creazione dei grafici.

- Dai menu dell'Editor dei dati o del Viewer, scegliere:
Modifica > Opzioni...

La finestra di dialogo Opzioni include varie impostazioni di configurazione. Fare clic sulla scheda Grafici per visualizzare le opzioni disponibili.

Figura 6-20
Scheda Grafici della finestra di dialogo Opzioni



Le opzioni controllano la modalità di creazione di un grafico. Per ogni nuovo grafico è possibile specificare:

- l'utilizzo o meno delle impostazioni correnti o di un modello.
- Il rapporto tra altezza e larghezza (proporzioni).
- Se non si utilizza un modello, le impostazioni di formattazione.
- I cicli stile per elementi del grafico.

I cicli stile consentono di specificare lo stile degli elementi nei nuovi grafici. Nell'esempio corrente verranno esaminati i dettagli del ciclo stile dei colori.

- ▶ Fare clic su Colori per aprire la finestra di dialogo Colori degli elementi dei dati.

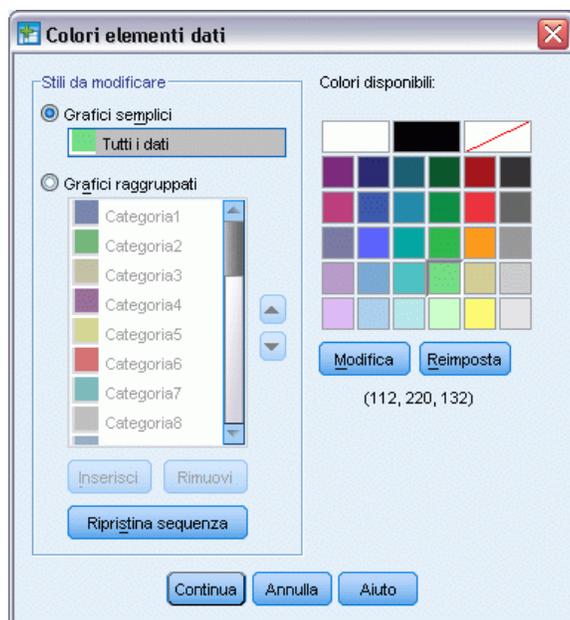
Per un grafico semplice, l'Editor dei grafici utilizza uno stile specificato dall'utente. Per i grafici raggruppati, viene utilizzato un gruppo di stili all'interno del quale l'Editor dei grafici scorre per ciascun gruppo (categoria) del grafico.

- ▶ Selezionare Grafici semplici.

- Selezionare il colore verde chiaro (il terzo da destra, sulla seconda riga a partire dal basso).

Figura 6-21

Finestra di dialogo Colori degli elementi dei dati



- Fare clic su Continua.
- Nella finestra di dialogo Opzioni fare clic su OK per salvare le modifiche al ciclo stile dei colori.

Agli elementi nei grafici semplici di nuova creazione sarà ora applicato il colore verde chiaro.

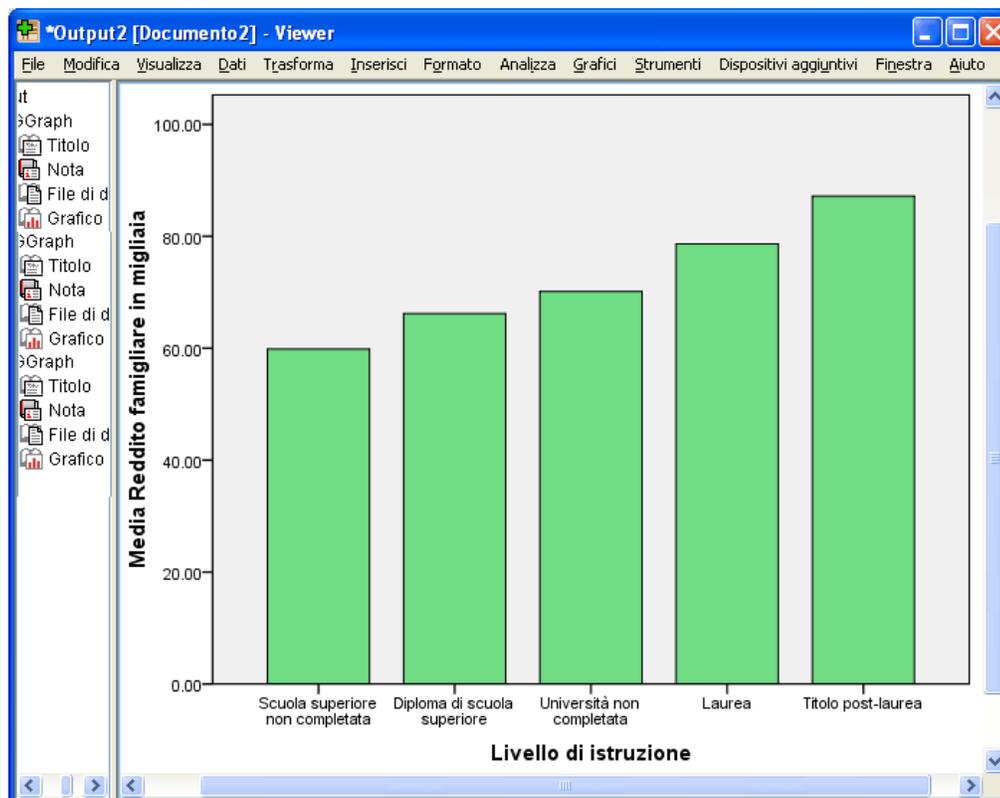
- Dai menu dell'Editor dei dati o del Viewer, scegliere:
Grafici > Generatore di grafici...

Il Generatore di grafici visualizza l'ultimo grafico creato. Si tenga presente che a questo grafico era associato un modello. Questo modello quindi, non verrà più utilizzato.

- Fare clic su Opzioni.
- Deselezionare il modello aggiunto precedentemente. Si noti che per eliminare il modello, si può anche fare clic sulla X rossa. Se si deseleziona il modello, invece di eliminarlo, esso rimane disponibile per un eventuale uso futuro.
- E fare clic su OK per creare il grafico.

Le barre del nuovo grafico sono visualizzate in verde chiaro. Questo grafico si differenzia dall'ultimo per aspetti diversi: non è presente alcun titolo, le etichette degli assi sono espresse in migliaia e non vi sono etichette dei dati. Tali differenze sono dovute alla mancata applicazione del modello nel grafico.

Figura 6-22
Grafico a barre aggiornato nel Viewer

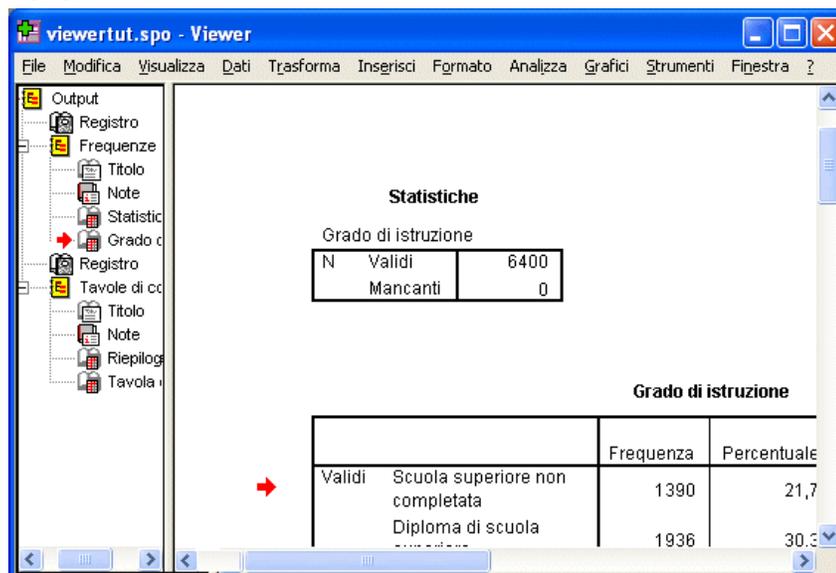


Utilizzo dell'output

I risultati dell'esecuzione di una procedura statistica vengono visualizzati nel Viewer. L'output può essere generato sottoforma di tabella, grafico o testo, a seconda delle opzioni scelte per l'esecuzione della procedura. In questo capitolo vengono utilizzati i file *viewertut.spv* e *demo.sav*. Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice A a pag. 152.

Utilizzo del Viewer

Figura 7-1
Viewer



La finestra del Viewer è suddivisa in due riquadri. La **vista riassuntiva**, che contiene tutte le informazioni memorizzate nel Viewer. L'**output**, che contiene tabelle statistiche, grafici e output testuale.

Utilizzare le barre di scorrimento per spostarsi nel contenuto di una finestra sia in senso verticale che orizzontale. Per una navigazione più facile, fare clic su un elemento nella vista riassuntiva per visualizzarlo nell'output.

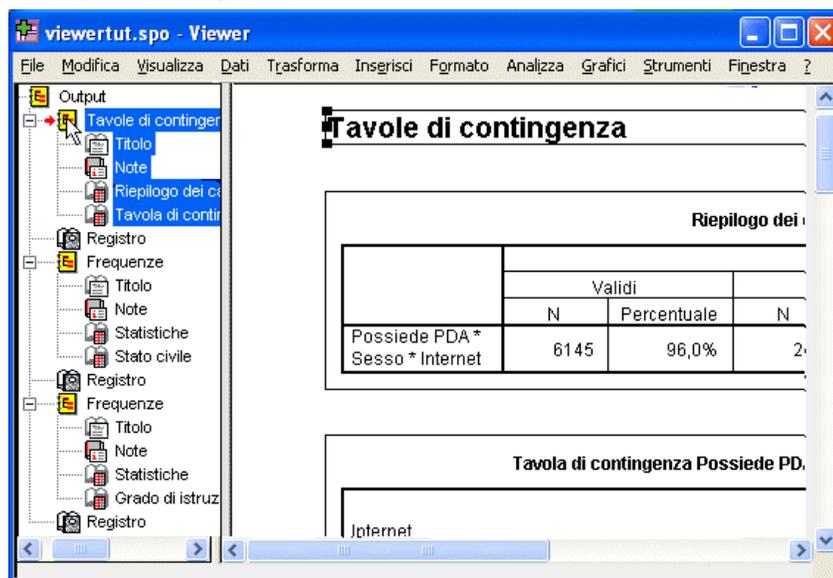
Se si ritiene che lo spazio nel Viewer sia insufficiente per poter visualizzare tutta una tabella o se la vista riassuntiva è troppo stretta, è possibile ridimensionare la finestra.

- Fare clic e trascinare il bordo destro della vista riassuntiva per modificarne l'ampiezza.

L'icona di un libro aperto nella vista riassuntiva indica che è correntemente visibile nel Viewer, sebbene in quel momento non si trovi nella parte visibile dell'output.

- ▶ Per nascondere una tabella o un grafico, fare doppio clic sull'icona a forma di libro nel riquadro della vista riassuntiva.
L'icona a forma di libro aperto si trasforma in libro chiuso, ad indicare che l'informazione ad esso associata è nascosta.
- ▶ Per visualizzare nuovamente l'output nascosto, fare doppio clic sull'icona a forma di libro chiuso. È possibile, inoltre, nascondere tutto l'output di una procedura statistica specifica oppure tutto l'output nel Viewer.
- ▶ Fare clic sulla casella con il segno meno (–) a sinistra della procedura di cui si desidera nascondere i risultati, oppure sulla casella accanto all'elemento più in alto nella vista riassuntiva per nascondere l'intero contenuto del Viewer.
La vista riassuntiva viene compressa, ad indicare che i risultati vengono nascosti.
È possibile modificare l'ordine in cui viene visualizzato l'output.
- ▶ Nel riquadro della vista riassuntiva, fare clic sugli elementi che si desidera spostare.
- ▶ Trascinare gli elementi selezionati in una posizione nuova della vista riassuntiva e rilasciare il pulsante del mouse.

Figura 7-2
Nuova posizione degli elementi nel Viewer



È anche possibile spostare gli elementi dell'output selezionandoli e trascinandoli nel riquadro dell'output.

Utilizzo dell'Editor di tabelle pivot

I risultati di molte procedure statistiche vengono visualizzati nelle **tabelle pivot**.

Accesso alle definizioni dell'output

Molti termini statistici vengono visualizzati nell'output. È possibile accedere alle definizioni relative a tali termini nel Viewer.

- ▶ Fare doppio clic sulla tabella *Tavola di contingenza Possiede PDA * Sesso * Internet*.
- ▶ Fare clic con il pulsante destro del mouse su *Conteggio atteso* e selezionare Guida rapida dal menu di scelta rapida.

La definizione viene visualizzata in una finestra popup.

Figura 7-3
Definizione popup

viewertut.spv [Documento1] - Viewer

File Modifica Visualizza Inserisci Pivot Formato Analizza Grafici Strumenti Dispositivi aggiuntivi Finestra Aiuto

Tavola di contingenza Possiede PDA * Sesso * Internet

				Sesso			
				Femmina	Maschio	Totale	
Internet	No	Possiede PDA	No	Conteggio	1897	1962	3859
			Conteggio atteso	1903.4	1955.6	3859.0	
		% entro Possiede PDA	49.2%	50.8%	100.0%		
		Totale		Conteggio	327	323	650
				Conteggio atteso	320.6	329.4	650.0
			% entro Possiede PDA	50.3%	49.7%	100.0%	
			Totale	Conteggio	2224	2285	4509
			Conteggio atteso	2224.0	2285.0	4509.0	
			% entro Possiede PDA	49.3%	50.7%	100.0%	
			Sì	Conteggio	513	512	1025
		Conteggio atteso	512.5	512.5	1025.0		
		% entro Possiede PDA	50.0%	50.0%	100.0%		
		Sì	Conteggio	305	306	611	
		Conteggio atteso	305.5	305.5	611.0		

Numero di casi previsti nella cella se le variabili di riga e di colonna sono statisticamente indipendenti o non correlate tra loro.

Table pivot

La presentazione delle informazioni nelle tabelle predefinite prodotte potrebbe non essere sufficientemente chiara o ordinata. Con le tabelle pivot, è possibile trasporre righe e colonne (“capovolgere” la tabella), regolare l’ordine dei dati in una tabella e modificare la tabella in molti altri modi. Ad esempio, è possibile modificare una tabella larga e corta in una lunga e stretta trasponendo righe e colonne. Modificare il layout della tabella non influisce sul risultato. Si tratta semplicemente di un meccanismo per visualizzare le informazioni in un modo diverso, più consono alle proprie aspettative.

- ▶ Se non è già attivata, attivare con un doppio clic la tabella *Tavola di contingenza Possiede di PDA * Sesso * Internet*.
- ▶ Se la finestra Pivoting non è visibile, dai menu selezionare:
Pivot > Pivoting

La funzionalità Pivoting consente di spostare i dati tra colonne, righe e strati.

Figura 7-4
Pivoting

Internet				Sesso		
				Femmina	Maschio	Totale
No	Possiede PDA	No	Conteggio	1897	1962	3859
			Conteggio atteso	1903.4	1955.6	3859.0
			% entro Possiede PDA	49.2%	50.8%	100.0%
Si	Possiede PDA		Conteggio	327	323	650

- Trascinare l'elemento *Statistiche* dalla dimensione Riga alla dimensione Colonna, sotto *Sesso*. La tabella viene immediatamente riconfigurata per riflettere le modifiche apportate.

Figura 7-5
Spostamento di righe su colonne

Internet				Sesso							
				Femmina			Maschio			Totale	
			Conteggio	Conteggio atteso	% entro Possiede PDA	Conteggio	Conteggio atteso	% entro Possiede PDA	Conteggio	Conteggio atteso	% entro Possiede PDA
No	Possiede PDA	No	1897						3859	3859.0	100.0%
		Si	327						650	650.0	100.0%
		Totale	2224						4509	4509.0	100.0%
Si	Possiede PDA	No	513						1025	1025.0	100.0%
		Si	305						611	611.0	100.0%
		Totale	818						1636	1636.0	100.0%

L'ordine degli elementi nel pivoting riflette l'ordine degli elementi nella tabella.

- Trascinare l'elemento *Possessore di PDA* prima dell'elemento *Internet* nella dimensione Riga per invertire l'ordine di queste due righe.

Figura 7-6
Scambiare righe

		Sesso								
		Femmina			Maschio			Totale		
Possiede PDA	Internet	Conteggio	Conteggio atteso	% entro Possiede PDA	Conteggio	Conteggio atteso	% entro Possiede PDA	Conteggio	Conteggio atteso	% entro Possiede PDA
			No	No	189					
	Si	51						1025	1025.0	100.0%
Si	No	32						650	650.0	100.0%
	Si	30						611	611.0	100.0%
Totale	No	222						4509	4509.0	100.0%
	Si	81						1636	1636.0	100.0%

Creazione e visualizzazione di strati

Gli strati sono utili nelle tabelle di grandi dimensioni con categorie di informazioni nidificate. La creazione di strati semplifica l'aspetto della tabella facilitandone la lettura.

- Trascinare l'elemento *Sesso* dalla dimensione Colonna alla dimensione Strato.

Figura 7-7
L'icona pivot *Sesso* nella dimensione *Strato*

		Sesso			
		Femmina			
Possiede PDA	Internet	Conteggio	Conteggio atteso	% entro Possiede PDA	
		No	No	1897	1903.4
	Si				
Si	No				
	Si				
Totale	No				
	Si				

Per visualizzare uno strato differente, selezionare una categoria dall'elenco a discesa della tabella.

Figura 7-8
Scelta di uno strato

Sesso		Internet	Conteggio	Conteggio atteso	% entro Possiede PDA
Possiede PDA	Totale	No	1897	1903.4	49.2%
		Sì	513	512.5	50.0%
	Sì	No	327	320.6	50.3%
		Sì	305	305.5	49.9%
	Totale	No	2224	2224.0	49.3%
		Sì	818	818.0	50.0%

Modifica di tabelle

A meno che non venga creato un TableLook personalizzato, le tabelle pivot vengono create con formattazione standard. È possibile modificare la formattazione di qualsiasi testo presente in una tabella. È possibile modificare il nome, la dimensione e lo stile (grassetto o corsivo) del carattere, nonché il colore.

- ▶ Fare doppio clic sulla tabella *Livello di istruzione*.
- ▶ Se la barra degli strumenti di formattazione non è visibile, dai menu selezionare: Visualizza > Barra degli strumenti
- ▶ Fare clic sul testo di un titolo, *Livello di istruzione*.
- ▶ Dall'elenco a discesa delle dimensioni del carattere sulla barra degli strumenti selezionare 12.
- ▶ Per cambiare il colore del testo del titolo, fare clic sullo strumento Colore testo e selezionare un nuovo colore.

Figura 7-9
Testo del titolo riformattato nella tabella pivot



		Frequenza	Percentuale	Percentuale valida	Percentuale cumulata
Validi	Scuola superiore non completata	1390	21.7	21.7	21.7
	Diploma di scuola superiore	1936	30.2	30.2	52.0
	Università non completata	1360	21.2	21.2	73.2
	Laurea	1355	21.2	21.2	94.4
	Titolo post-laurea	359	5.6	5.6	100.0
	Totale	6400	100.0	100.0	

È anche possibile modificare il contenuto di tabelle e livelli. Ad esempio, è possibile modificare il titolo di questa tabella.

- ▶ Fare doppio clic sul titolo.
- ▶ Digitare Livello di istruzione per la nuova etichetta.

Nota: Se si modificano i valori in una tabella, i totali e altre statistiche non vengono ricalcolate.

Nascondere righe e colonne

Alcuni dei dati visualizzati in una tabella potrebbero essere superflui o addirittura rendere inutilmente complessa la tabella. Fortunatamente, è possibile nascondere righe e colonne intere senza perdere dati.

- ▶ Se non è già attivata, attivare con un doppio clic la tabella *Livello di istruzione*.
- ▶ Fare clic sull'etichetta di colonna *Percentuale valida* per selezionarla.
- ▶ Dal menu Modifica o dal menu di scelta rapida scegliere:
Seleziona > Celle dati e celle etichette
- ▶ Dal menu Visualizza scegliere Nascondi oppure dal menu di scelta rapida scegliere Nascondi categoria.

La colonna viene nascosta, ma non eliminata.

Figura 7-10

Colonna Percentuale valida nascosta in una tabella

Grado di istruzione		Statistiche		
		Frequenza	Percentuale	Percentuale cumulata
Validi	Scuola superiore non completata	1390	21,7	21,7
	Diploma di scuola superiore	1936	30,3	52,0
	Frequenza all'università	1360	21,3	73,2
	Laurea universitaria	1355	21,2	94,4
	Specializzazione post-universitaria	359	5,6	100,0
	Totale	6400	100,0	

Per visualizzare nuovamente la colonna:

- ▶ Dai menu, scegliere:
Visualizza > Mostra tutto

È possibile nascondere e visualizzare le righe allo stesso modo delle colonne.

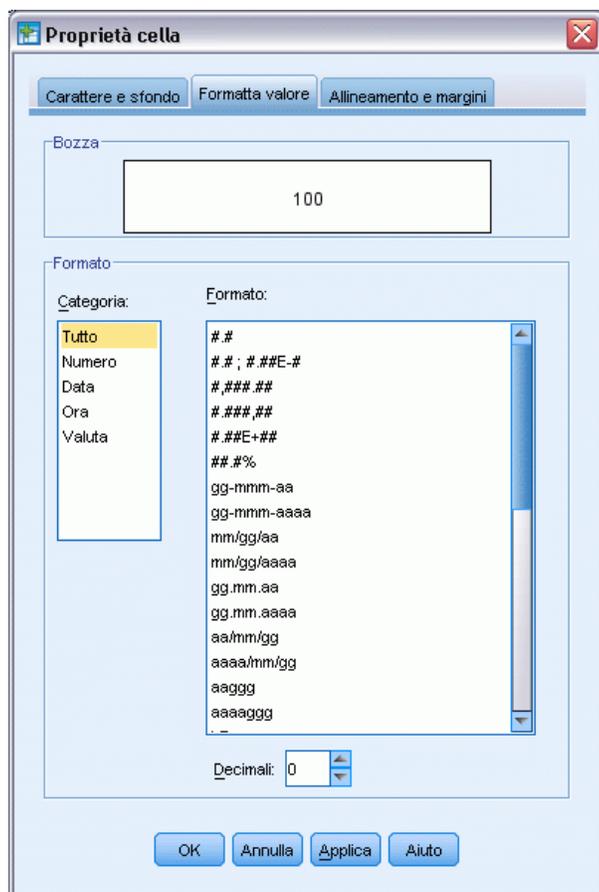
Modifica dei formati di visualizzazione dei dati

È possibile modificare con facilità il formato di visualizzazione dei dati nelle tabelle pivot.

- ▶ Se non è già attivata, attivare con un doppio clic la tabella *Livello di istruzione*.
- ▶ Fare clic sull'etichetta di colonna *Percentuale* per selezionarla.
- ▶ Dal menu Modifica o dal menu di scelta rapida scegliere:
Seleziona > Celle dati
- ▶ Dal menu Formato o dal menu di scelta rapida selezionare Proprietà cella.
- ▶ Fare clic sulla scheda Formatta valore.
- ▶ Digitare 0 nel campo Decimali per nascondere tutti i punti decimali nella colonna.

Figura 7-11

Proprietà cella: scheda Formatta Valore



È anche possibile modificare il tipo di dati e il formato in questa finestra di dialogo.

- ▶ Selezionare il tipo desiderato dall'elenco Categoria, quindi dall'elenco Formato selezionare il relativo formato.
- ▶ Fare clic su OK o su Applica per applicare le modifiche.

Figura 7-12
Decimali nascosti nella colonna Percentuale

		Frequenza	Percentuale	Percentuale cumulata
Validi	Scuola superiore non completata	1390	22	21,7
	Diploma di scuola superiore	1936	30	52,0
	Frequenza all'università	1360	21	73,2
	Laurea universitaria	1355	21	94,4
	Specializzazione post-universitaria	359	6	100,0
	Totale	6400	100	

Nella colonna *Percentuale* i decimali sono nascosti.

