IBM SPSS Statistics 31 -Kurzübersicht

Hinweis

Vor Verwendung dieser Informationen und des darin beschriebenen Produkts sollten die Informationen unter „Bemerkungen” auf Seite 93 gelesen werden.

Produktinformation

Diese Ausgabe bezieht sich auf Version 31, Release 0, Modifikation 2 von IBM® SPSS Statistik und auf alle nachfolgenden Releases und Modifikationen, bis dieser Hinweis in einer Neuausgabe geändert wird.

Inhaltsverzeichnis

Kapitel 1. Einführung

In diesem Handbuch wird Ihnen der Umgang mit vielen der verfügbaren Funktionen erklärt. Es wurde so konzipiert, dass Ihnen praktische und schrittweise Anleitungen gegeben werden. Alle in den Beispielen gezeigten Dateien werden zusammen mit der Anwendung installiert. Sie können also alles mitverfolgen und dieselben Analysen durchführen bzw. dieselben Ergebnisse erzielen, die hier gezeigt werden.

Detaillierte Beispiele für verschiedene statistische Analyseverfahren finden Sie in den in Einzelschritte unterteilten Fallstudien, die im Hilfemenü verfügbar sind.

Stichprobendateien

Für die meisten hier vorgestellten Beispiele wird die Datendatei *demo.sav* verwendet. Bei dieser Datendatei handelt es sich um eine fiktive Befragung von mehreren tausend Personen, die grundlegende demografische Daten und Verbraucherinformationen enthält.

Wenn Sie die Studentenversion verwenden, enthält Ihre Version von *demo.sav* lediglich einen repräsentativen Teil der ursprünglichen Datendatei, damit die Höchstgrenze von 1500 Fällen eingehalten wird. Die mit dieser Datendatei erzielten Ergebnisse weichen von den hier gezeigten Ergebnissen ab.

Die zusammen mit dem Produkt installierten Beispieldateien finden Sie im Unterverzeichnis *Samples* des Installationsverzeichnisses. Innerhalb des Unterverzeichnisses Samples gibt es für jede der folgenden Sprachen einen eigenen Ordner: Englisch, Französisch, Deutsch, Italienisch, Japanisch, Koreanisch, Polnisch, Vereinfachtes Chinesisch, Spanisch und Traditionelles Chinesisch.

Nicht alle Beispieldateien stehen in allen Sprachen zur Verfügung. Wenn eine Beispieldatei nicht in einer Sprache zur Verfügung steht, enthält der jeweilige Sprachordner eine englische Version der Beispieldatei.

Öffnen einer Datendatei

So öffnen Sie eine Datendatei:

1. Wählen Sie in den Menüs Folgendes aus:

   Datei > **Öffnen** > **Daten...**

   Ein Dialogfeld zum Öffnen von Dateien wird angezeigt.

   In der Standardeinstellung werden IBM SPSS Statistik -Datendateien (Erweiterung *.sav*) angezeigt.

   In diesem Beispiel wird die Datei *demo.sav* verwendet.

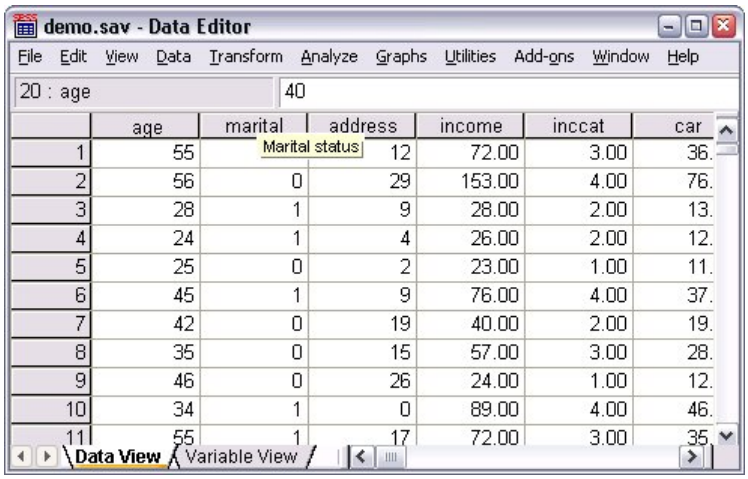


Abbildung 1. Datei demo.sav im Dateneditor

Die Datendatei wird im Dateneditor angezeigt. Wenn Sie in der Datenansicht mit dem Mauszeiger auf einen Variablennamen (die Spaltenüberschrift) zeigen, wird eine ausführliche Variablenbeschriftung angezeigt (sofern für diese Variable eine Beschriftung definiert wurde).

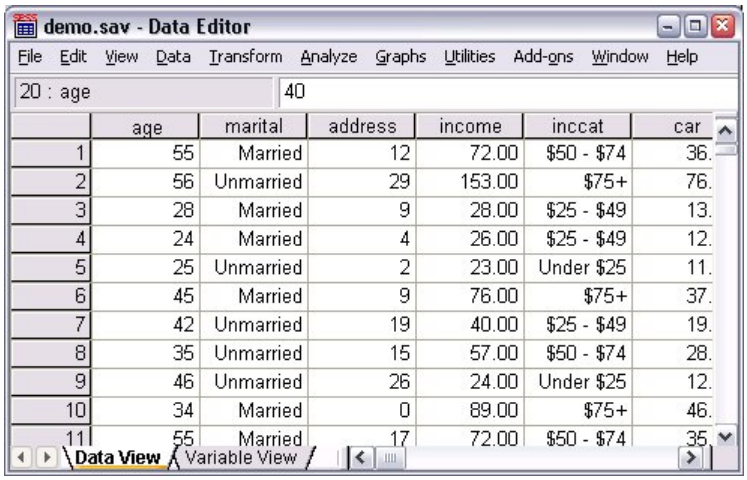
In der Standardeinstellung werden die tatsächlichen Datenwerte angezeigt. So zeigen Sie Beschriftungen an:

2. Wählen Sie in den Menüs Folgendes aus:

Anzeigen > **Wertbeschriftungen**

Abbildung 2. Schaltfläche "Wertbeschriftungen"

Sie können auch die Schaltfläche "Wertbeschriftungen" auf der Symbolleiste verwenden.



	age	marital	address	income	inccat	car
1	55	Married	12	72.00	$50 - $74	36.
2	56	Unmarried	29	153.00	$75+	76.
3	28	Married	9	28.00	$25 - $49	13.
4	24	Married	4	26.00	$25 - $49	12.
5	25	Unmarried	2	23.00	Under $25	11.
6	45	Married	9	76.00	$75+	37.
7	42	Unmarried	19	40.00	$25 - $49	19.
8	35	Unmarried	15	57.00	$50 - $74	28.
9	46	Unmarried	26	24.00	Under $25	12.
10	34	Married	0	89.00	$75+	46.
11	55	Married	17	72.00	$50 - $74	35.

Abbildung 3. Wertbeschriftungen im Dateneditor

Jetzt werden aussagekräftige Wertbeschriftungen angezeigt, um die Interpretation der Antworten zu erleichtern.

Durchführen einer Analyse

Wenn Sie Zusatzoptionen erworben haben, enthält das Menü "Analysieren" eine Liste von Kategorien für die Berichterstellung und die statistische Analyse.

Zunächst soll eine einfache Häufigkeitstabelle erstellt werden. Für dieses Beispiel ist die Option "Statistics Base" erforderlich.

1. Wählen Sie in den Menüs Folgendes aus:

   Analysieren > **Deskriptive Statistiken** > **Häufigkeiten...**

   Das Dialogfeld "Häufigkeiten" wird angezeigt.