Modelli di tabelle

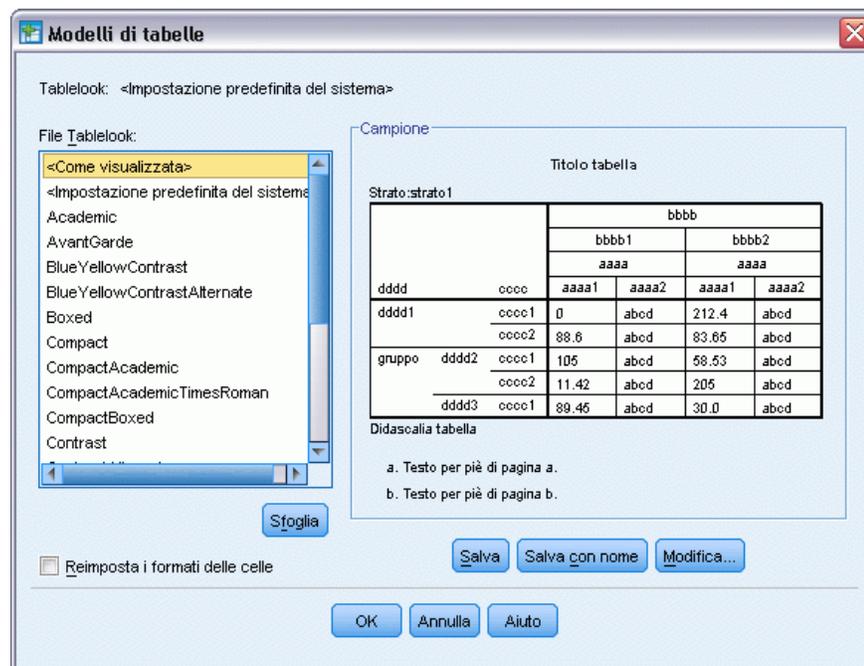
Il formato delle tabelle riveste particolare importanza ai fini di una produzione di risultati significativi, concisi e chiari. Se la lettura di una tabella è complicata, le informazioni contenute all'interno di essa potrebbero non essere facilmente comprensibili.

Utilizzo di formati predefiniti

- ▶ Fare doppio clic sulla tabella *Stato civile*.
- ▶ Dai menu, scegliere:
Formato > Tablelook...

Nella finestra di dialogo TableLook sono elencati vari stili predefiniti. Selezionare uno stile dall'elenco per visualizzarlo in anteprima nella finestra di esempio a destra.

Figura 7-13
Finestra di dialogo TableLook



È possibile utilizzare uno stile così com'è oppure modificarlo per adattarlo alle esigenze dell'utente.

- Per utilizzare uno stile esistente, selezionarlo e fare clic su OK.

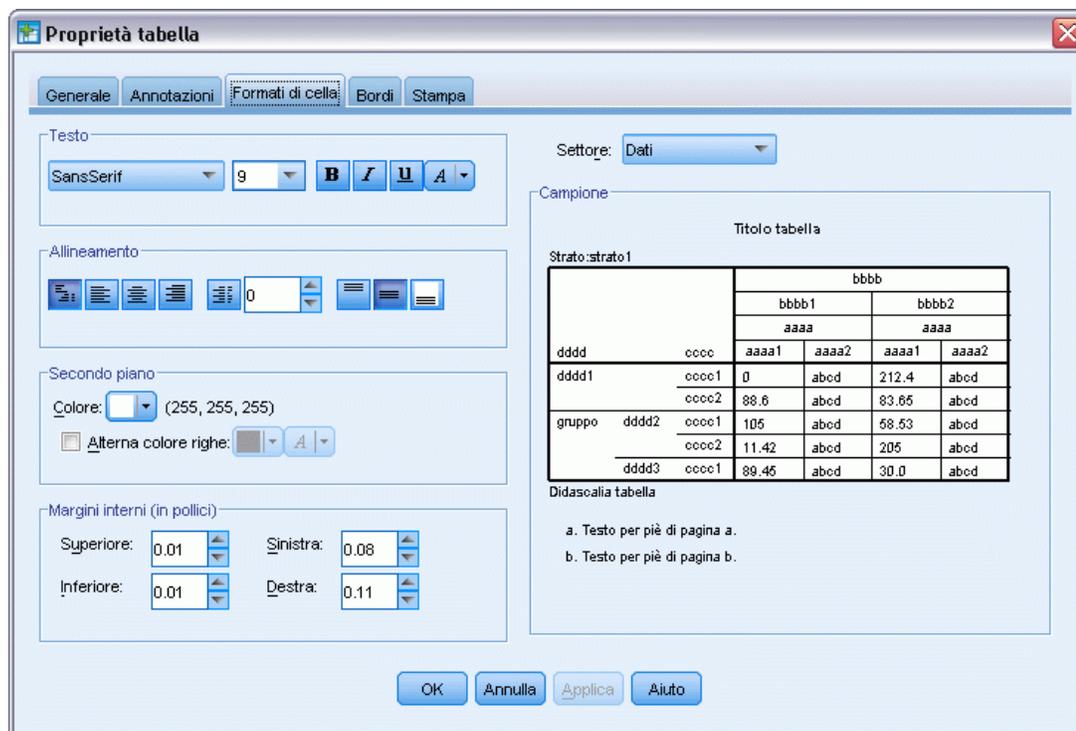
Personalizzazione di stili TableLook

È possibile modificare un formato per adattarlo a esigenze specifiche. Quasi tutti gli aspetti di una tabella, dal colore dello sfondo agli stili dei bordi, possono essere personalizzati.

- Fare doppio clic sulla tabella *Stato civile*.
- Dai menu, scegliere:
Formato > Tablelook...
- Selezionare lo stile più vicino al formato desiderato e fare clic su Modifica.

- Fare clic sulla scheda Formati di cella per visualizzare le opzioni di formattazione.

Figura 7-14
Finestra di dialogo Proprietà tabella



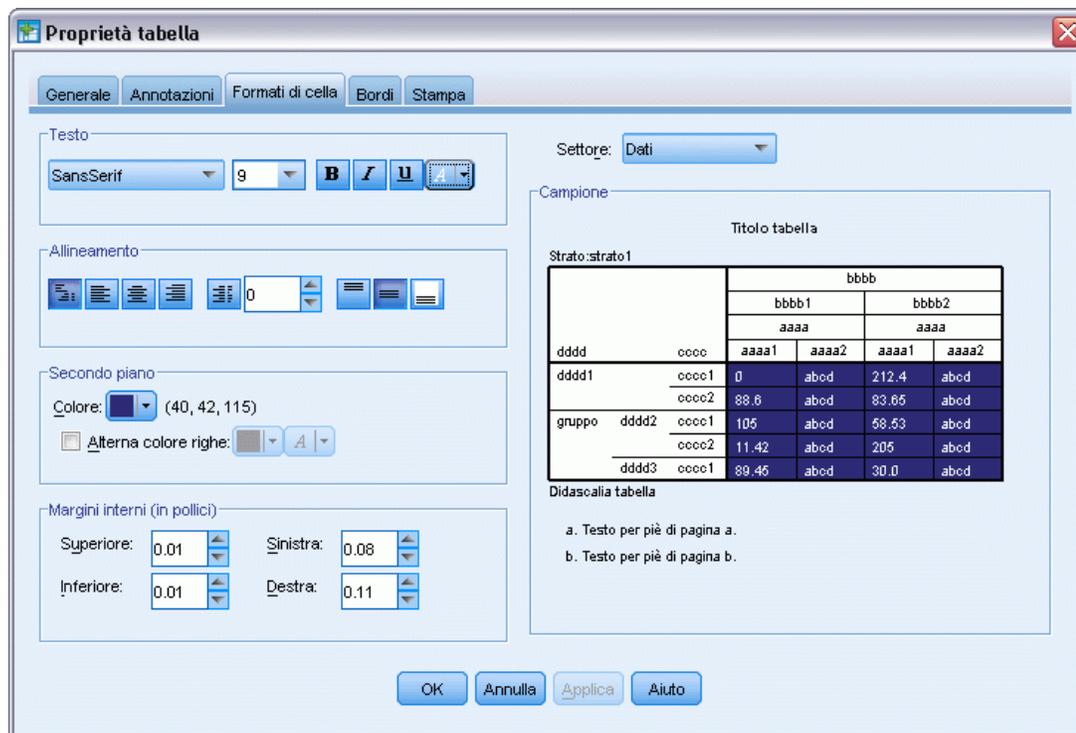
Le opzioni di formattazione riguardano il nome e le dimensioni del carattere, lo stile e il colore. Sono disponibili anche ulteriori opzioni per l'allineamento, i colori di primo piano e di sfondo e le dimensioni dei margini.

La finestra di esempio mostra a destra un'anteprima degli effetti della formattazione sulla tabella. Ogni area della tabella può avere stili di formattazione diversi. In tal modo, titolo e dati possono avere stili diversi. Per selezionare l'area della tabella che si desidera modificare, è possibile scegliere il nome dell'area nell'elenco a discesa Area oppure fare clic sull'area desiderata nella finestra di esempio.

- Selezionare Dati dall'elenco a discesa Area.
- Selezionare un nuovo colore dalla tavolozza Sfondo.
- Quindi, selezionare un nuovo colore del testo.

Il nuovo stile è visualizzato nella finestra di esempio.

Figura 7-15
Modifica dei formati di cella



- Fare clic su OK per chiudere la finestra di dialogo TableLook.

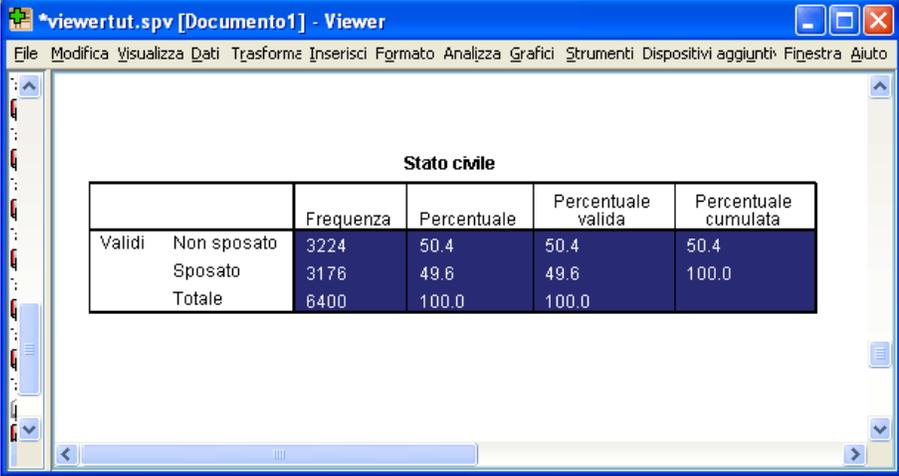
È possibile salvare il nuovo stile per applicarlo alle successive tabelle.

- Fare clic su Salva come.
- Andare alla directory di destinazione desiderata e digitare un nome per il nuovo stile nella casella di testo Nome file.
- Scegliere Salva.
- Fare clic su OK per applicare le modifiche e ritornare al Viewer.

Adesso la tabella contiene la formattazione personalizzata specificata.

Figura 7-16

Personalizzare il TableLook



The screenshot shows a window titled '*viewertut.spv [Documento1] - Viewer'. The menu bar includes File, Modifica, Visualizza, Dati, Trasforma, Inserisci, Formato, Analizza, Grafici, Strumenti, Dispositivi aggiunti, Finestra, and Aiuto. The main content area displays a PivotTable titled 'Stato civile'. The table has four columns: 'Validi', 'Frequenza', 'Percentuale', 'Percentuale valida', and 'Percentuale cumulata'. The rows represent 'Non sposato', 'Sposato', and 'Totale'.

Validi	Frequenza	Percentuale	Percentuale valida	Percentuale cumulata
Non sposato	3224	50.4	50.4	50.4
Sposato	3176	49.6	49.6	100.0
Totale	6400	100.0	100.0	

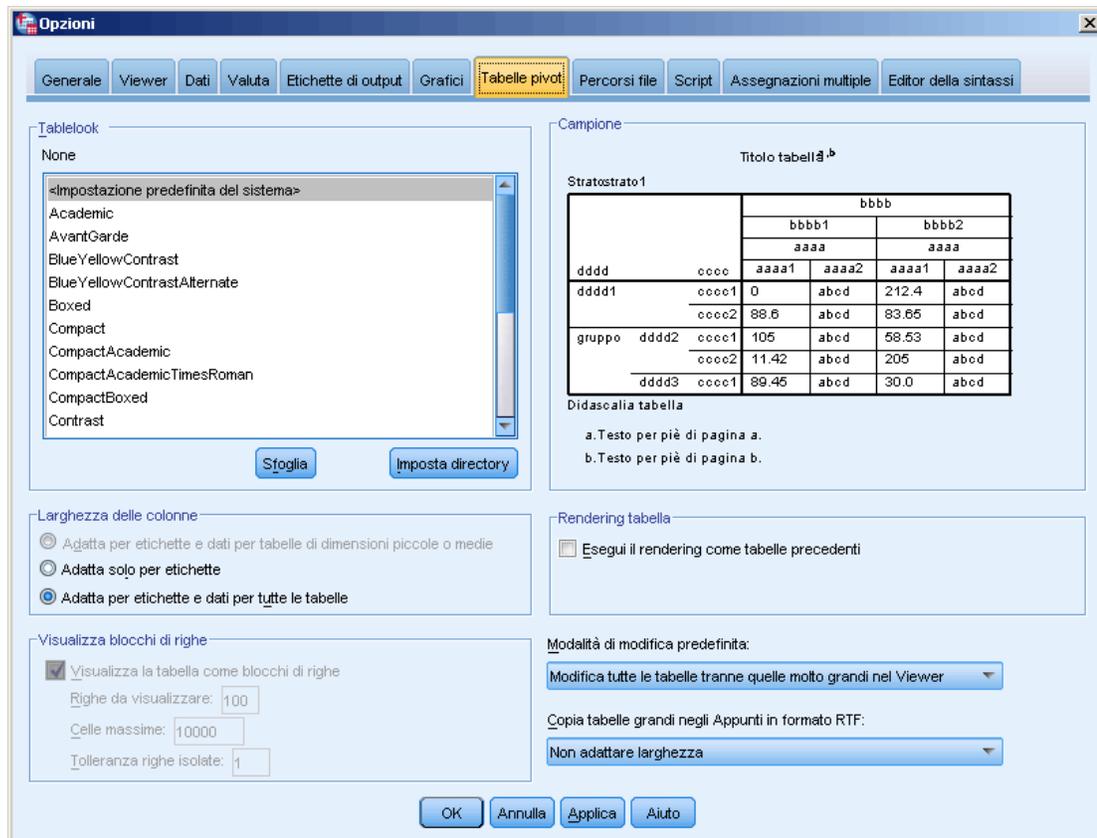
Modifica dei formati di tabella predefiniti

Sebbene sia possibile modificare il formato delle tabelle già create, potrebbe essere consigliabile modificare il TableLook predefinito, in modo da non dover intervenire sul formato ogni volta che si crea una tabella.

Per modificare lo stile TableLook predefinito per le tabelle pivot, scegliere:
Modifica > Opzioni...

- Fare clic sulla scheda Tabelle pivot nella finestra di dialogo Opzioni.

Figura 7-17
Finestra di dialogo Opzioni



- Selezionare lo stile TableLook che si desidera utilizzare per tutte le nuove tabelle.

A destra della finestra di esempio è visualizzata un'anteprima di ogni TableLook.

- Fare clic su OK per salvare le impostazioni e chiudere la finestra di dialogo.

Tutte le tabelle create dopo aver modificato il TableLook predefinito saranno conformi, automaticamente, alle nuove regole di formattazione.

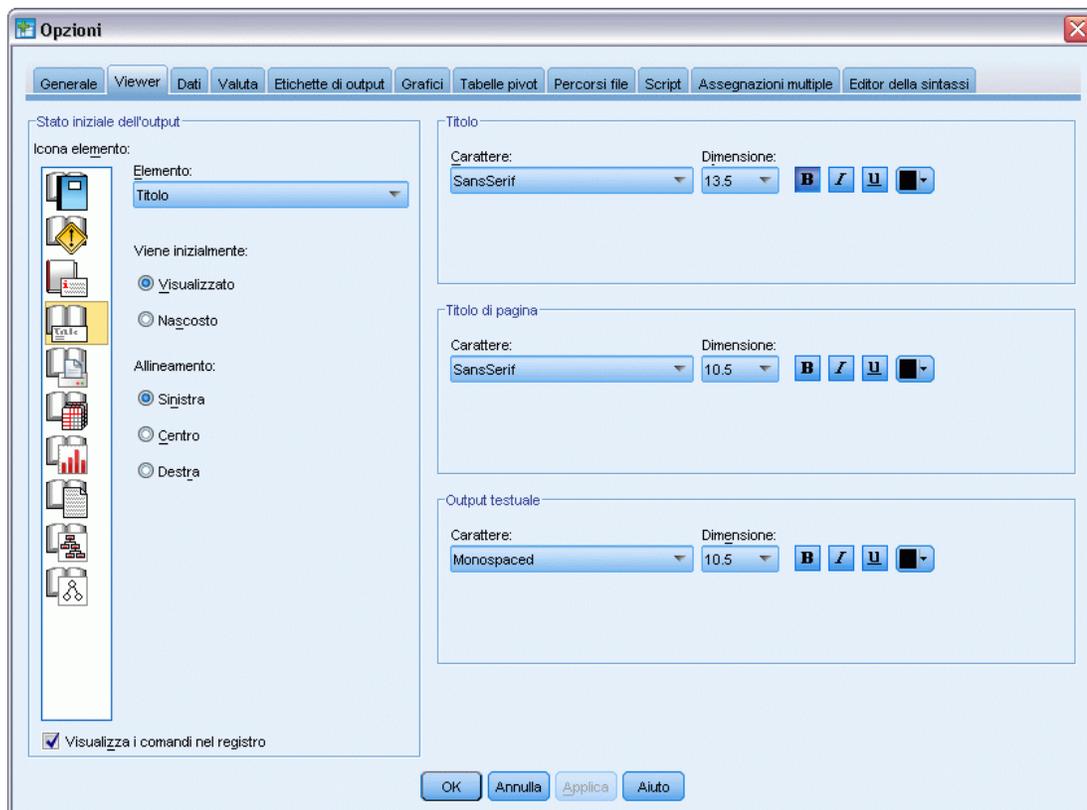
Personalizzazione delle impostazioni iniziali di visualizzazione

Le impostazioni iniziali di visualizzazione includono l'allineamento degli oggetti nel Viewer, che siano visibili o non visibili per impostazione predefinita, e l'ampiezza della finestra del Viewer. Per modificare tali impostazioni:

- Dai menu, scegliere:
Modifica > Opzioni...

- Fare clic sulla scheda Viewer.

Figura 7-18
opzioni Viewer



Le impostazioni vengono applicate per ogni singolo oggetto. Ad esempio, è possibile personalizzare la modalità di visualizzazione dei grafici senza apportare modifiche alla modalità di visualizzazione delle tabelle. È sufficiente selezionare l'oggetto che si desidera personalizzare e apportare le modifiche appropriate.

- Fare clic sull'icona Titolo per visualizzarne le impostazioni.
- Fare clic su Centra per visualizzare tutti i titoli al centro (orizzontale) del Viewer.

È anche possibile nascondere elementi quali i messaggi di registro e degli avvisi che non contengono informazioni realmente utili. Facendo doppio clic su un'icona si modifica automaticamente la proprietà di visualizzazione dell'oggetto.

- Fare doppio clic sull'icona Avvisi per nascondere i messaggi degli avvisi nell'output.
- Fare clic su OK per salvare le modifiche e chiudere la finestra di dialogo.

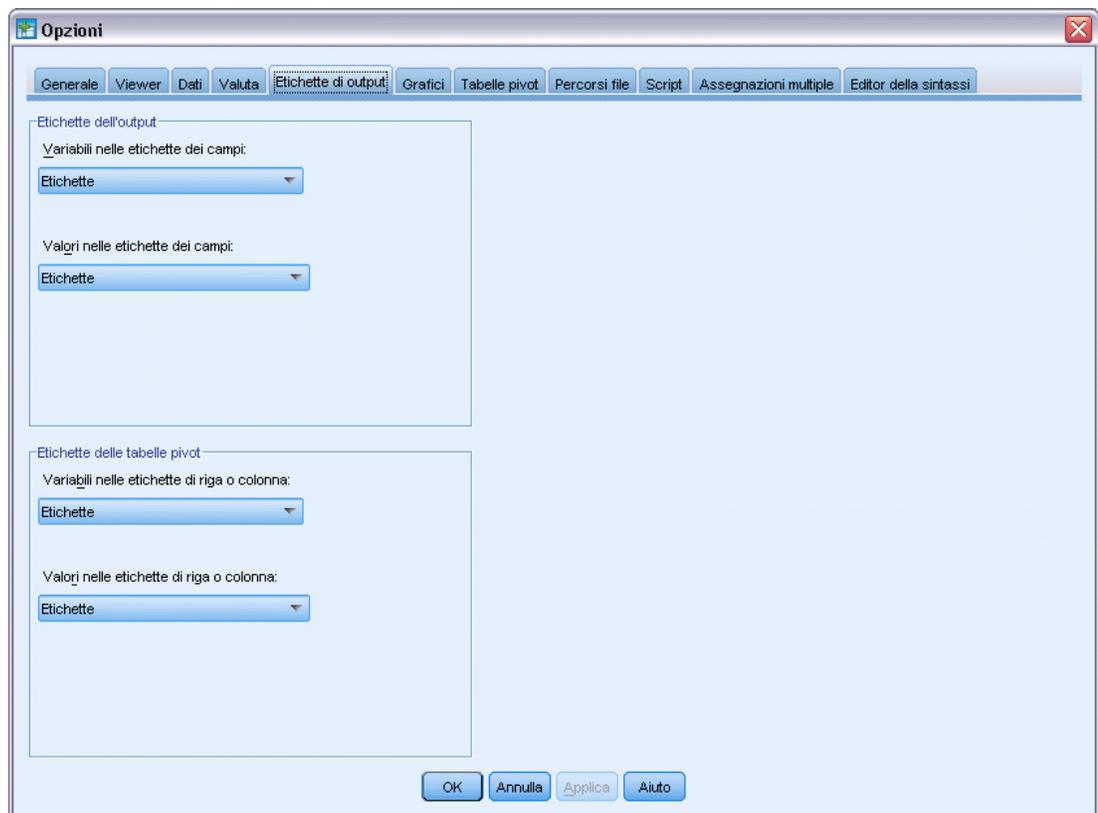
Visualizzazione delle etichette di valori e variabili

Nella maggior parte dei casi, la visualizzazione delle etichette di variabili e valori è più utile della visualizzazione del nome della variabile o del valore dei dati effettivi. In alcuni casi, tuttavia, può essere utile visualizzare sia i nomi che le etichette.

- ▶ Dai menu, scegliere:
Modifica > Opzioni...
- ▶ Selezionare la scheda Etichette di output.

Figura 7-19

Opzioni: Etichette di output



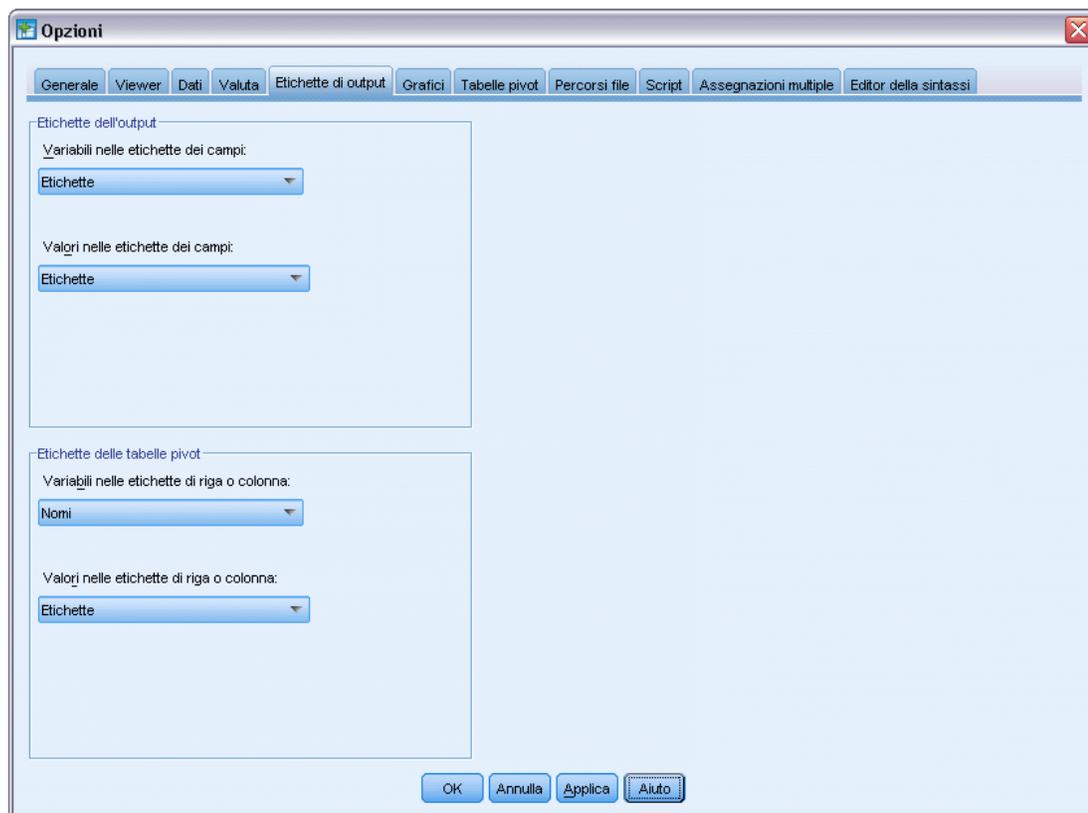
È possibile specificare impostazioni diverse per la vista riassuntiva e l'output. Per visualizzare, ad esempio, le etichette nella vista riassuntiva e i nomi delle variabili e i valori dei dati nell'output:

- ▶ Nel gruppo Etichette delle tabelle pivot selezionare Nomi dall'elenco a discesa Variabili nelle etichette di riga o colonna per visualizzare i nomi delle variabili anziché le etichette.

- Quindi selezionare Valori dall'elenco a discesa Valori nelle etichette di riga o colonna per visualizzare i valori dei dati anziché le etichette.

Figura 7-20

Impostazioni di Etichette delle tabelle pivot



Le tabelle seguenti generate nella sessione rispecchieranno tali modifiche.

Figura 7-21
Nomi e valori delle variabili visualizzati

The screenshot shows the SPSS Viewer window titled 'viewertut.spo - Viewer'. It displays two tables under the heading 'Statistiche'.

The first table, titled 'Stato civile', shows the number of valid and missing cases:

N	valido	6400
	mancanza	0

The second table, titled 'Stato civile', is a frequency table with the following data:

	Frequenza	Percentuale	Percentuale valida	Percentuale cumulata
Validi Maschio	1052	47,5	47,8	34,7
Femmina	1157	52,4	52,5	100,0
Totale	2209	100,0	100,0	

Utilizzo dell'output in altre applicazioni

I risultati possono essere utilizzati in molte applicazioni. Ad esempio, è possibile includere una tabella o un grafico in una presentazione o in un report.

Gli esempi seguenti sono specifici di Microsoft Word ma funzionano anche in altre applicazioni di elaborazione testi.

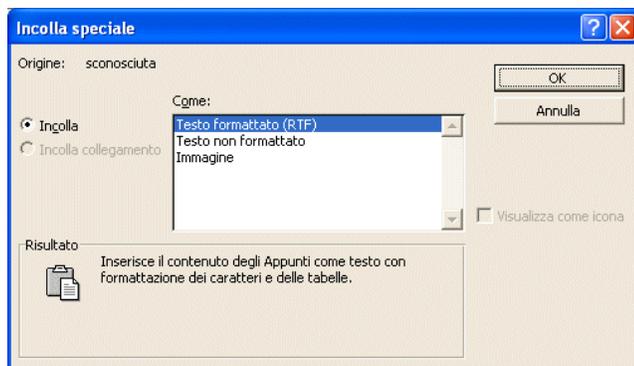
Incollare i risultati come tabelle Word

È possibile incollare tabelle pivot in Word come se fossero tabelle Word native. Tutti gli attributi della tabella, quali le dimensioni e il colore dei caratteri vengono preservati. Dato che la tabella viene incollata nel formato tabelle di Word, è possibile modificarla in Word proprio come qualsiasi altra tabella.

- ▶ Fare clic sulla tabella *Stato civile* nel Viewer.
- ▶ Dai menu, scegliere:
Modifica > Copia
- ▶ Aprire l'applicazione di elaborazione testi.
- ▶ Dai menu dell'applicazione scegliere:
Modifica > Incolla speciale...

- Selezionare Testo formattato (RTF) nella finestra di dialogo Incolla speciale.

Figura 7-22
Finestra di dialogo Incolla speciale



- Fare clic su OK per incollare i risultati nel documento corrente.

Figura 7-23
Tabella pivot visualizzata in Word

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulata
Validi	Celibe	3324	50,4	50,4	50,4
	Sposato	3176	49,6	49,6	100,0
	Totale	6400	100,0	100,0	

La tabella viene ora visualizzata nel documento. È possibile applicare la formattazione personalizzata, modificare i dati e ridimensionare la tabella per adeguarla alle necessità dell'utente.

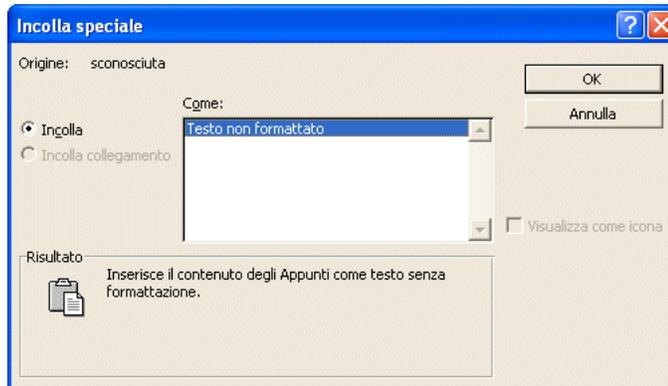
Incollare i risultati come testo

È possibile copiare le tabelle pivot ad altre applicazioni come testo normale. Gli stili di formattazione non vengono preservati con questo metodo, ma è possibile modificare i dati della tabella dopo averli incollati nell'applicazione di destinazione.

- Fare clic sulla tabella *Stato civile* nel Viewer.
- Dai menu, scegliere:
Modifica > Copia
- Aprire l'applicazione di elaborazione testi.

- ▶ Dai menu dell'applicazione scegliere:
Modifica > Incolla speciale...
- ▶ Selezionare Testo non formattato nella finestra di dialogo Incolla speciale.

Figura 7-24
Finestra di dialogo Incolla speciale



- ▶ Fare clic su OK per incollare i risultati nel documento corrente.

Figura 7-25
Tabella pivot visualizzata in Word

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulata
Validi	Celibe	3324	50,4	50,4	50,4
	Sposato	3176	49,6	49,6	100,0
	Totale	6400	100,0	100,0	

Ogni colonna della tabella viene separata da tabulazioni. È possibile modificare la larghezza delle colonne regolando le tabulazioni nell'applicazione di elaborazione testi.

Esportazione dei risultati in file di Microsoft Word, PowerPoint ed Excel

È possibile esportare i risultati a un file Excel, Microsoft Word o PowerPoint. È possibile esportare elementi selezionati o tutti gli elementi nel Viewer. In questa sezione vengono utilizzati i file *msouttut.spv* e *demo.sav*. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice A a pag. 152.](#)

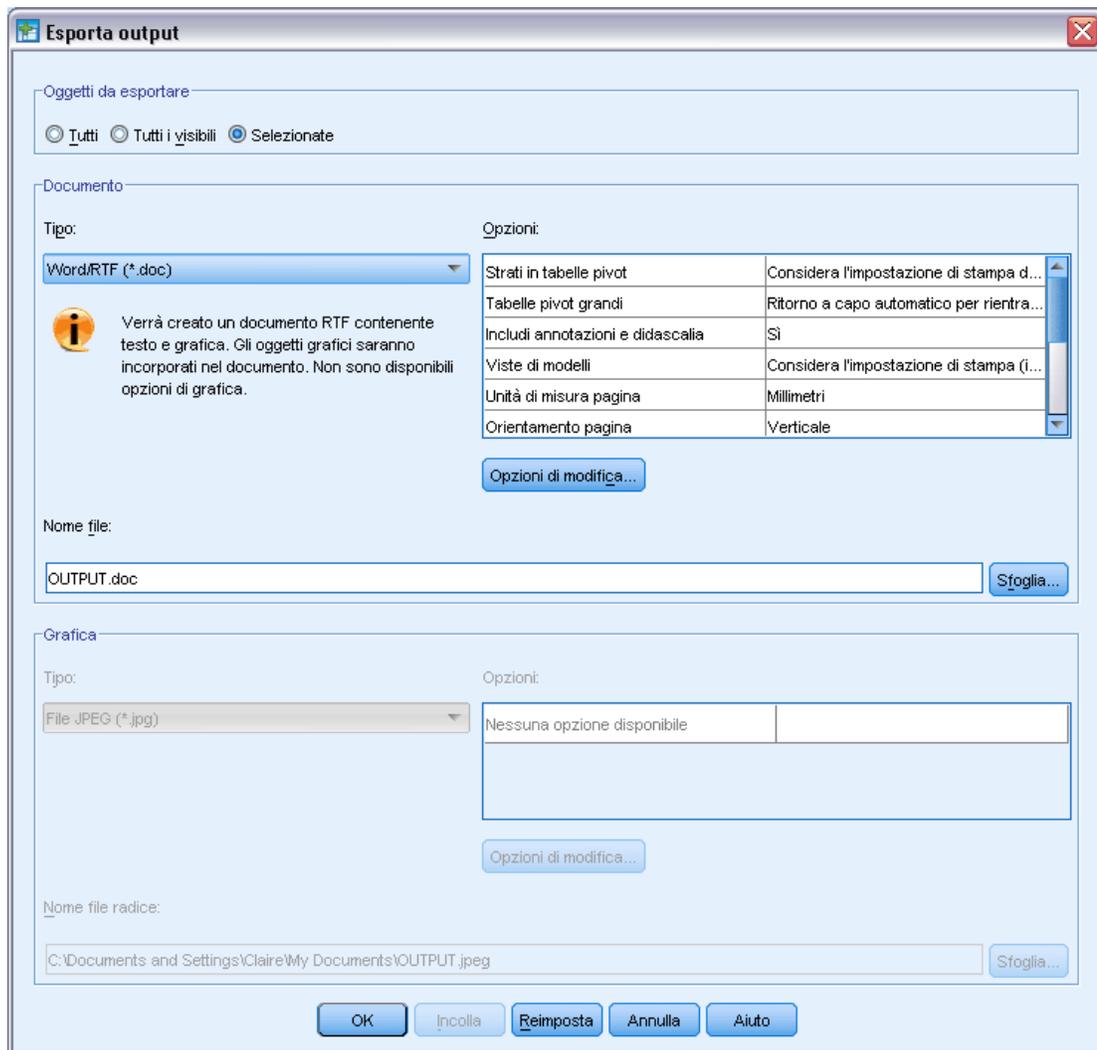
Nota: l'esportazione in PowerPoint è disponibile solo nei sistemi operativi Windows e non nella versione per studenti.

Nella vista riassuntiva del Viewer è possibile specificare gli elementi specifici che si desidera esportare oppure scegliere di esportare tutti gli elementi o tutti gli elementi visibili.

- ▶ Dai menu del Viewer scegliere:
File > Esporta...

Anziché esportare tutti gli oggetti nel Viewer, è possibile scegliere di esportare solo gli oggetti visibili (libri aperti nel riquadro della vista riassuntiva) oppure gli oggetti selezionati nella vista riassuntiva. Se non sono stati selezionati oggetti nel riquadro della vista riassuntiva, l'opzione che consente di esportare gli oggetti selezionati non sarà disponibile.

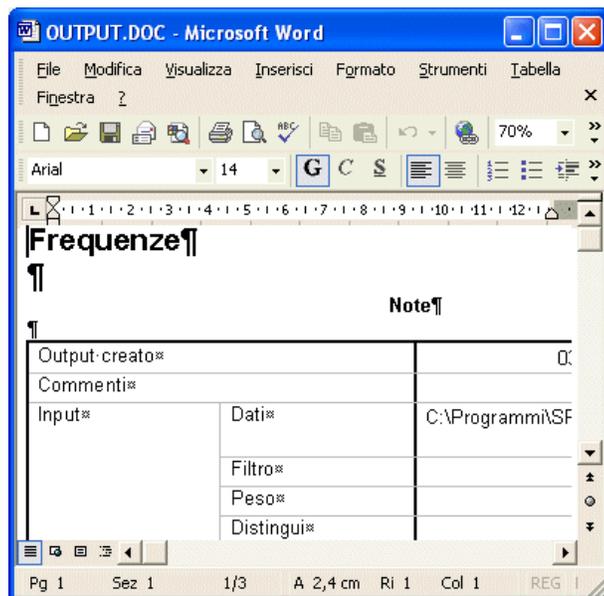
Figura 7-26
Finestra di dialogo *Esporta output*



- ▶ Nel gruppo Oggetti da esportare, selezionare Tutto.
- ▶ Dall'elenco a discesa Tipo file, selezionare File Word/RTF (*.doc).
- ▶ Fare clic su OK per generare il file di Word.

Quando viene aperto il file risultante con Word, è possibile vedere come sono stati esportati i risultati. Le note, che sono oggetti non visibili, vengono visualizzate in Word, perché è stato scelto di esportare tutti gli oggetti.

Figura 7-27
Output.doc in Word



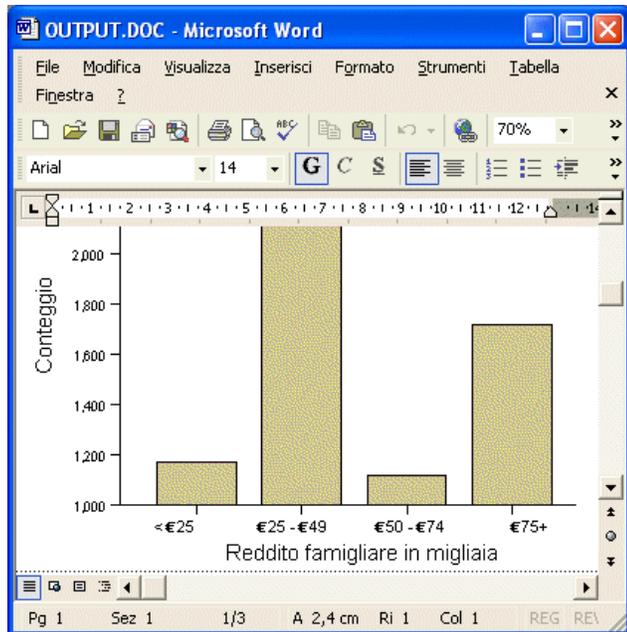
Le tabelle pivot diventano tabelle di Word e preservano la formattazione della tabella pivot originale, compresi i caratteri, i colori, i bordi e così via.

Figura 7-28
Tabelle pivot in Word

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulata
Validi	Maschio	1052	47,5	47,8	34,7
	Femmina	1157	52,4	52,5	100,0
	Totale	2208	100,0	100,0	

I grafici vengono incorporati nel documento di Word come immagini grafiche.

Figura 7-29
Grafici in Word



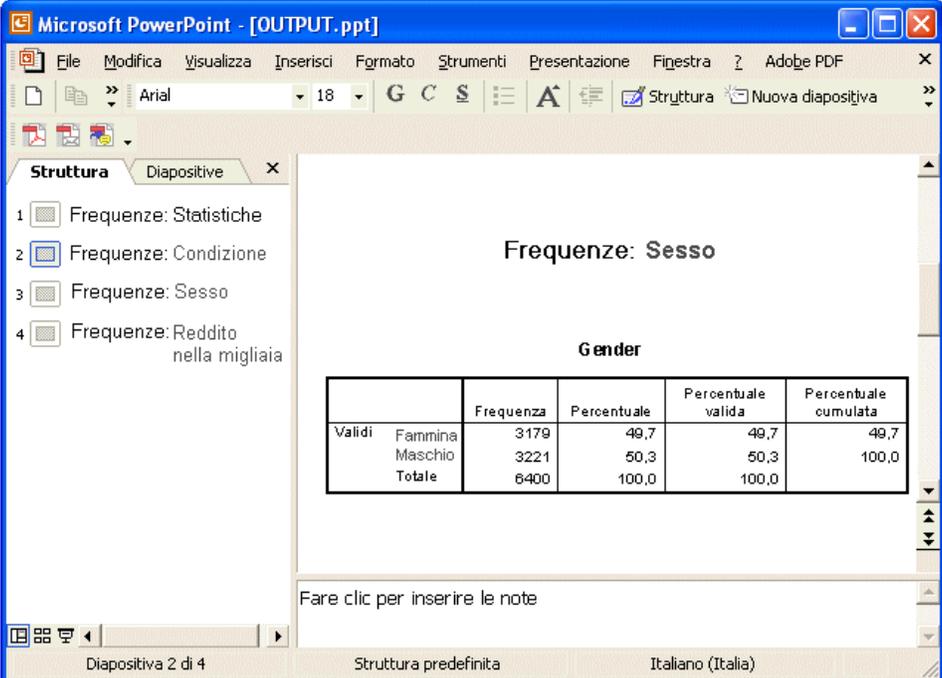
L'output testuale viene visualizzato con lo stesso carattere utilizzato per l'oggetto di testo nel Viewer. Per il corretto allineamento, per l'output testuale è necessario utilizzare un carattere a spaziatura fissa (proporzionale).

Figura 7-30
Output testuale in Word

Reddito familiare in migliaia	Età in anni Media	Livello di istruzione Media
< € 25	38	2
€ 25 - € 49	39	3
€ 50 - € 74	43	3
€ 75+	49	3
Totale	42	3

Se si esegue l'esportazione in un file di PowerPoint, ogni elemento esportato viene inserito in una diapositiva separata. Le tabelle pivot esportate in PowerPoint diventano tabelle di Word e mantengono la formattazione della tabella pivot originale, compresi i caratteri, i colori, i bordi e così via.

Figura 7-31
Tabelle pivot in PowerPoint



The screenshot shows a Microsoft PowerPoint window titled "Microsoft PowerPoint - [OUTPUT.ppt]". The slide is titled "Frequenze: Sesso" and contains a pivot table titled "Gender". The pivot table has the following data:

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulata
Validi	Fammina	3179	49,7	49,7	49,7
	Maschio	3221	50,3	50,3	100,0
	Totale	6400	100,0	100,0	

The slide also includes a navigation pane on the left with the following structure:

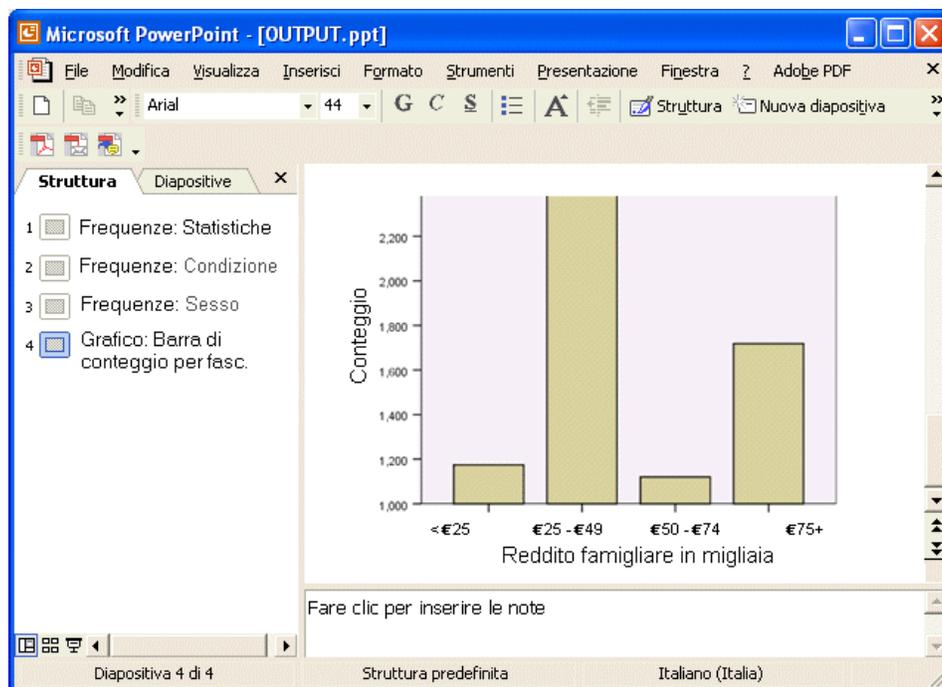
- 1 Frequenze: Statistiche
- 2 Frequenze: Condizione
- 3 Frequenze: Sesso
- 4 Frequenze: Reddito nella migliaia

The status bar at the bottom indicates "Diapositiva 2 di 4", "Struttura predefinita", and "Italiano (Italia)".

I grafici selezionati per l'esportazione in PowerPoint vengono incorporati nel file di PowerPoint.

Figura 7-32

Grafici in PowerPoint



Se l'esportazione è verso un file Excel, i risultati verranno esportati in modo diverso.

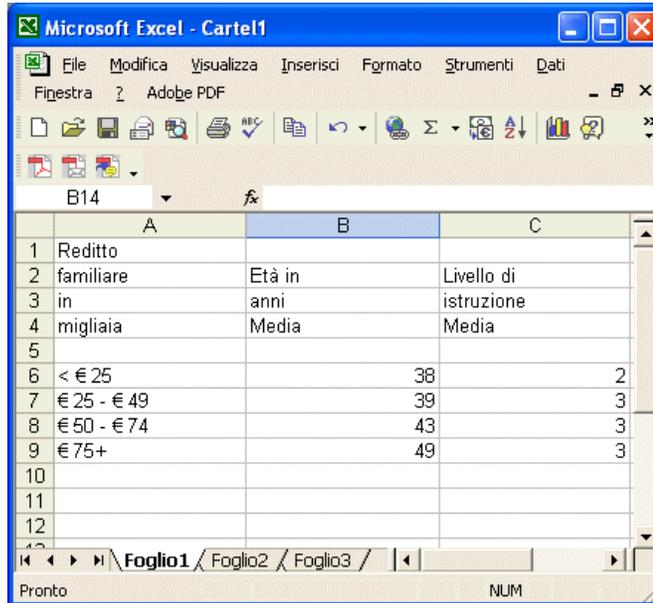
Figura 7-33
Output.xls in Excel

Sesso					
		Frequenza	Percentuale	Percentuale valida	Percentuale cumulata
Validi	f	3.179	49,7	49,7	49,7
	m	3.221	50,3	50,3	100,0
	Totale	6.400	100,0	100,0	
Categoria di reddito in migliaia					
		Frequenza	Percentuale	Percentuale valida	Percentuale cumulata
Validi	1,00	1.174	18,3	18,3	18,3
	2,00	2.388	37,3	37,3	55,7
	3,00	1.120	17,5	17,5	73,2
	4,00	1.718	26,8	26,8	100,0
	Totale	6.400	100,0	100,0	

Celle, colonne e righe della tabella pivot diventano celle, colonne e righe di Excel.

Ogni riga nell'output testuale è una riga del file di Excel, con tutti i contenuti di una riga raccolti in una singola cella.

Figura 7-34
Output testuale in Excel



The screenshot shows a Microsoft Excel window titled 'Microsoft Excel - Cartel1'. The spreadsheet contains the following data:

	A	B	C
1	Reddito		
2	familiare	Età in	Livello di
3	in	anni	istruzione
4	migliaia	Media	Media
5			
6	< € 25	38	2
7	€ 25 - € 49	39	3
8	€ 50 - € 74	43	3
9	€ 75+	49	3
10			
11			
12			

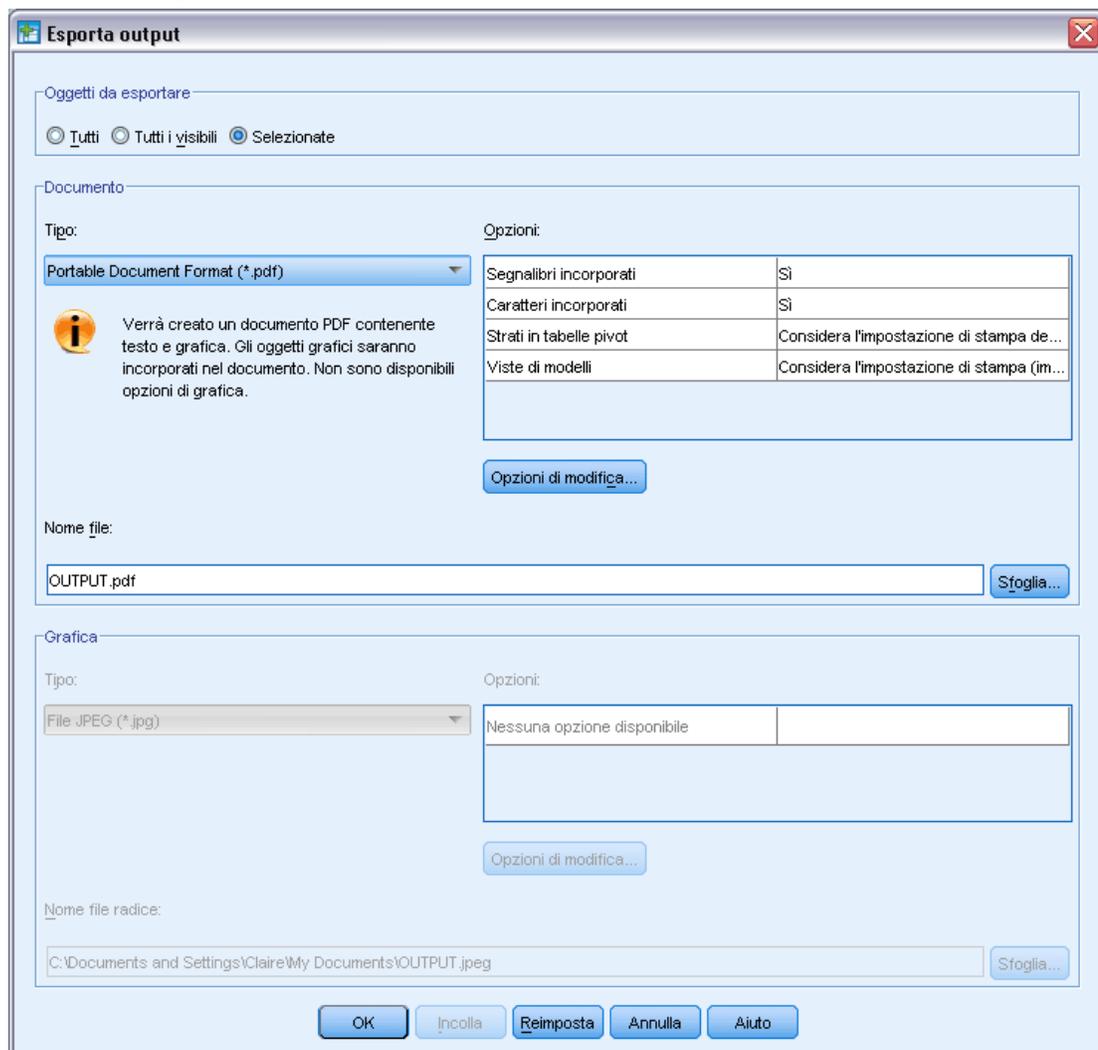
Esportazione dei risultati in formato PDF

È possibile esportare tutti gli elementi o solo gli elementi selezionati nel Viewer in un file PDF (portable document format).

- Dai menu della finestra del Viewer contenente il risultato che si desidera esportare in PDF scegliere:
File > Esporta...

- Nella finestra di dialogo *Esporta output*, nell'elenco a discesa *Tipo file* in *Esporta formato*, scegliere *Portable Document Format*.

Figura 7-35
Finestra di dialogo *Esporta output*

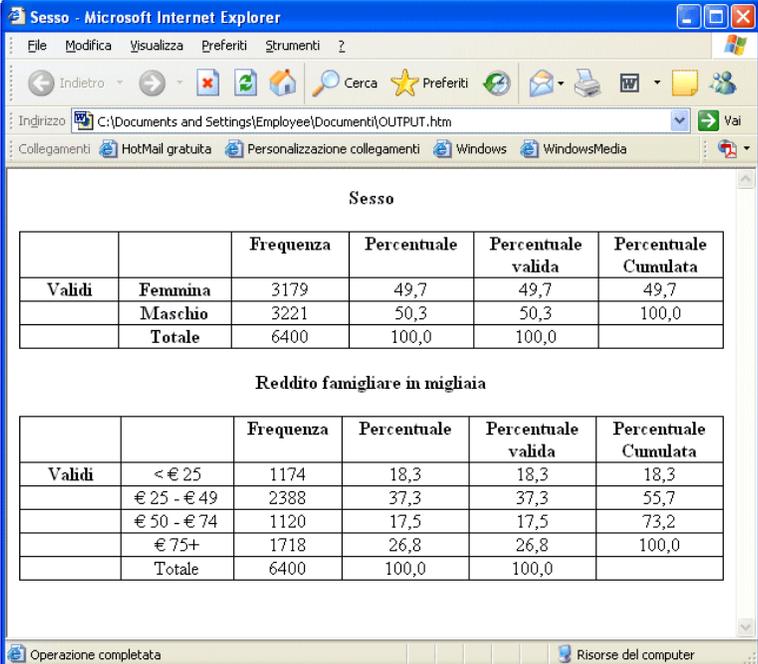


- La vista riassuntiva del documento del Viewer viene convertita in segnalibri nel file PDF per semplicità di navigazione.
- Le dimensioni della pagina, l'orientamento, i margini, il contenuto e la visualizzazione di intestazioni e piè di pagina nonché le dimensioni del grafico stampato nei documenti PDF sono opzioni controllate dell'impostazione della pagina (menu File, Imposta pagina nella finestra del Viewer).
- La risoluzione (DPI) del documento PDF corrisponde alla risoluzione corrente della stampante predefinita o correntemente selezionata (modificabile in Imposta pagina). La risoluzione massima è 1200 DPI. Se l'impostazione della stampante è maggiore, la risoluzione

Esportazione dei risultati in formato HTML

Si possono esportare i risultati anche in formato HTML (hypertext markup language). Durante le operazioni di salvataggio come HTML, l'intero output non grafico viene esportato in un file HTML unico.

Figura 7-37
Output.htm nel browser Web



The screenshot shows a Microsoft Internet Explorer window with the address bar pointing to 'C:\Documents and Settings\Employee\Documents\OUTPUT.htm'. The page content consists of two tables. The first table, titled 'Sesso', shows the frequency and percentage distribution of respondents by gender. The second table, titled 'Reddito familiare in migliaia', shows the frequency and percentage distribution of respondents by family income bracket.

Sesso					
		Frequenza	Percentuale	Percentuale valida	Percentuale Cumulata
Validi	Femmina	3179	49,7	49,7	49,7
	Maschio	3221	50,3	50,3	100,0
	Totale	6400	100,0	100,0	

Reddito familiare in migliaia					
		Frequenza	Percentuale	Percentuale valida	Percentuale Cumulata
Validi	< € 25	1174	18,3	18,3	18,3
	€ 25 - € 49	2388	37,3	37,3	55,7
	€ 50 - € 74	1120	17,5	17,5	73,2
	€ 75+	1718	26,8	26,8	100,0
	Totale	6400	100,0	100,0	

Quando si esegue l'esportazione in HTML, è possibile esportare anche i grafici ma non in un unico file.

Ogni grafico verrà salvato come un file del formato specificato e nel documento HTML verranno inseriti riferimenti ai file grafici indicati. Esiste inoltre la possibilità di esportare tutti i grafici, o i grafici selezionati, in file grafici distinti.

Utilizzo della sintassi

Il potente linguaggio a comandi consente di salvare e automatizzare molte delle operazioni più comuni. Inoltre, offre alcune funzionalità a cui non è possibile accedere tramite i menu e le finestre di dialogo. È possibile accedere alla maggior parte dei comandi tramite i menu e le finestre di dialogo. Tuttavia, alcuni comandi e opzioni sono disponibili solo utilizzando il linguaggio dei comandi. Questo linguaggio consente inoltre di salvare le sessioni in un file di sintassi, in modo da poter ripetere l'analisi successivamente.

Un file di sintassi dei comandi è un semplice file di testo che contiene comandi di sintassi IBM® SPSS® Statistics. È possibile aprire una finestra di sintassi e digitarvi direttamente i comandi, ma spesso è più semplice affidare parte o tutto il lavoro alle finestre di dialogo.

Gli esempi di questo capitolo utilizzano il file di dati *demo.sav*. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice A a pag. 152.](#)

Nota: la sintassi dei comandi non è disponibile nella versione per studenti.

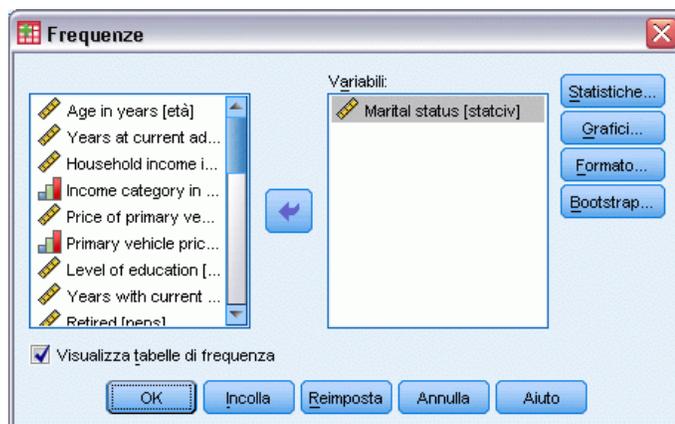
Incollare la sintassi

Il modo più semplice per creare la sintassi è utilizzare il pulsante Incolla situato nella maggior parte delle finestre di dialogo.

- ▶ Aprire il file di dati *demo.sav*. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice A a pag. 152.](#)
- ▶ Dai menu, scegliere:
Analizza > Statistiche descrittive > Frequenze...

Si apre la finestra di dialogo Frequenze.

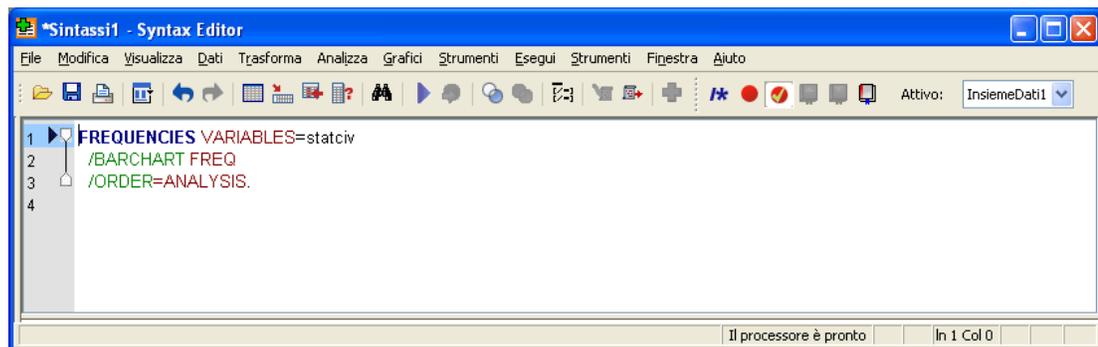
Figura 8-1
Finestra di dialogo Frequenze



- ▶ Selezionare *Stato civile [statciv]* e spostare la variabile nell'elenco Variabili.

- ▶ Fare clic su Grafici.
- ▶ Nella finestra di dialogo Grafici, selezionare Grafici a barre.
- ▶ Nel gruppo Valori nel grafico, selezionare Percentuali.
- ▶ Fare clic su Continua.
- ▶ Fare clic su Incolla per copiare la sintassi creata come risultato delle selezioni della finestra di dialogo nell'Editor di sintassi.

Figura 8-2
Sintassi di frequenze

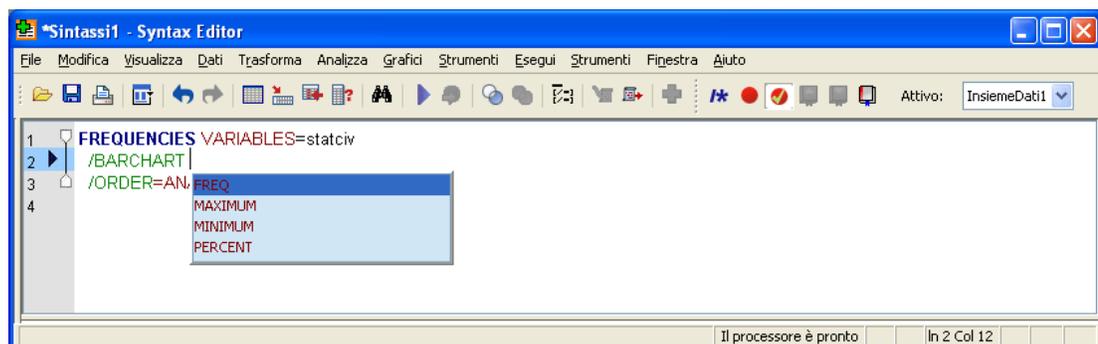


- ▶ Per eseguire la sintassi correntemente visualizzata, scegliere:
Esegui > Selezione

Modificare la sintassi

Nella finestra di sintassi è possibile modificare la sintassi. Ad esempio, è possibile modificare il sottocomando `/BARCHART` in modo da visualizzare le frequenze anziché le percentuali. (I sottocomandi sono indicati da una barra). Se si conosce la parola chiave per la visualizzazione delle frequenze, è possibile immetterla direttamente. In caso contrario, è possibile ottenere un elenco delle parole chiave disponibili per il sottocomando posizionando il cursore in un punto qualsiasi dopo il nome del sottocomando e premendo `Ctrl+Barra` spaziatrice. In questo modo si visualizza il controllo di autocompletamento per il sottocomando.

Figura 8-3
Controllo di autocompletamento che visualizza le parole chiave



- Fare clic sulla voce etichettata **FREQ** per le frequenze. Facendo clic su una voce nel controllo di autocompletamento si inserirà quella voce nella posizione del cursore (la parola chiave **PERCENT** originale è stata eliminata manualmente).

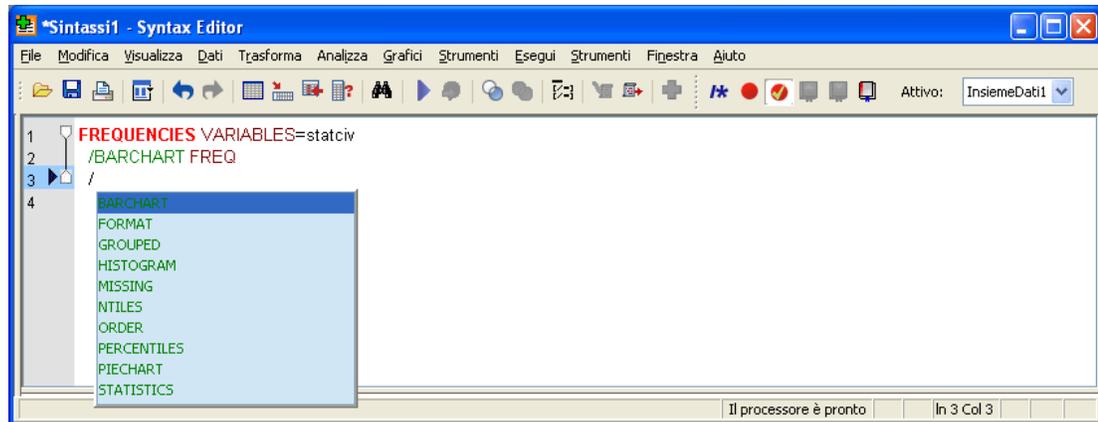
Per impostazione predefinita, il controllo di autocompletamento proporrà un elenco di termini disponibili durante la digitazione. Ad esempio, si potrebbe decidere di includere un grafico a torta insieme a un grafico a barre. Il grafico a torta viene specificato con un sottocomando distinto.

- Premere Invio dopo la parola chiave **FREQ** e digitare una barra (/) per indicare l'inizio di un sottocomando.

L'Editor della sintassi suggerisce l'elenco dei sottocomandi per il comando corrente.

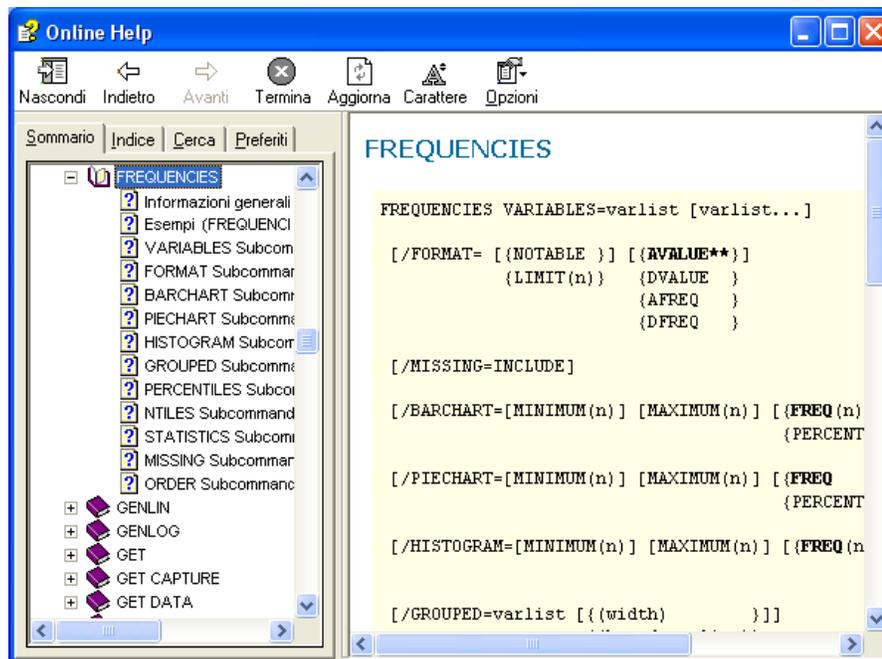
Figura 8-4

Controllo di autocompletamento che visualizza i sottocomandi



Per ottenere una guida più dettagliata per il comando corrente, premere il tasto F1. In tal modo si ottengono direttamente le informazioni di riferimento sulla sintassi dei comandi relative al comando corrente.

Figura 8-5
Aiuto sintassi di FREQUENCIES



Si è forse notato come il testo visualizzato nella finestra della sintassi sia colorato. La codifica colori consente di individuare rapidamente i termini non riconosciuti, dal momento che solo i termini riconosciuti sono colorati. Ad esempio, il sottocomando `FORMAT` viene erroneamente digitato come `FRMAT`. Per impostazione predefinita, i sottocomandi vengono colorati in verde, ma il testo `FRMAT` apparirà senza colori, poiché non è un termine riconosciuto.

Apertura ed esecuzione di un file di sintassi

- Per aprire un file di sintassi salvato, dai menu scegliere:
File > Apri > Sintassi...

Viene visualizzata una finestra di dialogo standard per l'apertura dei file.

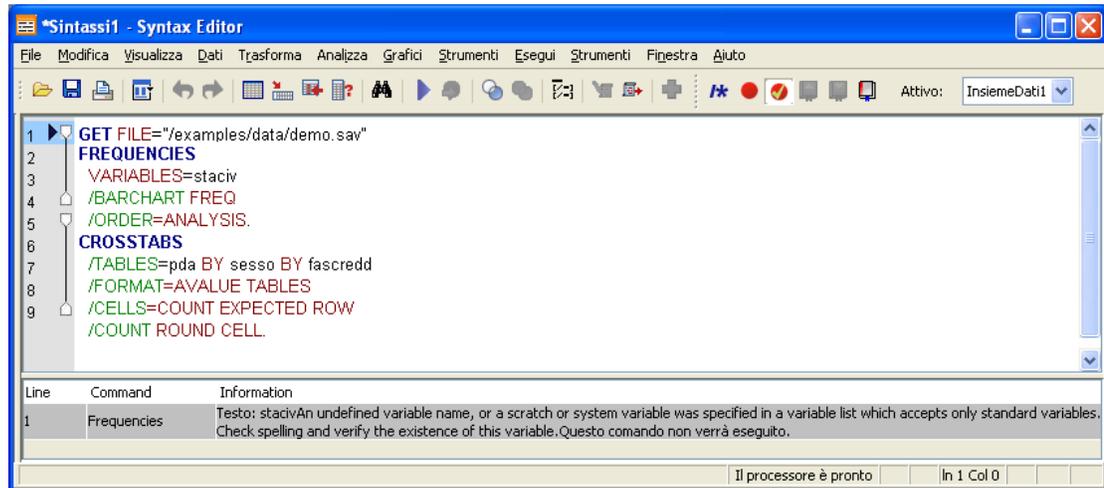
- Selezionare un file di sintassi. Se non vengono visualizzati file di sintassi, accertarsi che venga selezionato Sintassi (*.sps) come tipo di file da visualizzare.
- Fare clic su Apri.
- Utilizzare il menu Esegui nella finestra della sintassi per eseguire i comandi.

Se i comandi sono validi per un tipo di file specifico, prima di eseguirli è necessario aprire il file di dati oppure inserire un comando che apre tale file. Questo tipo di comando può essere incollato dalle finestre di dialogo nelle quali è possibile aprire i file di dati.

Informazioni sul riquadro degli errori

Il riquadro degli errori visualizza gli errori di runtime dall'esecuzione più attuale. Contiene i dettagli di ciascun errore, nonché il numero di riga del comando in cui si è verificato l'errore.

Figura 8-6
Riquadro degli errori visualizzato nell'Editor della sintassi



Facendo clic sulla voce di un errore, il cursore si posiziona sulla prima riga di comando in cui si è verificato l'errore.

Utilizzo dei punti di interruzione

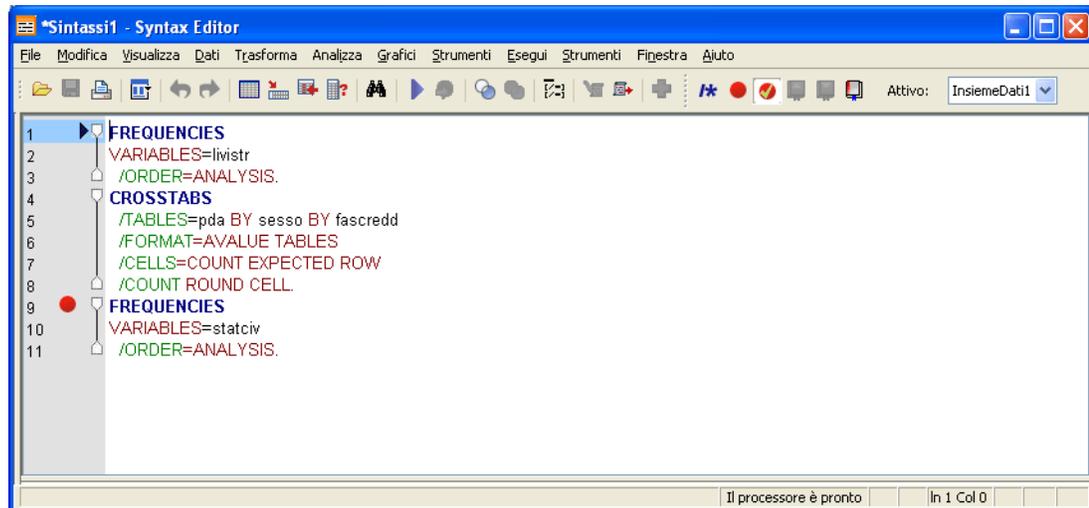
I punti di interruzione consentono di interrompere l'esecuzione della sintassi dei comandi in punti specifici all'interno della sintassi e di riprendere l'esecuzione quando si è pronti. In questo modo è possibile visualizzare l'output o i dati in un punto intermedio di un processo di sintassi, oppure di eseguire una sintassi di comando che visualizza le informazioni sullo stato corrente dei dati, ad esempio `FREQUENCIES`. I punti di interruzione possono essere impostati solo a livello di un comando, non su righe specifiche all'interno di un comando.

Per inserire un punto di interruzione in un comando:

- Fare clic in un punto qualsiasi dell'area alla sinistra del testo associato al comando.

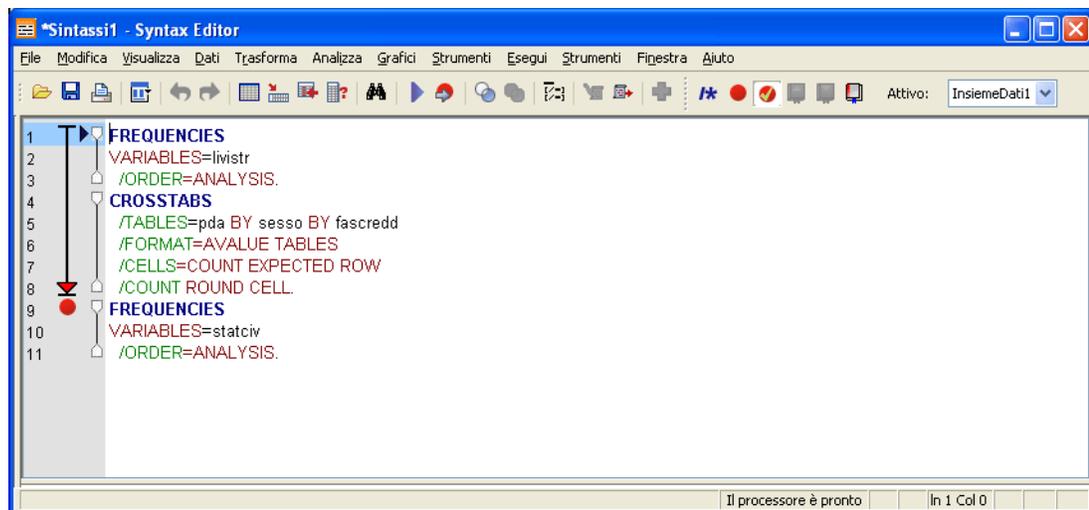
Il punto di interruzione è rappresentato da un cerchio rosso nell'area alla sinistra del testo del comando e sulla stessa riga del nome del comando indipendentemente dal punto in cui si è fatto clic con il mouse.

Figura 8-7
Punto di interruzione visualizzato nella finestra Editor della sintassi



Quando si esegue la sintassi dei comandi che contiene punti di interruzione, l'esecuzione viene interrotta prima di ciascun comando contenente un punto di interruzione.

Figura 8-8
Esecuzione interrotta in corrispondenza di un punto di interruzione



La freccia che punta verso il basso alla sinistra del testo del comando mostra l'avanzamento dell'esecuzione della sintassi. Abbraccia l'area dall'esecuzione del primo comando fino all'esecuzione dell'ultimo comando ed è particolarmente utile quando si esegue la sintassi dei comandi contenente punti di interruzione.

Per riprendere l'esecuzione in seguito a un punto di interruzione:

- ▶ Dai menu dell'Editor della sintassi, scegliere:
Esegui > Continua

Modifica dei valori dei dati

I dati da cui si parte non sempre sono organizzati nel modo più consono alle esigenze analitiche o di reporting. Ad esempio, è possibile che si desideri eseguire una delle operazioni seguenti:

- Creare una variabile categoriale da una variabile di scala.
- Unire più categorie di risposta in un'unica categoria.
- Creare una variabile nuova che rappresenta la differenza calcolata tra due variabili esistenti.
- Calcolare il periodo di tempo tra due date.

In questo capitolo viene utilizzato il file di dati *demo.sav*. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice A a pag. 152.](#)

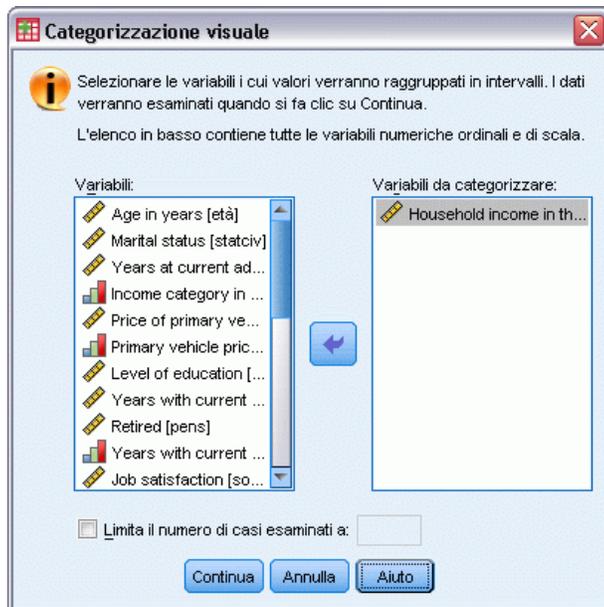
Creazione di una variabile categoriale da una variabile di scala.

Varie variabili categoriali presenti nel file di dati *demo.sav* derivano infatti da variabili di scala presenti nello stesso file di dati. Ad esempio, la variabile *catredd* è semplicemente il *reddito* raggruppato in quattro categorie. Tale variabile categoriale usa i valori interi 1–4 per rappresentare le seguenti categorie di reddito (in migliaia): inferiore a \$25, \$25–\$49, \$50–\$74 e \$75 o superiore.

Per creare la variabile categoriale *inccat*:

- ▶ Dai menu nella finestra dell'Editor dei dati, scegliere:
Trasforma > Categorizzazione visuale...

Figura 9-1
Finestra di dialogo iniziale Categorizzazione visuale



Nella finestra di dialogo iniziale Categorizzazione visuale, selezionare le variabili ordinali e/o di scala per cui si desiderano creare nuove variabili categorizzate. Con il termine **categorizzazione** si intende l'operazione che consente di raggruppare due o più valori contigui nella stessa categoria.

Poiché Categorizzazione visuale si basa sui valori reali del file di dati per eseguire scelte di raggruppamento corrette, prima deve leggere il file di dati. Poiché tale operazione potrebbe richiedere tempi lunghi se il file di dati contiene un numero molto elevato di casi, la finestra di dialogo iniziale consente inoltre di limitare il numero di casi da leggere ("esaminare"). Non si tratta di un'operazione necessaria per il file di dati campione. Sebbene i casi siano oltre 6.000, non occorre molto tempo per esaminarli.

- Trascinare *Reddito familiare in migliaia [reddito]* dall'elenco Variabili all'elenco Variabili da categorizzare, quindi fare clic su Continua.

Figura 9-2
Finestra di dialogo principale Categorizzazione visuale

Categorizzazione visuale

Elenco delle variabili esaminate:
Household income in th...

Nome: Etichetta:
Variabile corrente: reddito Household income in thousands
Variabile categorizzata: inccat2 Household income in thousands(Categorizzato)

Minimo: 9,00 Valori non mancanti Massimo: 1116,00

Inserire i punti di divisione dell'intervallo o fare clic su Crea punti divisione per gli intervalli automatici.
Un valore del punto di divisione di 10, per esempio, definisce un intervallo che inizia sopra a quello precedente e termina a 10.

Griglia:

	Valore	Etichetta
1		ALTO
2		

Punti finali superiori
 Inclusi (<=)
 Esclusi (<)

Crea punti divisione...
Crea etichette

Scala inversa

Casi esaminati: 6400
Valori mancanti: 0

Copia intervalli
Da un'altra variabile...
Ad altre variabili...

OK Incolla Reimposta Annulla Aiuto

- Nella finestra di dialogo principale Categorizzazione visuale, selezionare *Reddito familiare in migliaia [reddito]* nell'elenco delle variabili esaminate.

Un istogramma visualizza la distribuzione della variabile selezionata (che in questo caso è molto asimmetrica).

- Inserire *catredd2* per la nuova variabile categorizzata e *Categoria reddito [in migliaia]* per l'etichetta della variabile.
- Fare clic su *Crea punti divisione*.

Figura 9-3
Finestra di dialogo dei punti di Categorizzazione visuale

Crea punti divisione

Intervalli di larghezza uguale

Intervalli - inserire dati in almeno due campi

Posizione primo punto divisione: 25.00

Numero punti di divisione: 3

Lunghezza: 25

Posizione ultimo punto divisione: 75.00

Percentili uguali basati sui casi esaminati

Intervalli - inserire dati in entrambi i campi

Numero punti di divisione:

Larghezza (%):

Punti di divisione alla media e deviazioni standard selezionati in base ai casi esaminati

Deviazione standard +/- 1

Deviazione standard +/- 2

Deviazione standard +/- 3

i Applica sostituirà le definizioni di punti di divisione correnti con questa impostazione.
Un intervallo finale includerà tutti i valori rimanenti: N punti di divisione creano N+1 intervalli.

Applica Annulla Aiuto

- ▶ Selezionare Intervalli di larghezza uguale
- ▶ Inserire 25 per la posizione del primo punto di divisione, 3 per il numero di punti di divisione e 25 per la larghezza.

Il numero di categorie categorizzate corrisponde al numero di punti di divisione più uno. Quindi, in questo esempio, la nuova variabile categorizzata avrà quattro categorie: le prime tre contengono ognuna intervalli di 25 (in migliaia) e l'ultima contiene tutti i valori sopra al valore del punto di divisione più alto, uguale a 75 (in migliaia).

- ▶ Fare clic su Applica.

Figura 9-4
Finestra di dialogo principale Categorizzazione visuale con punti di divisione definiti

Categorizzazione visuale

Elenco delle variabili esaminate:
Household income in th...

Nome: Etichetta:
Variabile corrente: reddito Household income in thousands
Variabile categorizzata: inccat2 Household income in thousands (Categorizzato)

Minimo: 9,00 Valori non mancanti Massimo: 1116,00

Inserire i punti di divisione dell'intervallo o fare clic su Crea punti divisione per gli intervalli automatici.
Un valore del punto di divisione di 10, per esempio, definisce un intervallo che inizia sopra a quello precedente e termina a 10.

Griglia:

	Valore	Etichetta
1	25.000	
2	50.000	
3	75.000	
4	ALTO	
5		

Punti finali superiori
 Inclusi (<=)
 Esclusi (<)

Crea punti divisione...
 Crea etichette

Scala inversa

Casi esaminati: 6400
 Valori mancanti: 0

Copia intervalli
 Da un'altra variabile...
 Ad altre variabili...

OK Incolla Reimposta Annulla Aiuto

I valori visualizzati adesso nella griglia rappresentano i punti di divisione definiti, ovvero i punti finali superiori di ogni categoria. Le linee verticali nell'istogramma indicano inoltre le posizioni dei punti di divisione.

Per impostazione predefinita, i valori di questi punti di divisione vengono inclusi nelle categorie corrispondenti. Ad esempio, il primo valore pari 25 includerebbe tutti i valori inferiori o uguali a 25. Tuttavia, in questo esempio, si desidera ottenere categorie che corrispondono a valori inferiori a 25, 25–49, 50–74 e 75 o valori superiori.

- ▶ Nel gruppo dei punti finali superiori, selezionare Esclusi (<)
- ▶ Quindi fare clic su Crea etichette.

Figura 9-5
Etichette di variabili generate automaticamente

Categorizzazione visuale

Elenco delle variabili esaminate:
Household income in th...

Nome: reddito Etichetta: Household income in thousands

Variabile corrente: reddito

Variabile categorizzata: inccat2 Etichetta: Household income in thousands(Categorizzato)

Minimo: 9.00 Valori non mancanti Massimo: 1116.00

Inserire i punti di divisione dell'intervallo o fare clic su Crea punti divisione per gli intervalli automatici.
Un valore del punto di divisione di 10, per esempio, definisce un intervallo che inizia sopra a quello precedente e termina a 10.

Casi esaminati: 6400
Valori mancanti: 0

Copia intervalli
Da un'altra variabile...
Ad altre variabili...

	Valore	Etichetta
1	25.000	< 25.00
2	50.000	25.00 - 49.00
3	75.000	50.00 - 74.00
4	ALTO	75.00+
5		

Punti finali superiori
 Inclusi (<=)
 Esclusi (<)

Crea punti divisione...
Crea etichette

Scala inversa

OK Incolla Reimposta Annulla Aiuto

Questa operazione genera automaticamente etichette di valore descrittive per ogni categoria. Poiché i valori effettivi assegnati alla nuova variabile categorizzata sono semplicemente interi sequenziali che iniziano con 1, le etichette di valore possono risultare molto utili.

È inoltre possibile specificare o modificare manualmente i punti di divisione e le etichette nella griglia, modificare le posizioni dei punti di divisione trascinando le linee dei punti di divisione nell'istogramma ed eliminare i punti di divisione trascinando le linee dei punti di divisione all'esterno dell'istogramma.

- Fare clic su OK per creare la nuova variabile categorizzata.

La nuova variabile viene visualizzata nell'Editor dei dati. Poiché la variabile viene aggiunta alla fine del file, viene visualizzata nella colonna più a destra della Visualizzazione dati e nell'ultima riga della Visualizzazione variabili.

Figura 9-6
Nuova variabile visualizzata nell'Editor dei dati

	pc	fax	quotid	risposta	inccat2
1	No	No	No	Sì	50,00 - 74,00
2	No	No	No	No	75,00+
3	Sì	No	Sì	Sì	25,00 - 49,00
4	Sì	Sì	Sì	Sì	25,00 - 49,00
5	No	No	Sì	Sì	< 25,00
6	Sì	No	No	Sì	75,00+
7	No	No	No	Sì	25,00 - 49,00
8	Sì	No	No	Sì	50,00 - 74,00
9	No	No	Sì	Sì	< 25,00
10	No	Sì	Sì	No	75,00+

Calcolo di nuove variabili

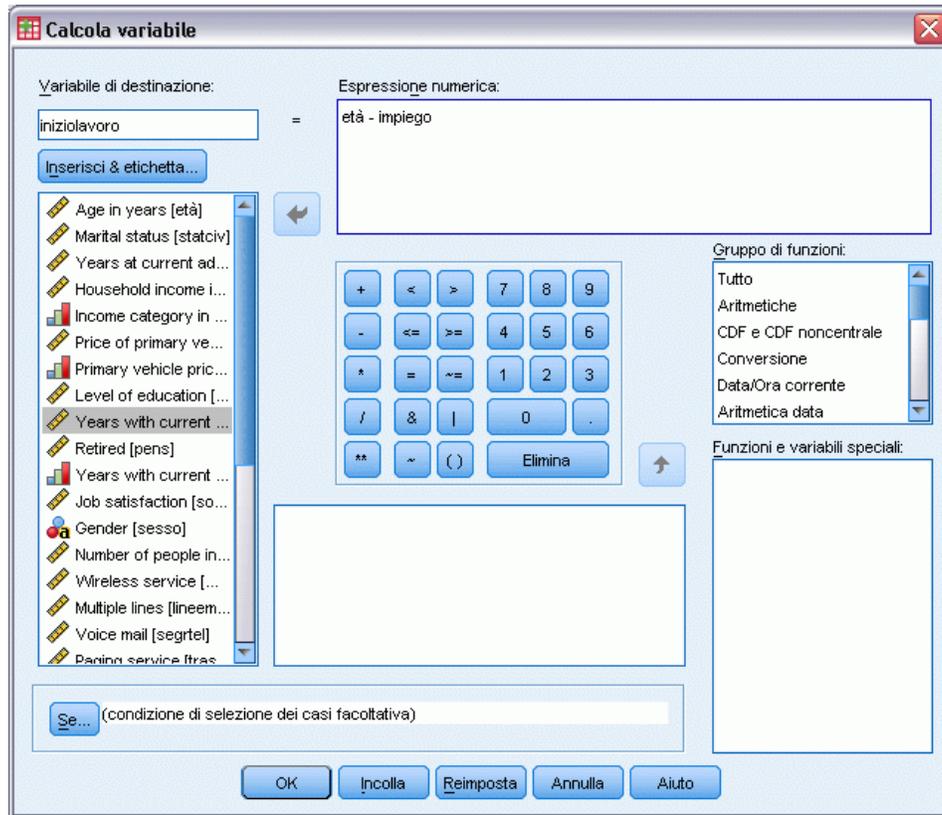
È possibile calcolare nuove variabili basandosi su equazioni altamente complesse, utilizzando una grande varietà di funzioni matematiche. Tuttavia, in questo esempio verrà semplicemente calcolata una nuova variabile che è la differenza tra i valori di due variabili esistenti.

Il file di dati *demo.sav* contiene una variabile per l'età corrente del rispondente e una variabile per il numero di anni trascorsi nell'attuale posto di lavoro. Tuttavia, non contiene una variabile per l'età del rispondente nel momento in cui iniziò tale lavoro. È possibile creare una nuova variabile che è la differenza calcolata tra l'età corrente e il numero di anni trascorsi nell'attuale posto di lavoro, che dovrebbe essere l'età approssimativa in cui il rispondente ha iniziato tale lavoro.

- ▶ Dai menu nella finestra dell'Editor dei dati, scegliere:
Trasforma > Calcola variabile...
- ▶ Per la Variabile di destinazione, digitare *iniziolavoro*.
- ▶ Selezionare *Età in anni [età]* nell'elenco delle variabili sorgente e fare clic sul pulsante freccia per copiarla nella casella di testo *Espressione numerica*.
- ▶ Fare clic sul pulsante meno (–) della calcolatrice nella finestra di dialogo (oppure premere il tasto meno della tastiera).

- Selezionare *Anni di permanenza nell'impiego attuale [impiego]* e fare clic sul pulsante freccia per copiarla nell'espressione.

Figura 9-7
Finestra di dialogo *Calcola variabile*



Nota: Fare attenzione e selezionare la variabile occupazionale corretta. Esiste anche una versione categoriale ricodificata della variabile, che *non* è quella desiderata. L'espressione numerica deve essere *età-impiego*, non *età-catimp*.

- Fare clic su OK per calcolare la nuova variabile.

La nuova variabile viene visualizzata nell'Editor dei dati. Poiché la variabile viene aggiunta alla fine del file, viene visualizzata nella colonna più a destra della Visualizzazione dati e nell'ultima riga della Visualizzazione variabili.

Figura 9-8

Nuova variabile visualizzata nell'Editor dei dati

	fax	quotid	risposta	inccat2	jobstart	
1	No	No	No	Sì	50,00 - 74,00	52,00
2	No	No	No	No	75,00+	53,00
3	Sì	No	Sì	Sì	25,00 - 49,00	27,00
4	Sì	Sì	Sì	Sì	25,00 - 49,00	23,00
5	No	No	Sì	Sì	< 25,00	23,00
6	Sì	No	No	Sì	75,00+	43,00
7	No	No	No	Sì	25,00 - 49,00	40,00
8	Sì	No	No	Sì	50,00 - 74,00	34,00
9	No	No	Sì	Sì	< 25,00	44,00
10	No	Sì	Sì	No	75,00+	32,00
11	Sì	No	Sì	Sì	50,00 - 74,00	54,00

Utilizzo delle funzioni nelle espressioni

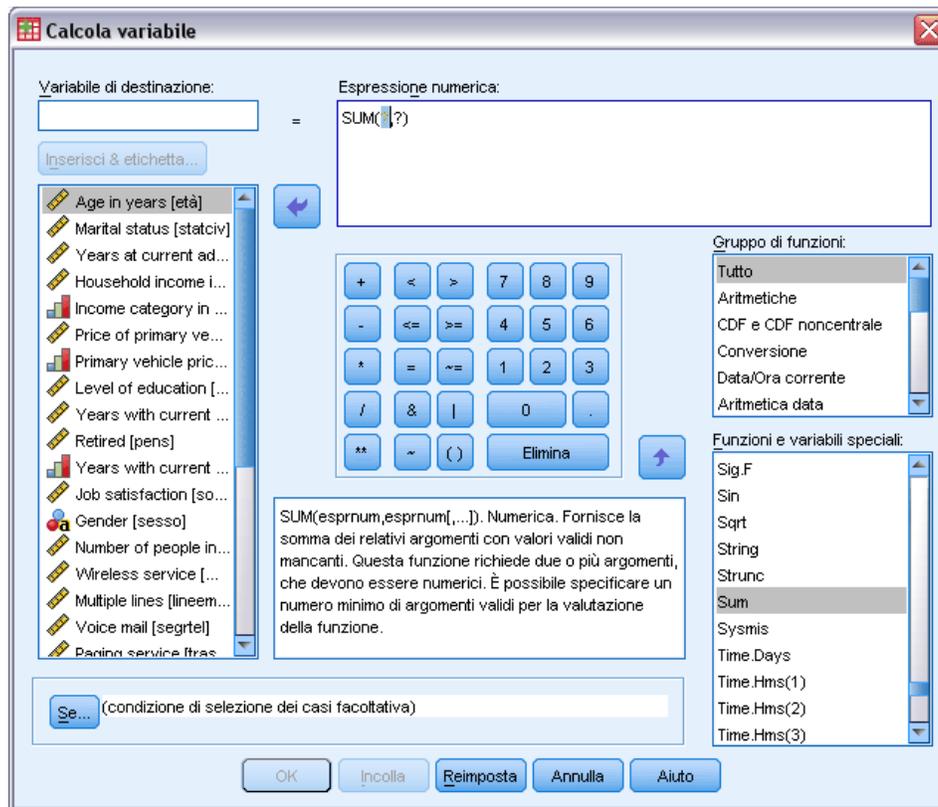
È inoltre possibile utilizzare le funzioni predefinite nelle espressioni. Sono disponibili oltre 70 funzioni incorporate, tra le quali:

- Funzioni aritmetiche
- Funzioni statistiche
- Funzioni di distribuzione
- Funzioni logiche
- Funzioni di aggregazione e di estrazione relative alla data e all'ora
- Funzioni per valori mancanti

- Funzioni incrociate
- Funzioni stringa

Figura 9-9

Finestra di dialogo Calcola variabile con raggruppamento delle funzioni visualizzato



Le funzioni sono organizzate in gruppi divisi in modo logico, ad esempio uno per le operazioni aritmetiche e un altro per il calcolo delle metriche statistiche. Per comodità, anche alcune variabili di sistema di uso comune, come *\$TIME* (ora e data correnti) sono incluse in gruppi di funzioni appropriate. Una breve descrizione della funzione (in questo caso, *SUM*) o della variabile di sistema attualmente selezionata è visualizzata in un'area riservata nella finestra di dialogo Calcola variabile.

Incollare una funzione in un'espressione

Per incollare una funzione in un'espressione:

- ▶ Posizionare il cursore all'interno dell'espressione, nel punto in cui si desidera visualizzare la funzione.
- ▶ Selezionare il gruppo appropriato dall'elenco Gruppo di funzioni. Il gruppo con etichetta Tutti fornisce un elenco di tutte le funzioni e di tutte le variabili di sistema disponibili.
- ▶ Fare doppio clic sulla funzione desiderata nell'elenco Funzioni e Variabili speciali, oppure selezionare la funzione e fare clic sulla freccia accanto all'elenco Gruppo di funzioni.

La funzione verrà inserita nell'espressione. Se si seleziona una parte dell'espressione e quindi si inserisce la funzione, la parte selezionata diventerà il primo argomento della funzione.

Modifica di una funzione di un'espressione

La funzione non è completa fino a quando non si inseriscono gli argomenti, rappresentati da punti interrogativi nella funzione incollata. Il numero dei punti interrogativi indica il numero minimo di argomenti necessari per completare la funzione.

- ▶ Evidenziare i punti interrogativi nella funzione incollata.
- ▶ Inserire gli argomenti. Se si tratta di nomi di variabili, è possibile incollarli dall'elenco di variabili.

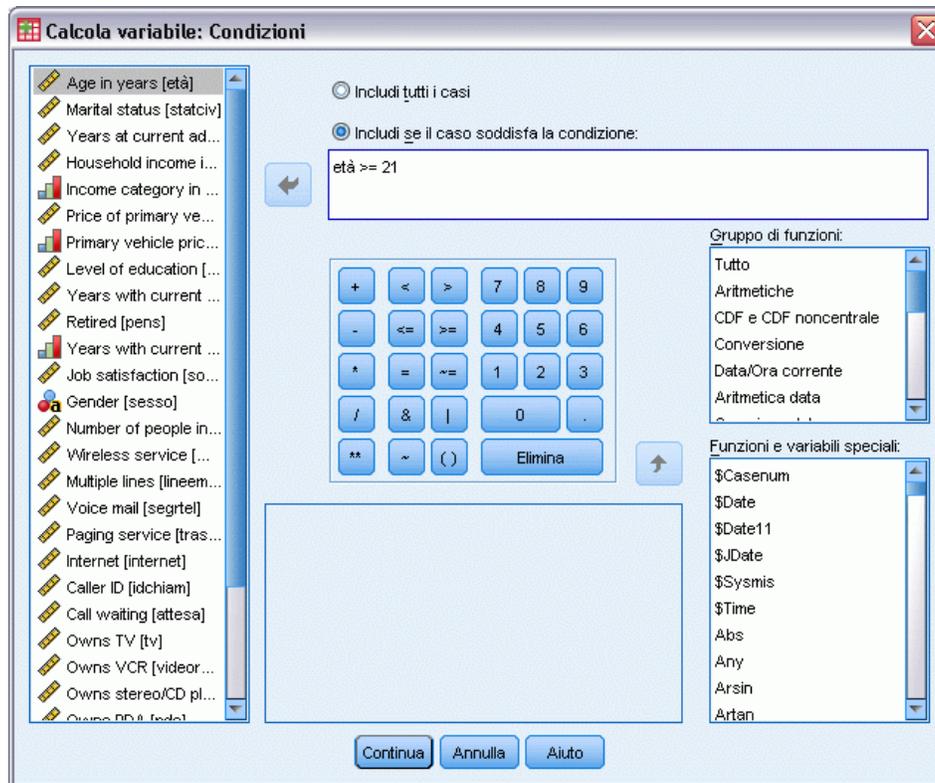
Utilizzo delle espressioni logiche

Le espressioni logiche consentono di applicare le trasformazioni ai sottoinsiemi di casi selezionati. Un'espressione logica restituisce un valore vero, falso o mancante per ciascun caso. Se il risultato dell'espressione logica è vero, la trasformazione verrà applicata al caso. Se il risultato è falso o mancante, la trasformazione non verrà applicata al caso.

Per specificare un'espressione logica:

- ▶ Fare clic su **Se** nella finestra di dialogo **Calcola variabile**. Viene aperta la finestra di dialogo **Condizioni**.

Figura 9-10
Finestra di dialogo Condizioni



- Selezionare Includi se il caso soddisfa la condizione.
- Specificare l'espressione condizionale.

La maggior parte delle espressioni logiche contiene almeno un operatore relazionale, ad esempio

`età >= 21`

o

`reddito * 3 < 100`

Nel primo esempio, verranno selezionati solo i casi con un valore uguale o maggiore a 21 per *Età [età]*. Nel secondo esempio, *Reddito familiare in migliaia [reddito]* moltiplicato per 3 deve essere inferiore a 100 affinché venga selezionato un caso.

È inoltre possibile collegare due o più espressioni logiche mediante gli operatori logici, ad esempio:

`età >= 21 | ed >= 4`

o

`reddito * 3 < 100 & istruz = 5`

Nel primo esempio verranno selezionati i casi che soddisfano la condizione di *Età [età]* oppure il *Livello di istruzione [istrucz]*. Nel secondo esempio, entrambe le condizioni *Reddito familiare in migliaia [reddito]* e *Livello di istruzione [istrucz]* devono essere soddisfatte affinché venga selezionato un caso.

Utilizzo di date e orari

Numerose operazioni eseguite comunemente con date e orari possono essere realizzate facilmente utilizzando l'Impostazione guidata data e ora. L'Impostazione guidata offre le funzioni seguenti:

- Creazione di una variabile di data/ora da una variabile stringa che contiene una data o un orario.
- Creazione di una variabile di data/ora dall'unione di variabili contenenti varie parti della data o dell'ora.
- Aggiunta o sottrazione di valori da variabili di data/ora, inclusa aggiunta o sottrazione di due variabili di data/ora.
- Estrazione di una parte di una variabile di data o ora; ad esempio, il giorno del mese da una variabile di data/ora con formato mm/gg/aaaa.

Gli esempi di questa sezione utilizzano il file di dati *upgrade.sav*. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice A a pag. 152.](#)

Per utilizzare l'Impostazione guidata data e ora

- ▶ Dai menu, scegliere:
Trasforma > Procedura guidata Data e ora...

Figura 9-11
Schermata introduttiva dell'Impostazione guidata data e ora



Nella schermata introduttiva dell'Impostazione guidata data e ora è visualizzato un gruppo di operazioni di carattere generale. Le operazioni non applicabili ai dati correnti sono disabilite. Ad esempio, il file di dati *upgrade.sav* non contiene variabili stringa, quindi l'attività di creazione di una variabile di data da una stringa è disattivata.

Se non si ha familiarità con le operazioni relative a date e ore in IBM® SPSS® Statistics, selezionare Visualizzare informazioni sulla rappresentazione di date e ore e fare clic su Successivo. Verrà visualizzata una schermata che fornisce una breve panoramica sulle variabili di data e ora e un collegamento, tramite il pulsante Aiuto, a informazioni più dettagliate.

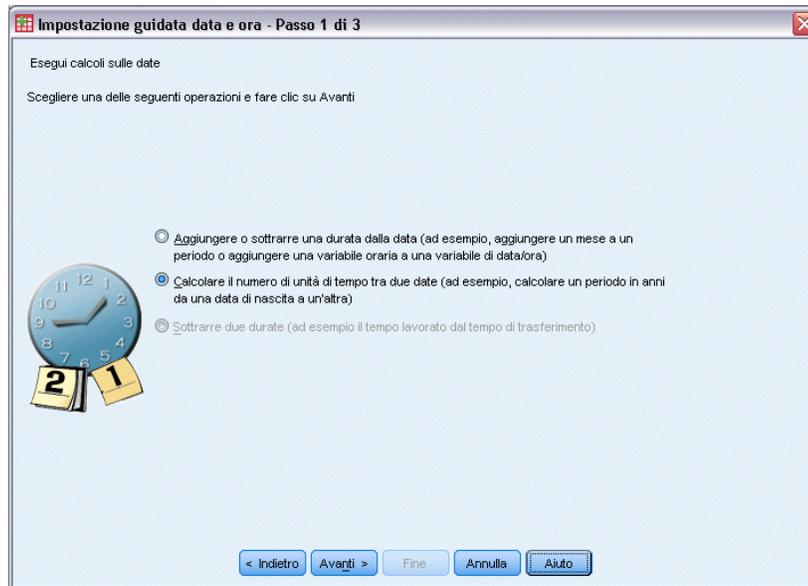
Calcolo del periodo di tempo tra due date

Una delle attività più comuni relative alle date è il calcolo del periodo di tempo tra due date. Si consideri ad esempio un'azienda produttrice di software interessata ad analizzare gli acquisti di licenze di aggiornamento tramite la determinazione del numero di anni dall'ultimo acquisto di un aggiornamento da parte di ogni cliente. Il file di dati *upgrade.sav* contiene una variabile per la data in cui ogni cliente ha acquistato l'ultimo aggiornamento ma non il numero di anni trascorsi da tale data. Una nuova variabile che rappresenti il periodo di tempo in anni tra la data dell'ultimo aggiornamento e quella della prossima release del prodotto fornirebbe una misura di questa quantità.

Per calcolare il periodo di tempo tra due date:

- Selezionare Eseguire calcoli con date e ore nella schermata introduttiva dell'Impostazione guidata data e ora e fare clic su Avanti.

Figura 9-12
Calcolo del periodo di tempo tra due date: Passo 1



- Selezionare *Calcolare il numero di unità di tempo tra due date* e fare clic su *Avanti*.

Figura 9-13
Calcolo del periodo di tempo tra due date: Passo 2



- Selezionare *Data della prossima release* per *Data1*.
- Selezionare *Data dell'ultimo aggiornamento* per *Data2*.
- Selezionare *Anni* per l'*Unità* e *Tronca a numero intero* per il *Trattamento risultato*. Si tratta di selezioni predefinite.

- Fare clic su Avanti.

Figura 9-14

Calcolo del periodo di tempo tra due date: Passo 3

Impostazione guidata data e ora - Passo 3 di 3

Calcolo: ProssimaRelease - UltimoUpgrade

Variabile risultante: AnniUltAgg Unità: Anni

Etichetta variabile: Anni dall'ultimo aggiornamento

Esecuzione

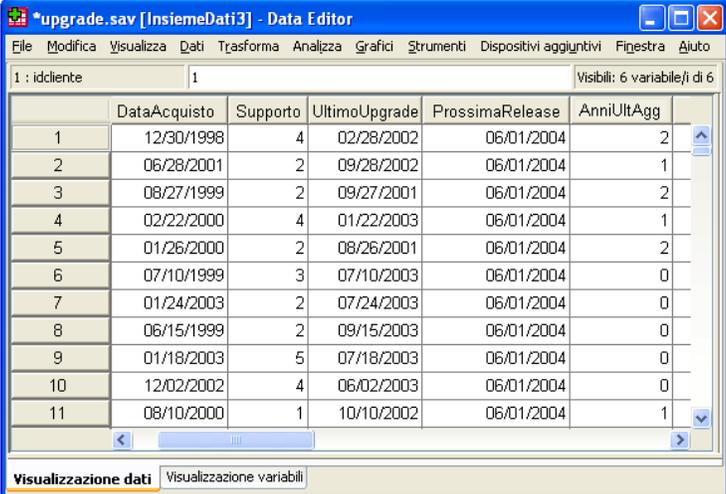
Crea la variabile ora Incolla la sintassi nell'apposita finestra

< Indietro Avanti > Fine Annulla Aiuto

- Inserire *AnniUltAgg* come nome della variabile di risultato. Le variabili di risultato non possono avere lo stesso nome di una variabile esistente.
- Digitare *Anni dall'ultimo aggiornamento* come etichetta per la variabile di risultato. Le etichette per le variabili di risultato sono facoltative.
- Non modificare la selezione predefinita di Crea la variabile ora e fare clic su Fine per creare la nuova variabile.

La nuova variabile, *AnniUltAgg*, visualizzata nell'Editor dei dati, rappresenta il numero intero di anni tra le due date. Le parti frazionarie di un anno vengono troncate.

Figura 9-15
Nuova variabile visualizzata nell'Editor dei dati



	DataAcquisto	Supporto	UltimoUpgrade	ProssimaRelease	AnniUltAgg
1	12/30/1998	4	02/28/2002	06/01/2004	2
2	06/28/2001	2	09/28/2002	06/01/2004	1
3	08/27/1999	2	09/27/2001	06/01/2004	2
4	02/22/2000	4	01/22/2003	06/01/2004	1
5	01/26/2000	2	08/26/2001	06/01/2004	2
6	07/10/1999	3	07/10/2003	06/01/2004	0
7	01/24/2003	2	07/24/2003	06/01/2004	0
8	06/15/1999	2	09/15/2003	06/01/2004	0
9	01/18/2003	5	07/18/2003	06/01/2004	0
10	12/02/2002	4	06/02/2003	06/01/2004	0
11	08/10/2000	1	10/10/2002	06/01/2004	1

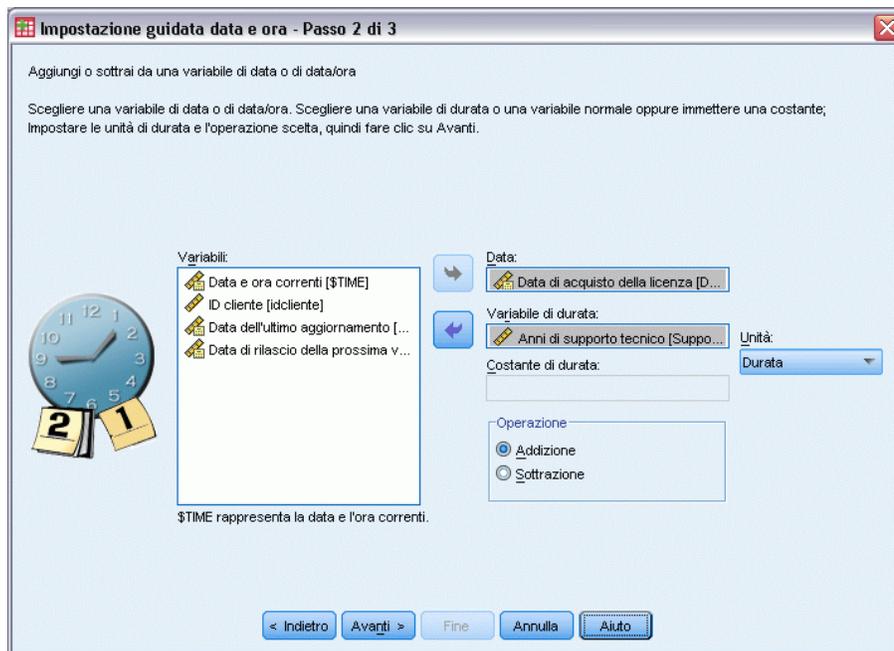
Aggiunta di una durata a una data

È possibile aggiungere o sottrarre a una data durate come 10 giorni o 12 mesi. Continuando con l'esempio della sezione precedente, si consideri di voler determinare la data di termine del contratto iniziale di assistenza tecnica di ogni cliente. Il file di dati *upgrade.sav* contiene una variabile per il numero di anni di durata del contratto di assistenza e una variabile per la data di acquisto iniziali. È quindi possibile determinare la data finale del supporto iniziale aggiungendo gli anni di assistenza alla data di acquisto.

Per aggiungere una durata a una data:

- ▶ Selezionare Eseguire calcoli con date e ore nella schermata introduttiva dell'Impostazione guidata data e ora e fare clic su Avanti.
- ▶ Selezionare Aggiungere o sottrarre una durata da una data e fare clic su Avanti.

Figura 9-16
Aggiunta di una durata a una data: Passo 2



► Selezionare *Data della licenza prodotto iniziale* come Data.

► Selezionare *Anni di assistenza tecnica* come variabile di durata.

Poiché *Anni di assistenza tecnica* è semplicemente una variabile numerica, è necessario indicare le unità da utilizzare quando si aggiunge tale variabile come una durata.

► Selezionare Anni dall'elenco a discesa Unità.

► Fare clic su Avanti.

Figura 9-17
 Aggiunta di una durata a una data: Passo 3

- ▶ Inserire *DataFinAss* come nome della variabile di risultato. Le variabili di risultato non possono avere lo stesso nome di una variabile esistente.
- ▶ Inserire *Data finale assistenza* come etichetta della variabile di risultato. Le etichette per le variabili di risultato sono facoltative.
- ▶ Fare clic su *Fine* per creare la nuova variabile.

La nuova variabile viene visualizzata nell'Editor dei dati.

Figura 9-18
 Nuova variabile visualizzata nell'Editor dei dati

1 : idcliente	Supporto	UltimoUpgrade	ProssimaRelease	AnniUltAgg	DataFinAss
1	4	02/28/2002	06/01/2004	2	12/30/2002
2	2	09/28/2002	06/01/2004	1	06/28/2003
3	2	09/27/2001	06/01/2004	2	08/27/2001
4	4	01/22/2003	06/01/2004	1	12/22/2003
5	2	08/26/2001	06/01/2004	2	01/26/2002
6	3	07/10/2003	06/01/2004	0	07/10/2002
7	2	07/24/2003	06/01/2004	0	01/24/2005
8	2	09/15/2003	06/01/2004	0	06/15/2001
9	5	07/18/2003	06/01/2004	0	01/18/2008
10	4	06/02/2003	06/01/2004	0	12/02/2006
11	1	10/10/2002	06/01/2004	1	08/10/2001

Ordinamento e selezione di dati

I file di dati non sempre sono organizzati nella forma ottimale per esigenze specifiche. È tuttavia disponibile una vasta gamma di trasformazioni di file che consentono di preparare i dati per l'analisi, ad esempio:

- **Ordinare i dati.** È possibile ordinare i dati in base al valore di una o più variabili.
- **Selezionare sottoinsiemi di casi.** È possibile limitare l'analisi a un sottoinsieme di casi oppure eseguire contemporaneamente più analisi in diversi sottoinsiemi.

Gli esempi di questo capitolo utilizzano il file di dati *demo.sav*. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice A a pag. 152.](#)

Ordinamento di dati

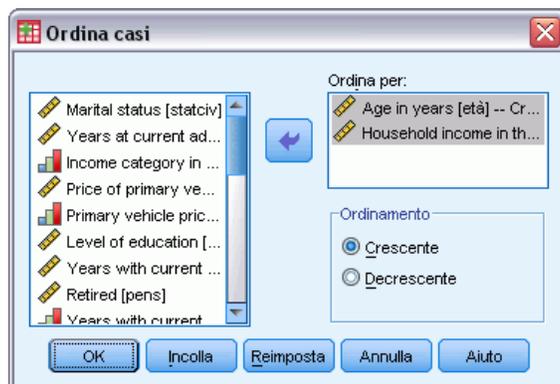
In alcuni tipi di analisi è spesso utile e a volte necessario ordinare i dati, ovvero ordinare le righe del file di dati.

Per riordinare la sequenza dei casi del file di dati in base al valore di una o più variabili di ordinamento:

- Dai menu, scegliere:
Dati > Ordina casi...

Si visualizza la Finestra di dialogo Ordina casi.

Figura 10-1
Finestra di dialogo Ordina casi.



- Aggiungere le variabili *Età in anni [età]* e *Reddito familiare in migliaia [reddito]* all'elenco Ordinamento per.

Se si selezionano più variabili di ordinamento, l'ordinamento dei casi sarà determinato dall'ordine in cui tali variabili sono visualizzate nell'elenco Ordinamento per. In questo esempio, in base all'elenco Ordinamento per, i casi verranno ordinati in base al valore di *Reddito familiare in migliaia [reddito]* all'interno delle categorie di *Età in anni [età]*. Le variabili stringa verranno ordinate in modo che le lettere maiuscole precedano le minuscole corrispondenti. Ad esempio, la stringa *Si* precederà la stringa *si*.

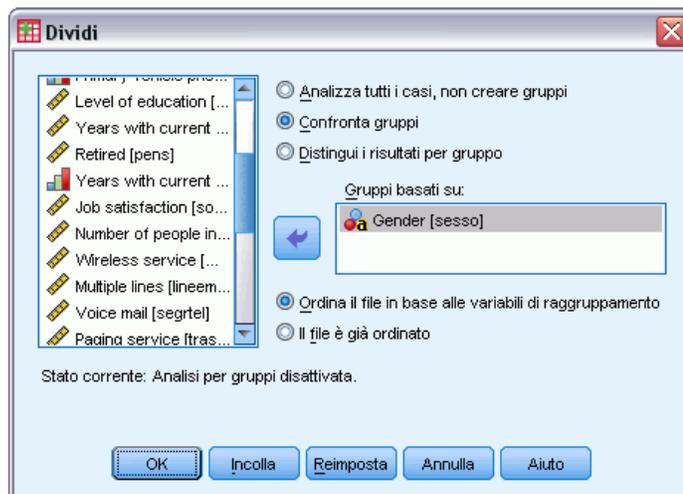
Distinzione per gruppi

Per suddividere il file di dati in gruppi distinti per l'analisi:

- ▶ Dai menu, scegliere:
Dati > Dividi...

Viene visualizzata la finestra di dialogo Dividi.

Figura 10-2
Finestra di dialogo Dividi



- ▶ Selezionare Confronta gruppi o Distingui i risultati per gruppo. (Nell'esempio seguente vengono illustrate le differenze tra queste due opzioni).
- ▶ Selezionare *Sesso [sesso]* per distinguere il file in gruppi separati in base a tali variabili.

È possibile utilizzare variabili numeriche, stringa corta e stringa lunga come variabili di raggruppamento. Per ciascun sottogruppo definito dalle variabili di raggruppamento verrà eseguita un'analisi distinta. Se si selezionano più variabili di raggruppamento, la modalità di raggruppamento dei casi sarà determinata dall'ordine in cui tali variabili sono visualizzate nell'elenco Gruppi basati su.

Se si seleziona Confronta gruppi, i risultati di tutti i gruppi di analisi verranno inclusi nella stessa tabella, come indicato nella tabella delle statistiche riassuntive seguente, generata dalla procedura Frequenze.

Figura 10-3
Output di distinzione file con una singola tabella pivot

Statistiche^a

Fascia di reddito in migliaia

Femmina	N	Validi	1157
		Mancanti	0
	Media		2,1530
	Mediana		2,0000
	Deviazione std.		1,02896
Maschio	N	Validi	1051
		Mancanti	0
	Media		2,1970
	Mediana		2,0000
	Deviazione std.		1,05299

Se si seleziona Distingui i risultati per gruppo e si esegue la procedura Frequenze, verranno create due tabelle pivot, una per le femmine e una per i maschi.

Figura 10-4
Output di distinzione file con una tabella pivot per le femmine

Statistiche^a

Fascia di reddito in migliaia

N	Validi	1157
	Mancanti	0
Media		2,1530
Mediana		2,0000
Deviazione std.		1,02896

a. Sesso = Femmina

Figura 10-5
Output di distinzione file con una tabella pivot per i maschi

Statistiche^a

Fascia di reddito in migliaia

N	Validi	1051
	Mancanti	0
Media		2,1970
Mediana		2,0000
Deviazione std.		1,05299

a. Sesso = Maschio

Ordinamento dei casi per la distinzione per gruppi

La procedura Dividi crea un nuovo sottogruppo ogni volta che rileva un valore diverso per una delle variabili di raggruppamento. È quindi importante ordinare i casi in base ai valori delle variabili di raggruppamento prima di avviare la distinzione per gruppi.

Per impostazione predefinita la procedura Dividi ordina automaticamente il file di dati in base ai valori delle variabili di raggruppamento. Se al file è già applicato l'ordinamento corretto, è possibile ridurre i tempi di elaborazione selezionando il file è già ordinato.

Attivazione e disattivazione della distinzione per gruppi

Dopo essere stata avviata, la distinzione per gruppi rimane attiva per il resto della sessione fino a quando non viene disattivata.

- **Analizza tutti i casi.** Questa opzione disattiva la distinzione per gruppi.
- **Confronta gruppi e Distingui i risultati per gruppo.** Questa opzione attiva la distinzione per gruppi.

Se la distinzione per gruppi è attiva, nella barra di stato nella parte inferiore della finestra dell'applicazione verrà visualizzato il messaggio Dividi attivo.

Selezione di sottoinsiemi di casi

È possibile limitare l'analisi a un sottogruppo specifico in base a criteri che includono variabili ed espressioni complesse. È inoltre possibile selezionare un campione casuale di casi. I criteri in base ai quali viene definito un sottogruppo includono:

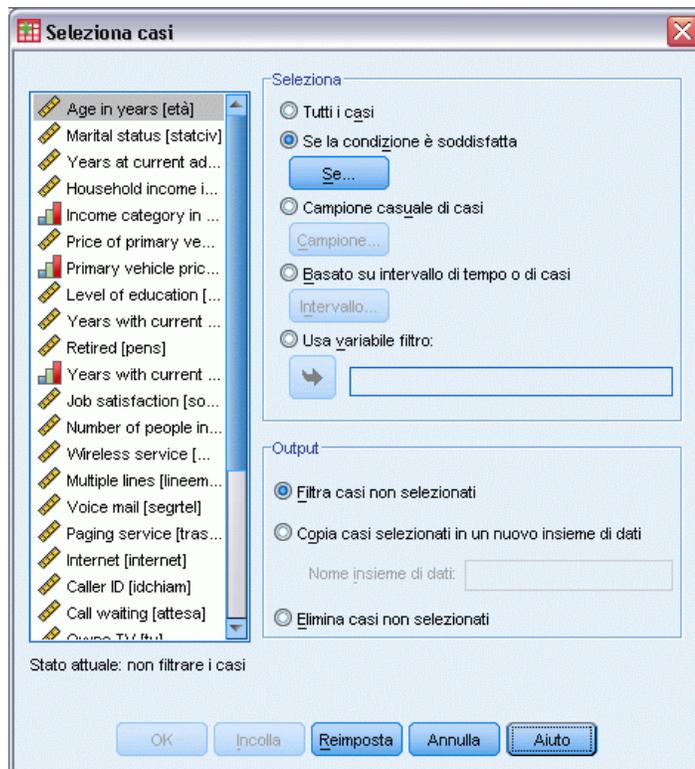
- Valori di variabili e intervalli
- Intervalli di date e orari
- Numeri di casi (righe)
- Espressioni aritmetiche
- Espressioni logiche
- Funzioni

Per selezionare un sottoinsieme di casi per l'analisi:

- ▶ Dai menu, scegliere:
Dati > Seleziona casi...

Verrà aperta la finestra di dialogo Seleziona casi.

Figura 10-6
Finestra di dialogo *Seleziona casi*

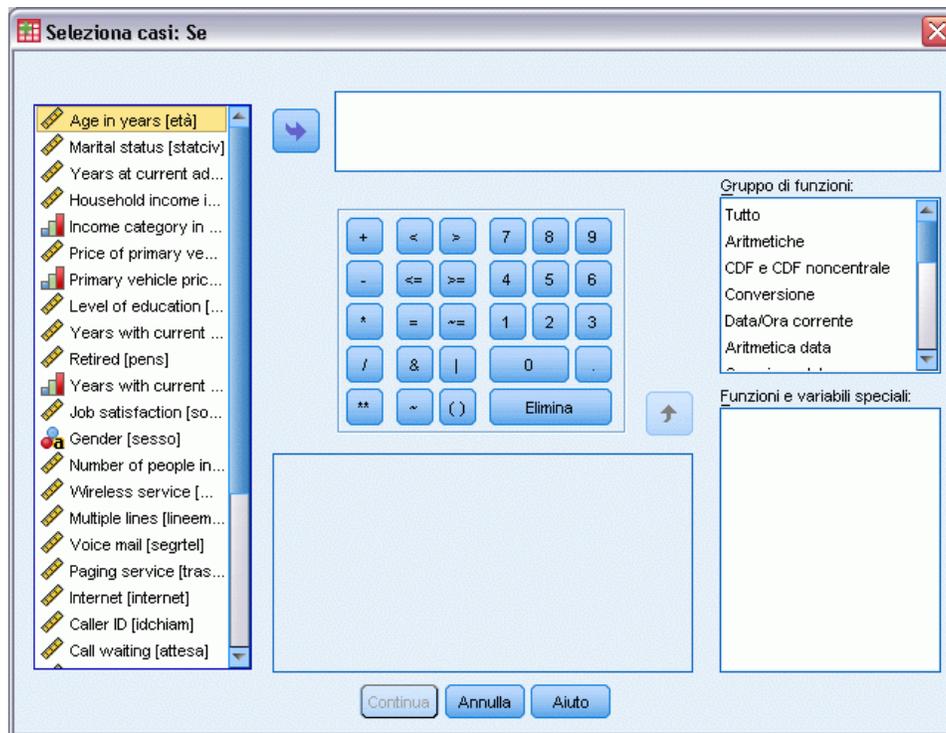


Per selezionare sottoinsiemi di casi in base a un'espressione logica

Per selezionare casi in base a un'espressione logica:

- Selezionare **Se la condizione è soddisfatta** e fare clic su **Se** nella finestra di dialogo **Seleziona casi**.
Verrà aperta la finestra di dialogo **Seleziona casi: Se**.

Figura 10-7
Finestra di dialogo *Seleziona casi: Se*



L'espressione logica può utilizzare nomi di variabili esistenti, costanti, operatori aritmetici, operatori logici, operatori relazionali e funzioni. È possibile digitare e modificare l'espressione nella casella di testo allo stesso modo del testo nella finestra dell'output. È inoltre possibile utilizzare la calcolatrice, l'elenco delle variabili e l'elenco delle funzioni per incollare elementi nell'espressione. [Per ulteriori informazioni, vedere l'argomento Utilizzo delle espressioni logiche in il capitolo 9 a pag. 133.](#)

Per selezionare un campione casuale di dati

Per ottenere un campione casuale di dati:

- ▶ Selezionare Campione casuale di casi nella finestra di dialogo *Seleziona casi*.
- ▶ Fare clic su *Campione*.

Verrà aperta la finestra di dialogo *Seleziona casi: Campione casuale*.

Figura 10-8
Finestra di dialogo *Seleziona casi: Campione casuale*



È possibile selezionare una delle opzioni seguenti per la dimensione del campione:

- **Approssimativamente.** Una percentuale specificata dall'utente. Questa opzione genera un campione casuale che include approssimativamente la percentuale di casi specificata.
- **Esattamente.** Un numero di casi specificato dall'utente. È necessario specificare anche il numero di casi in base al quale generare il campione. Il secondo numero deve essere uguale o minore del numero totale di casi del file dati. Se questo numero supera il numero totale di casi del file di dati, in proporzione il campione conterrà un numero di dati inferiore a quello specificato.

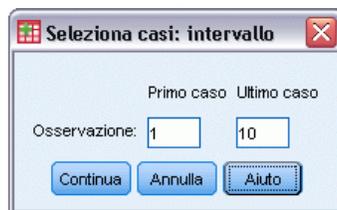
Selezione di un intervallo di tempo o di casi

Per selezionare un intervallo di casi in base alla data, all'ora o al numero di osservazione (riga):

- ▶ Selezionare Basato su intervallo di tempo o di casi e fare clic su Intervallo nella finestra di dialogo *Seleziona casi*.

Verrà aperta la finestra di dialogo *Seleziona casi: intervallo*, in cui è possibile selezionare un intervallo di numeri di osservazione (riga).

Figura 10-9
Finestre di dialogo *Seleziona casi: intervallo*

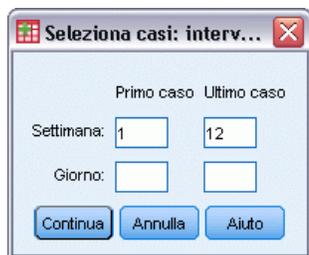


- **Primo caso.** Specificare i valori della data e/o dell'orario iniziale relativi all'intervallo. Se non sono definite variabili della data, specificare il numero dell'osservazione iniziale (numero di riga nell'Editor dei dati a meno che non sia attivo Dividi). Se non si specifica un valore per Ultimo caso, verranno selezionati tutti i casi a partire dalla data o dell'orario iniziale fino alla fine della serie storica.
- **Ultimo caso.** Specificare i valori della data e/o dell'orario finale relativi all'intervallo. Se non sono definite variabili della data, specificare il numero dell'osservazione finale (numero di riga nell'Editor dei dati a meno che non sia attivo Dividi). Se non si specifica un valore

per Primo caso, verranno selezionati tutti i casi a partire dall'inizio della serie storica fino alla data o all'orario finale.

Per i dati di serie storiche con variabili di data definite, è possibile selezionare un intervallo di date e/o orari basato sulle variabili di data definite. Ciascun caso rappresenta osservazioni eseguite in momenti diversi e l'ordinamento del file è cronologico.

Figura 10-10
Finestre di dialogo *Seleziona casi: intervallo (serie storiche)*



Per generare variabili di data per i dati di serie storiche:

- ▶ Dai menu, scegliere:
Dati > Definisci date...

Trattamento dei casi non selezionati

Per il trattamento dei casi non selezionati sono disponibili le opzioni seguenti:

- **Rimuovi casi non selezionati.** I casi non selezionati non saranno inclusi nell'analisi, ma rimarranno nel file di dati. e sarà possibile utilizzarli successivamente nella sessione se si disattivano le funzioni di filtro. Se si seleziona un campione casuale o si selezionano casi in base a un'espressione logica, verrà creata una variabile denominata *filtro_\$* con un valore 1 per i casi selezionati e un valore 0 per i casi non selezionati.
- **Copia casi selezionati in un nuovo file di dati.** I casi selezionati vengono copiati in un nuovo file di dati senza modificare quello originale. I casi non selezionati non vengono inclusi nel nuovo file di lavoro e vengono lasciati nello stato originale nel file di dati originale.
- **Elimina casi non selezionati.** I casi non selezionati verranno cancellati dal file di dati. Per recuperare i casi cancellati è necessario uscire dal file senza salvare le modifiche e quindi riaprire il file. Se si salvano le modifiche al file di dati, l'eliminazione dei casi sarà definitiva.

Nota: Se si eliminano i casi non selezionati e si salva il file, non sarà possibile recuperare i casi.

Criteri di selezione dei casi

Se si è selezionato un sottoinsieme di casi senza eliminare i casi non selezionati, nell'Editor dei dati questi ultimi verranno contrassegnati da una linea diagonale sopra il numero di riga.

Figura 10-11
Criteri di selezione dei casi

*demo.sav [InsiemeDati1] - Data Editor

File Modifica Visualizza Dati Trasforma Analizza Grafici Strumenti Dispositivi aggiuntivi Finestra Aiuto

13 : catimp 1 Visibili: 29 variabile/i di 29

	età	statciv	indirizz	reddito	fascredd	auto	catauto	
1	18	Non spos...	0	13,00	< €25	6,60	Scuola ...	D
2	18	Sposato	0	14,00	< €25	6,80	Scuola ...	D
3	18	Sposato	0	18,00	< €25	8,70	Scuola ...	D
4	18	Non spos...	0	13,00	< €25	6,40	Scuola ...	D
5	18	Non spos...	0	24,00	< €25	12,00	Scuola ...	D
6	18	Sposato	0	14,00	< €25	6,90	Scuola ...	D
7	18	Sposato	0	22,00	< €25	11,10	Scuola ...	D
8	18	Non spos...	0	17,00	< €25	8,60	Scuola ...	S
9	18	Non spos...	0	18,00	< €25	9,10	Scuola ...	D
10	18	Non spos...	0	13,00	< €25	6,60	Scuola ...	D

Visualizzazione dati Visualizzazione variabili

File di esempio

Il file di esempio installato con il prodotto si trova nella sottodirectory *Samples* della directory di installazione. La sottodirectory *Samples* contiene cartelle separate per ciascuna delle seguenti lingue: Inglese, Francese, Tedesco, Italiano, Giapponese, Coreano, Polacco, Russo, Cinese semplificato, Spagnolo e Cinese tradizionale.

Non tutti i file di esempio sono disponibili in tutte le lingue. Se un file di esempio non è disponibile in una lingua, la cartella di tale lingua contiene una versione inglese del file.

Descrizioni

Questa sezione contiene brevi descrizioni dei file di esempio utilizzati negli esempi riportati in tutta la documentazione.

- **accidents.sav.** File di dati ipotetici che prende in esame una compagnia di assicurazioni impegnata nello studio dei fattori di rischio correlati all'età e al sesso per gli incidenti automobilistici che si verificano in una determinata regione. Ciascun caso corrisponde a una classificazione incrociata della categoria relativa età e del sesso.
- **adl.sav.** File di dati ipotetici che prende in esame l'impegno richiesto per determinare i vantaggi di un tipo di terapia proposto per i pazienti con problemi di cuore. I medici hanno assegnato in modo casuale i pazienti con problemi di cuore di sesso femminile a uno di due gruppi. Al primo gruppo è stata assegnata la terapia fisica standard; al secondo gruppo, un'ulteriore terapia di supporto psicologico. Dopo tre mesi di trattamenti, a ciascuna capacità dei pazienti che consente di riprendere le normali attività giornaliere è stato assegnato un punteggio come variabile ordinale.
- **advert.sav.** File di dati ipotetici che prende in esame l'impegno di un rivenditore al dettaglio che desidera esaminare la relazione tra il denaro speso per la pubblicità e le vendite risultanti. Finora sono stati raccolti i dati delle vendite precedenti e i relativi costi pubblicitari.
- **aflatoxin.sav.** File di dati ipotetici che prende in esame il test di raccolti di mais con presenza di Aflatossina, un veleno la cui concentrazione varia notevolmente nei raccolti. Una macchina per la lavorazione dei cereali ha ricevuto 16 campioni da ciascuno degli otto raccolti di mais e ha misurato i livelli di Aflatossina in parti per miliardo (PPB).
- **anorectic.sav.** Per trovare una sintomatologia standardizzata del comportamento anoressico/bulimico, i ricercatori hanno condotto uno studio basato su 55 adolescenti affetti da disordini alimentari conosciuti. Ogni paziente è stato visitato quattro volte in quattro anni, per un totale di 220 visite. Durante ogni visita, ai pazienti sono stati assegnati punteggi per ciascuno dei 16 sintomi. I punteggi relativi ai sintomi sono assenti per il paziente 71 alla visita 2, il paziente 76 alla visita 2 e il paziente 47 alla visita 3, con 217 osservazioni valide.
- **bankloan.sav.** File di dati ipotetici che prende in esame l'impegno di una banca nel tentativo di ridurre il tasso di inadempienza nel rimborso di un prestito. Il file contiene informazioni finanziarie e demografiche su 850 vecchi e potenziali clienti. I primi 700 casi riguardano i

clienti a cui sono stati concessi dei prestiti precedentemente. Gli ultimi 150 casi riguardano i potenziali clienti che la banca deve classificare come rischi di credito positivi o negativi.

- **bankloan_binning.sav.** File di dati ipotetici che contiene informazioni finanziarie e demografiche su 5000 vecchi clienti.
- **behavior.sav.** In un classico esempio , è stato chiesto a 52 studenti di classificare una combinazione di 15 situazioni e 15 comportamenti utilizzando una scala da 0=“molto appropriato” a 9=“molto inadeguato”. I valori medi riferiti ai partecipanti sono stati considerati dissimilarità.
- **behavior_ini.sav.** Questo file di dati contiene la configurazione iniziale di una soluzione a due dimensioni per *behavior.sav*.
- **brakes.sav.** File di dati ipotetici che prende in esame il controllo di qualità di un'industria che produce freni a disco per automobili con elevate prestazioni. Il file di dati contiene le misurazioni del diametro di 16 dischi da ciascuna delle otto macchine di produzione. L'obiettivo finale è ottenere un diametro dei dischi pari a 322 millimetri.
- **breakfast.sav.** In uno studio classico , è stato chiesto a 21 studenti MBA della Wharton School e ai loro consorti di classificare 15 cibi da colazione in ordine di preferenza, dove il valore 1 corrispondeva all'alimento preferito in assoluto e il valore 15 a quello meno preferito. Le loro preferenze sono state registrate per sei diversi scenari, che comprendevano tutti gli scenari compresi tra “Preferenza generale” e “Solo snack con bibita”.
- **breakfast-overall.sav.** Questo file contiene le preferenze degli alimenti della colazione solo per il primo scenario, “Preferenza generale”.
- **broadband_1.sav.** File di dati ipotetici che contiene il numero di sottoscrittori, per area, di un provider di servizi a banda larga nazionale. Il file di dati contiene il numero dei sottoscrittori mensili di 85 aree in un periodo di quattro anni.
- **broadband_2.sav.** Questo file è identico al file *broadband_1.sav*, ma contiene i dati per ulteriori tre mesi.
- **car_insurance_claims.sav.** Un insieme di dati presentato e analizzato altrove riguarda le richieste di risarcimento auto. La quantità media di richieste di risarcimento può essere adattata come avente una distribuzione gamma, utilizzando una funzione di collegamento inverso per correlare la media della variabile dipendente a una combinazione lineare di età del contraente della polizza e tipo e anni del veicolo. Il numero delle richieste di risarcimento specificato può essere utilizzato come peso scalato.
- **car_sales.sav.** Questo file di dati ipotetici contiene le stime sulle vendite, i prezzi di listino e le specifiche fisiche di numerose marche e modelli di veicoli. I prezzi di listino e le specifiche fisiche sono state ottenute dal sito *edmunds.com* e dai siti dei produttori.
- **car_sales_uprepared.sav.** Questa è una versione modificata di *car_sales.sav* che non comprende versioni trasformate dei campi.
- **carpet.sav.** Come esempio tipico , un'azienda interessata alla commercializzazione di un nuovo battitappeto desidera esaminare l'influenza di cinque fattori sulle preferenze del consumatore, ovvero design della confezione, marca, prezzo, la presenza di un *marchio di qualità* e una garanzia “Soddisfatti o rimborsati”. Esistono tre livelli di fattore per il design della confezione, che differiscono per la posizione della spazzola dell'applicatore; tre marchi (*K2R*, *Glory* e *Bissell*); tre livelli di prezzo e due livelli (no o sì) per ciascuno degli ultimi due fattori. Dieci consumatori sono classificati in 22 profili definiti da questi fattori. La variabile

Preferenza include il rango delle classificazioni medie per ogni profilo. Classificazioni basse corrispondono a una preferenza elevata. La variabile riflette una misura globale della preferenza per ogni profilo.

- **carpet_prefs.sav.** Questo file di dati si basa sullo stesso esempio del file *carpet.sav*, ma contiene le classificazioni effettive raccolte da ciascuno dei 10 clienti. Ai clienti è stato chiesto di classificare 22 profili di prodotti in ordine di preferenza. Le variabili da *PREF1* a *PREF22* contengono gli ID dei profili associati, come definito nel file *carpet_plan.sav*.
- **catalog.sav.** File di dati ipotetico che contiene le cifre sulle vendite mensili di tre prodotti venduti da una società di vendita per corrispondenza. Il file include anche i dati di cinque possibili variabili predittore.
- **catalog_seasfac.sav.** Questo file di dati è uguale al file *catalog.sav* con l'eccezione che contiene un insieme di fattori stagionali calcolati dalla procedura Decomposizionale stagionale insieme a variabili di dati.
- **cellular.sav.** File di dati ipotetici che prende in esame l'impegno di un'azienda di telefonia cellulare nel tentativo di ridurre il churn, ovvero l'abbandono dei clienti. Agli account vengono applicati i punteggi relativi alla propensione al churn, con valori compresi tra 0 e 100. Gli account con punteggio pari a 50 o superiore è probabile che stiano cercando nuovi provider.
- **ceramics.sav.** File di dati ipotetici che prende in esame l'impegno di un produttore che desidera stabilire se una nuova lega premium ha una maggiore resistenza al calore rispetto alla lega standard. Ciascun caso rappresenta il test separato di una delle leghe. È indicata la temperatura massima alla quale può essere sottoposto il cuscinetto.
- **cereal.sav.** File di dati ipotetici che prende in esame le preferenze relative agli alimenti della colazione di un campione di 880 persone. Il file riporta anche l'età, il sesso e lo stato civile del campione e se le persone conducono uno stile di vita attivo (in base a un'attività sportiva con frequenza di due volte alla settimana). Ogni caso rappresenta un rispondente separato.
- **clothing_defects.sav.** File di dati ipotetici che prende in esame il processo di controllo di qualità di un'industria di abbigliamento. Per ciascun lotto prodotto nella fabbrica, gli ispettori prelevano un campione di abiti per contare il numero dei capi che non sono accettabili per la vendita.
- **coffee.sav.** Questo file di dati contiene informazioni sulle immagini percepite di sei marche di caffè freddo. Per ciascuno dei 23 attributi dell'immagine del caffè freddo, sono state selezionate tutte le marche descritte da tale attributo. Le sei marche sono indicate dalle sigle AA, BB, CC, DD, EE e FF per tutelare la confidenzialità dei dati.
- **contacts.sav.** File di dati ipotetici che prende in esame l'elenco dei contatti di un gruppo di rappresentanti di vendita di computer aziendali. Ciascun contatto è classificato in base al reparto della società in cui lavora e dalle relative categorie aziendali. Il file riporta anche l'importo dell'ultima vendita effettuata, il tempo trascorso dall'ultima vendita e le dimensioni della società del contatto.
- **creditpromo.sav.** File di dati ipotetici che prende in esame l'impegno di un grande magazzino nel tentativo di valutare l'efficacia di una recente promozione con carta di credito. A tale scopo, sono stati selezionati 500 titolari di carta in modo casuale. Alla metà di questi è stato inviato un annuncio promozionale che comunica la riduzione del tasso d'interesse nel caso di acquisti effettuati entro i tre mesi successivi. All'altra metà è stato inviato un annuncio stagionale standard.

- **customer_dbase.sav.** File di dati ipotetico che prende in esame l'impegno di una società nel tentativo di utilizzare le informazioni contenute nel proprio database dei dati per creare offerte speciali per i clienti che più probabilmente risponderanno all'offerta. È stato selezionato in modo casuale un sottoinsieme della base dei clienti a cui è stata inviata l'offerta speciale e sono state registrate le risposte ricevute.
- **customer_information.sav.** File di dati ipotetici contenente le informazioni postali del cliente, ad esempio il nome e l'indirizzo.
- **customer_subset.sav.** Un sottoinsieme di 80 casi da *customer_dbase.sav*.
- **debate.sav.** File di dati ipotetici che prende in esame le risposte appaite a un'indagine da parte dei partecipanti a un dibattito politico prima e dopo il dibattito. Ogni caso rappresenta un rispondente separato.
- **debate_aggregate.sav.** File di dati ipotetici che aggrega le risposte contenute nel file *debate.sav*. Ciascun caso corrisponde a una classificazione incrociata della preferenza prima e dopo il dibattito.
- **demo.sav.** File di dati ipotetici che prende in esame un database di clienti che hanno fatto acquisti al fine di inviare offerte mensili tramite il metodo del direct mailing. Viene registrata la risposta dei clienti, sia che abbiano aderito all'offerta o meno, insieme a diverse informazioni demografiche.
- **demo_cs_1.sav.** File di dati ipotetici che prende in esame il primo passo che una società intraprende per compilare un database con informazioni ricavate dai sondaggi. Ogni caso rappresenta una diversa città. Sono registrate anche le informazioni sulla regione, provincia, distretto e città.
- **demo_cs_2.sav.** File di dati ipotetici che prende in esame il secondo passo che una società intraprende per compilare un database con informazioni ricavate dai sondaggi. Ogni caso rappresenta una diversa unità di abitazione, ricavata dalle città selezionate nel primo passo. Sono registrate anche le informazioni sulla regione, provincia, distretto, città, suddivisione e unità. Il file include inoltre informazioni sul campionamento ottenute dai primi due stadi del disegno.
- **demo_cs.sav.** File di dati ipotetici che contiene informazioni sulle indagini raccolte utilizzando un disegno di campionamento complesso. Ogni caso rappresenta una diversa unità di abitazione. Sono registrate diverse informazioni demografiche e sul campionamento.
- **dmdata.sav.** File di dati ipotetici che contiene informazioni demografiche e di acquisto di una società di direct marketing. *dmdata2.sav* contiene informazioni su un sottoinsieme di contatti che hanno ricevuto un mailing di prova e *dmdata3.sav* contiene informazioni sui contatti rimanenti che non hanno ricevuto il mailing di prova.
- **dietstudy.sav.** File di dati ipotetici che contiene il risultato di uno studio ipotetico sulla dieta chiamato "Stillman diet". Ogni caso rappresenta un diverso soggetto e ne riporta il peso prima e dopo la dieta in libbre e i livelli dei trigliceridi in mg/100 ml.
- **dvdplayer.sav.** File di dati ipotetici che prende in esame lo sviluppo di un nuovo lettore DVD. Utilizzando un prototipo, il personale addetto al marketing ha raccolto dati sui gruppi di interesse. Ogni caso rappresenta un diverso utente che è stato sottoposto all'indagine e include informazioni demografiche personali dell'utente e sulle risposte che ha fornito riguardo al prototipo.

- **german_credit.sav.** Questo file di dati contiene informazioni ricavate dall'insieme di dati "German Credit" del Repository of Machine Learning Databases presso la University of California, Irvine.
- **grocery_1month.sav.** Questo file di dati ipotetici corrisponde al file di dati *grocery_coupons.sav* con gli acquisti settimanali organizzati in modo che ogni caso corrisponda a un cliente separato. Alcune delle variabili che cambiano settimanalmente non vengono riportate nei risultati; l'importo speso registrato corrisponde ora alla somma degli importi spesi durante le quattro settimane dello studio.
- **grocery_coupons.sav.** File di dati ipotetici che contiene i dati sui sondaggi raccolti da una catena di drogherie interessata alle abitudini di acquisto dei suoi clienti. Ciascun cliente viene seguito per quattro settimane e ciascun caso corrisponde a una settimana per cliente con informazioni sul luogo degli acquisti e i tipi di acquisti, incluso l'importo speso nelle drogherie durante la settimana.
- **guttman.sav.** Bell ha presentato una tabella per illustrare i possibili gruppi sociali. Guttman ha utilizzato una parte di tale tabella, in cui cinque variabili che descrivono elementi come l'interazione sociale, i sentimenti di appartenenza a un gruppo, la vicinanza fisica dei membri e il grado di formalità della relazione, sono state incrociate con cinque gruppi sociali teorici, compresi folla (ad esempio, le persone presenti a una partita di calcio), uditorio (ad esempio, di uno spettacolo teatrale o di una lezione universitaria), pubblico (ad esempio televisivo), calca (come una folla, ma con un'interazione molto maggiore), gruppi primari (intimi), gruppi secondari (volontari) e la comunità moderna (unione non stretta derivante da una vicinanza fisica elevata e dall'esigenza di servizi specializzati).
- **health_funding.sav.** File di dati ipotetici che contiene i dati sui fondi di assistenza sanitaria (importo per 100 persone), sui tassi di malattie (tasso per 10.000 persone) e sulle visite ai fornitori di assistenza sanitaria (tasso per 10.000 persone). Ogni caso rappresenta una diversa città.
- **hivassay.sav.** File di dati ipotetici che prende in esame l'impegno di un'industria farmaceutica nel tentativo di sviluppare un'analisi che riesca a rilevare in tempi brevi l'infezione da virus HIV. I risultati dell'analisi sono otto sfumature di colore rosso sempre più intenso; le sfumature più intense indicano la maggiore probabilità di infezione. Un esperimento di laboratorio è stato condotto su 2000 campioni di sangue. La metà di questi è risultata infetta al virus HIV, l'altra metà non è risultata infetta.
- **hourlywagedata.sav.** File di dati ipotetici che prende in esame la paga oraria degli infermieri occupati presso uffici e ospedali e in base ai diversi livelli di esperienza.
- **insurance_claims.sav.** File di dati ipotetici che prende in esame una compagnia di assicurazioni impegnata nella creazione di un modello per contrassegnare le richieste di risarcimento sospette e potenzialmente fraudolente. Ogni caso rappresenta una richiesta di risarcimento separata.
- **insure.sav.** File di dati ipotetici che prende in esame una compagnia di assicurazioni impegnata nello studio dei fattori di rischio, che indicano l'eventualità che un cliente presenti una domanda di indennizzo in un contratto assicurativo sulla vita della durata di dieci anni. Ogni caso nel file di dati rappresenta una coppia di contratti. In un contratto sono contenute informazioni su una richiesta di risarcimento, l'altro sull'età e sul sesso.

- **judges.sav.** File di dati ipotetici che prende in esame il punteggio assegnato, da giurie qualificate (più un appassionato) a 300 prestazioni sportive. Ciascuna riga rappresenta una diversa prestazione; i giudici hanno esaminato le stesse prestazioni.
- **kinship_dat.sav.** Rosenberg e Kim si prefiggono di analizzare 15 termini indicanti parentela (zia, fratello, cugino, padre, nipote femmina, di nonni, nonno, nonna, nipote maschio di nonni, madre, nipote maschio di zii), nipote femmina di zii, sorella, figlio, zio). Hanno richiesto a quattro gruppi di studenti universitari (due composti da femmine e due da maschi) di ordinare questi termini in base alla similarità. A due gruppi (uno femminile e uno maschile) è stato richiesto di effettuare l'ordinamento due volte, con il secondo ordinamento basato su un criterio diverso rispetto al primo. Di conseguenza, sono state ottenute sei "sorgenti" in totale. Ogni sorgente corrisponde a una matrice di prossimità 15×15 , le cui celle sono uguali al numero delle persone in una sorgente meno il numero di volte in cui gli oggetti sono stati ripartiti insieme nella sorgente.
- **kinship_ini.sav.** Questo file di dati contiene la configurazione iniziale di una soluzione a tre dimensioni per *kinship_dat.sav*.
- **kinship_var.sav.** Questo file di dati contiene variabili indipendenti relative a *sexo*, *generazione* e *grado* di separazione che possono essere utilizzate per interpretare le dimensioni di una soluzione per *kinship_dat.sav*. In modo specifico, tali variabili possono essere utilizzate per limitare lo spazio della soluzione a una combinazione lineare di tali variabili.
- **marketvalues.sav.** File di dati che prende in esame le vendite di abitazioni in un nuovo centro abitato in Algonquin, Ill., durate gli anni 1999–2000. Tali vendite sono una questione di dominio pubblico.
- **nhis2000_subset.sav.** Il National Health Interview Survey (NHIS) è un sondaggio di grandi dimensioni condotto sulla popolazione civile americana. Le interviste vengono realizzate di persona e si basano su un campione rappresentativo di famiglie a livello nazionale. Per ogni membro di una famiglia vengono raccolte osservazioni e informazioni di carattere demografico relative allo stato di salute. Questo file di dati contiene un sottoinsieme delle informazioni ottenute dall'indagine del 2000. National Center for Health Statistics. National Health Interview Survey, 2000. File di dati e documentazione di dominio pubblico. ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHIS/2000/. Accesso 2003.
- **ozone.sav** I dati includono 330 osservazioni basate su sei variabili meteorologiche per quantificare la concentrazione dell'ozono dalle variabili rimanenti. I precedenti ricercatori, e , hanno rilevato non linearità tra queste variabili, che impediscono un approccio di regressione standard.
- **pain_medication.sav.** File di dati ipotetici che contiene i risultati di un test clinico per stabilire la cura antinfiammatoria per il trattamento del dolore generato dall'artrite cronica. Di particolare interesse, il test ha evidenziato il tempo che impiega il farmaco ad avere effetto e il confronto con altri farmaci esistenti.
- **patient_los.sav.** File di dati ipotetici che contiene informazioni sul trattamento dei pazienti ricoverati per sospetto di infarto del miocardio. Ogni caso corrisponde a un diverso paziente e contiene diverse variabili correlate alla degenza nell'ospedale.
- **patlos_sample.sav.** File di dati ipotetici che contiene informazioni sul trattamento di un campione di pazienti curato con trombolitici durante la degenza per infarto del miocardio. Ogni caso corrisponde a un diverso paziente e contiene diverse variabili correlate alla degenza nell'ospedale.

- **poll_cs.sav.** File di dati ipotetici che prende in esame i sondaggi per stabilire il livello di sostegno pubblico nei confronti di un disegno di legge prima che diventi una legge vera e propria. I casi corrispondono ai votanti registrati. Ciascun caso riporta informazioni sulla contea, sul comune e sul quartiere in cui vive il votante.
- **poll_cs_sample.sav.** File di dati ipotetici che contiene un campione dei votanti elencati nel file *poll_cs.sav*. Il campione è stato selezionato in base al disegno specificato nel file di piano *poll.csplan* e questo file di dati contiene le probabilità di inclusione e i pesi del campione. Tuttavia, notare che poiché fa uso del metodo PPS (probability-proportional-to-size, probabilità proporzionale alla dimensione), esiste anche un file contenente le probabilità di selezione congiunte (*poll_jointprob.sav*). Le ulteriori variabili corrispondenti ai dati demografici dei votanti e alla loro opinione sul disegno di legge, sono state raccolte e aggiunte al file di dati dopo aver acquisito il campione.
- **property_assess.sav.** File di dati ipotetici che prende in esame l'impegno di un perito di una contea nel tentativo di mantenere gli accertamenti sui valori delle proprietà aggiornati in base alle risorse limitate. I casi rappresentano le proprietà vendute nella contea nello scorso anno. Ogni caso nel file di dati contiene informazioni sul comune in cui si trova la proprietà, il perito che per ultimo ha visitato la proprietà, il tempo trascorso dall'accertamento, la valutazione fatta in tale momento e il valore di vendita della proprietà.
- **property_assess_cs.sav.** File di dati ipotetici che prende in esame l'impegno di un perito di uno stato nel tentativo di mantenere aggiornati gli accertamenti sui valori delle proprietà in base alle risorse limitate. I casi corrispondono alle proprietà nello stato. Ogni caso nel file di dati include informazioni sulla contea, il comune e il quartiere in cui risiede la proprietà, la data dell'ultimo accertamento e la valutazione fatta in tale data.
- **property_assess_cs_sample.sav.** File di dati ipotetici che contiene un campione delle proprietà elencate nel file *property_assess_cs.sav*. Il campione è stato selezionato in base al disegno specificato nel file di piano *property_assess.csplan* e questo file di dati contiene le probabilità di inclusione e i pesi del campione. L'ulteriore variabile *Valore corrente* è stata raccolta e aggiunta al file di dati dopo aver acquisito il campione.
- **recidivism.sav.** File di dati ipotetici che prende in esame l'impegno delle Forze dell'Ordine nel tentativo di valutare il tasso di recidività nella propria area di giurisdizione. Ogni caso corrisponde a un precedente trasgressore e include le informazioni demografiche, alcuni dettagli sul primo crimine, il tempo trascorso fino al secondo arresto e se tale arresto è avvenuto entro due anni dal primo.
- **recidivism_cs_sample.sav.** File di dati ipotetici che prende in esame l'impegno delle Forze dell'Ordine nel tentativo di valutare il tasso di recidività nella propria area di giurisdizione. Ogni caso corrisponde a un trasgressore precedente, rilasciato dopo il primo arresto durante il mese di giugno del 2003 e registra le relative informazioni demografiche, alcuni dettagli sul primo crimine commesso e i dati del secondo arresto, se si è verificato prima della fine di giugno del 2006. I trasgressori sono stati selezionati dai dipartimenti sottoposti a campione in base al piano di campionamento specificato nel file *recidivism_cs.csplan*. Poiché viene utilizzato un metodo PPS (Probability-Proportional-to-Size, probabilità proporzionale alla dimensione), esiste anche un file contenente le probabilità di selezione congiunte (*recidivism_cs_jointprob.sav*).
- **rfm_transactions.sav.** File di dati ipotetici contenente i dati delle transazioni di acquisto, inclusa la data di acquisto, gli articoli acquistati e il valore monetario di ciascuna transazione.

- **salesperformance.sav.** File di dati ipotetici che prende in esame la valutazione di due nuovi corsi di formazione alle vendite. Sessanta dipendenti, divisi in tre gruppi, ricevono tutti la formazione standard. In più, al gruppo 2 viene assegnato un corso di formazione tecnica e al gruppo 3 un'esercitazione pratica. Alla fine del corso di formazione, ciascun dipendente viene sottoposto a un esame e il punteggio conseguito viene registrato. Ciascun caso nel file di dati rappresenta un diverso partecipante. Il file di dati include il gruppo a cui è assegnato il partecipante e il punteggio conseguito all'esame finale.
- **satisf.sav.** File di dati ipotetico che prende in esame un'indagine sulla soddisfazione dei clienti condotta da una società di vendita al dettaglio presso 4 negozi. Sono stati intervistati 582 clienti e ciascun caso rappresenta le risposte ottenute da un singolo cliente.
- **screws.sav.** Questo file di dati contiene informazioni sulle caratteristiche di viti, bulloni, dadi e puntine .
- **shampoo_ph.sav.** File di dati ipotetici che prende in esame il processo di controllo di qualità di un'industria di prodotti per capelli. A intervalli di tempo regolari, vengono misurati sei diversi lotti prodotti e ne viene registrato il relativo pH. I valori accettati sono compresi tra 4,5 e 5,5.
- **ships.sav.** Ad esempio, un insieme di dati presentato e analizzato altrove riguarda i danni subiti dalle navi da carico a causa delle onde. I conteggi degli incidenti possono essere presentati con un tasso di Poisson in base al tipo di nave, al periodo di costruzione e al periodo di servizio. I mesi di servizio aggregati di ciascuna cella della tabella generata dalla classificazione incrociata dei fattori fornisce i valori di esposizione al rischio.
- **site.sav.** File di dati ipotetici che prende in esame l'impegno di una società nella scelta di nuovi siti in cui espandere la propria presenza. La società ha incaricato due consulenti separati che, oltre a valutare i siti e presentare un report completo, devono classificarli come potenzialmente "molto adatti", "adatti" o "poco adatti".
- **smokers.sav.** Questo file di dati è un estratto del 1998 National Household Survey of Drug Abuse e rappresenta un campione probabile di famiglie americane. (<http://dx.doi.org/10.3886/ICPSR02934>) Il primo passo nell'analisi di questo file di dati consiste quindi nel pesare i dati per rispecchiare le tendenze della popolazione.
- **stocks.sav** Questo file di dati ipotetici contiene i prezzi e i volumi delle scorte riferiti a un anno.
- **stroke_clean.sav.** File di dati ipotetici che riporta lo stato di un database medico dopo averne eseguito la pulizia utilizzando le procedure del modulo Data Preparation.
- **stroke_invalid.sav.** File di dati ipotetici che riporta lo stato iniziale di un database medico e contiene numerosi errori di immissione dati.
- **stroke_survival.** Questo file di dati ipotetici riguarda i tempi di sopravvivenza per i pazienti che, dopo avere completato un programma riabilitativo in seguito a un ictus postischemico, affrontano alcune sfide. Dopo l'attacco, viene annotata l'occorrenza dell'infarto miocardico, dell'ictus ischemico o emorragico e viene registrata l'ora dell'evento. Questo campione viene troncato a sinistra perché include solo i pazienti che sono sopravvissuti fino alla fine del programma riabilitativo post-ictus.
- **stroke_valid.sav.** File di dati ipotetici che riporta lo stato di un database medico dopo il controllo dei valori eseguito con la procedura Convalida i dati. Il database contiene comunque casi potenzialmente anomali.

- **survey_sample.sav.** File di dati che contiene i dati dell'indagine, compresi i dati demografici e varie misure dell'atteggiamento. Si basa su un sottoinsieme di variabili tratte dal 1998 NORC General Social Survey, benché i valori di alcuni dati siano stati modificati e siano state aggiunte variabili fittizie a scopo dimostrativo.
- **telco.sav.** File di dati ipotetici che prende in esame l'impegno di un'azienda di telecomunicazioni nel tentativo di ridurre il churn, ovvero l'abbandono dei propri clienti. Ciascun caso rappresenta un cliente separato e riporta diverse informazioni demografiche e sull'uso del servizio.
- **telco_extra.sav.** Questo file di dati è simile al file *telco.sav*, ma le variabili "tenure" e spesa del cliente trasformata tramite logaritmo sono state sostituite dalle variabili di spesa del cliente trasformata tramite logaritmo standardizzate.
- **telco_missing.sav.** Questo file di dati è un sottoinsieme del file di dati *telco.sav*, ma alcuni dei valori di dati demografici sono stati sostituiti con valori mancanti.
- **testmarket.sav.** File di dati ipotetici che prende in esame i piani di una catena di fast food per aggiungere un nuovo prodotto al proprio menu. Sono previste tre campagne promozionali del nuovo prodotto. Il prodotto viene introdotto in diversi mercati selezionati in modo casuale. Per ogni sede viene utilizzata una promozione differente registrando le vendite settimanali della nuova voce per le prime quattro settimane. Ogni caso rappresenta un luogo e una settimana diversi.
- **testmarket_1month.sav.** Questo file di dati ipotetici corrisponde al file *testmarket.sav* con le vendite settimanali organizzate in modo che ogni caso corrisponda a un luogo separato. Alcune delle variabili che cambiano settimanalmente non vengono riportate nei risultati; le vendite registrate corrispondono ora alla somma delle vendite conseguite durante le quattro settimane dello studio.
- **tree_car.sav.** File di dati ipotetici che contiene dati demografici e sul prezzo di acquisto dei veicoli.
- **tree_credit.sav.** File di dati ipotetici che contiene dati demografici e sulla cronologia dei mutui di una banca.
- **tree_missing_data.sav.** File di dati ipotetici che contiene dati demografici e sulla cronologia dei mutui di una banca con un numero elevato di valori mancanti.
- **tree_score_car.sav.** File di dati ipotetici che contiene dati demografici e sul prezzo di acquisto dei veicoli.
- **tree_textdata.sav.** File di dati semplice con due variabili destinato principalmente per mostrare lo stato predefinito delle variabili prima dell'assegnazione dei livelli di misurazione e delle etichette dei valori.
- **tv-survey.sav.** File di dati ipotetici che prende in esame un sondaggio condotto da una emittente televisiva che deve stabilire se estendere la durata di un programma di successo. A un campione di 906 intervistati è stato chiesto se preferisce guardare il programma con diverse condizioni. Ciascuna riga rappresenta un diverso intervistato e ciascuna colonna una diversa condizione.
- **ulcer_recurrence.sav.** Questo file contiene informazioni parziali su uno studio svolto per mettere a confronto l'efficacia di due terapie preventive per la recidiva delle ulcere. Fornisce un ottimo esempio di dati acquisiti a intervalli ed è stato presentato e analizzato in altri luoghi .

- **ulcer_recurrence_recoded.sav.** In questo file sono contenute le informazioni del file *ulcer_recurrence.sav* riorganizzate per consentire di presentare la probabilità degli eventi per ciascun intervallo dello studio, anziché solo alla fine. È stato presentato e analizzato in altri luoghi .
- **verd1985.sav.** Questo file di dati prende in esame un'indagine . Sono state registrate le risposte di quindici soggetti a otto variabili. Le variabili di interesse sono suddivise in tre insiemi. L'insieme 1 include *età* e *statociv*, l'insieme 2 include *andom* e *giornale* e l'insieme 3 include *musica* e *vicinato*. *Andom* viene scalata come nominale multipla ed *età* come ordinale; tutte le altre variabili vengono scalate come nominali singole.
- **virus.sav.** File di dati ipotetici che prende in esame l'impegno di un ISP (Internet Service Provider) nel tentativo di determinare gli effetti che un virus può generare nelle sue reti. Si è tenuta traccia della percentuale (approssimativa) di traffico e-mail infettato da virus sulla rete in un lasso di tempo, dal momento dell'individuazione fino alla soppressione della minaccia.
- **wheeze_steubenville.sav.** Questo file è un sottoinsieme di uno studio longitudinale degli effetti che l'inquinamento provoca sulla salute dei bambini . I dati contengono misure binarie ripetute del livello di asma dei bambini della città di Steubenville, Ohio, di 7, 8, 9 e 10 anni. I dati indicano anche se la mamma dei bambini era fumatrice durante il primo anno dello studio.
- **workprog.sav.** File di dati ipotetici che prende in esame un programma di lavoro governativo il cui obiettivo è fornire attività più adatte alle persone diversamente abili. È stato seguito un campione di potenziali partecipanti al programma, alcuni dei quali sono stati selezionati in modo casuale e altri no. Ogni caso rappresenta un diverso partecipante al programma.
- **worldsales.sav** Questo file di dati ipotetici contiene i ricavi suddivisi per continenti e prodotti.

Note legali

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Marchi commerciali

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

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.

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.

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.

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

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

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

Questo prodotto utilizza WinWrap Basic, Copyright 1993-2007, Polar Engineering and Consulting, <http://www.winwrap.com>.

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

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

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



- Access (Microsoft), 13
- attributi di variabili
 - riutilizzare, 41
- calcolo di nuove variabili, 129
- casi
 - ordinamento, 143, 146
 - selezione, 146
- conteggi
 - tabelle di conteggi, 55
- copiare gli attributi delle variabili, 41
- creare etichette delle variabili, 32
- Creazione guidata di query, 13

- dati categoriali, 55
 - misure riassuntive, 55
- dati continui, 55
- dati di intervallo, 55
- dati di rapporto, 55
- dati di scala, 55
- dati nominali, 55
- dati numerici, 27
- dati ordinali, 55
- dati qualitativi, 55
- dati quantitativi, 55
- dati stringa
 - inserimento di dati, 30
- distinzione per gruppi, 144

- Editor dei dati
 - file di dati multipli aperti, 51
 - inserimento di dati non numerici, 30
 - inserimento di dati numerici, 27
- esportazione dei risultati
 - a Excel, 105
 - a Word, 105
 - HTML, 115
 - in PowerPoint, 105
- espressioni logiche, 133
- etichette dei valori
 - assegnazione, 33, 35
 - controllo visualizzazione nel Viewer, 33, 35
 - variabili non numeriche, 35
 - variabili numeriche, 33
- etichette di variabili
 - creazione, 32
- Excel (Microsoft)
 - esportazione dei risultati, 105

- file di database
 - lettura, 13
- file di dati
 - file di dati multipli aperti, 51
- file di dati multipli aperti, 51
 - soppressione, 54

- file di dati testuali
 - lettura, 19
- file di esempio
 - posizione, 152
- file di Excel
 - lettura, 11
- file di fogli elettronici
 - lettura, 11
 - lettura dei nomi di variabili, 11
- file di sintassi
 - apertura, 119
- finestre di sintassi
 - autocompletamento, 117
 - codifica colori, 119
 - esecuzione di comandi, 117, 119
 - incollare i comandi, 116
 - modifica di comandi, 117
 - punti di interruzione, 120
- funzioni nelle espressioni, 131

- grafici
 - barra, 57, 62
 - creazione di grafici, 62
 - grafici: opzioni, 81
 - istogrammi, 60
 - modelli, 76
 - modifica dei grafici, 68
 - modifica di grafici, 68
 - grafici a barre, 57

- HTML
 - esportazione dei risultati, 115

- Importazione guidata di testo, 19
- Impostazione guidata data e ora, 135
- incollare la sintassi
 - da una finestra di dialogo, 116
- inserimento di dati, 27, 30
 - non numerici, 30
 - numerico, 27
- insiemi di dati
 - ridenominazione, 53
- istogrammi, 60

- livello di misurazione, 55

- marchi commerciali, 163
- misure riassuntive
 - dati categoriali, 55
 - variabili di scala, 58
- modifica di tabelle pivot, 90

- nascondere righe e colonne in tabelle pivot, 91
- note legali, 162

- ordinamento dei casi, 143
- PowerPoint (Microsoft)
 - esportazione dei risultati, 105
- ricodifica di valori, 123
- ridenominazione degli insiemi di dati, 53
- selezione di casi, 146
- sintassi, 116
- sottoinsiemi di casi
 - basati su date e orari, 149
 - campione casuale, 148
 - eliminazione di casi non selezionati, 150
 - espressioni logiche, 147
 - filtro di casi non selezionati, 150
 - se la condizione è soddisfatta, 147
 - selezione, 146
- sovrapposte, 27
 - etichette, 32
 - tipi di dati, 33
- spostamento
 - elementi del Viewer, 85
 - elementi delle tabelle pivot, 87
- strati
 - creazione in tabelle pivot, 89
- strumento Aiuto sintassi, 119
- tabella di frequenza, 55
- tabelle pivot
 - accesso alle definizioni, 87
 - formati di celle, 92
 - formattazione, 90
 - modifica, 90
 - nascondere punti decimali, 92
 - nascondere righe e colonne, 91
 - pivoting, 87
 - strati, 89
 - tipi di dati cella, 92
 - trasposizione di righe e colonne, 87
- tipi di dati
 - per variabili, 33
- trasposizione (capovolgimento) di righe e colonne nelle tabelle pivot, 87
- valori mancanti
 - mancanti di sistema, 37
 - per variabili non numeriche, 40
 - per variabili numeriche, 38
- valori mancanti di sistema, 37
- variabili di data e ora, 135
- variabili di scala
 - misure riassuntive, 58
- Viewer
 - nascondere e visualizzare l'output, 85
- spostamento dell'output, 85
- Word (Microsoft)
 - esportazione dei risultati, 105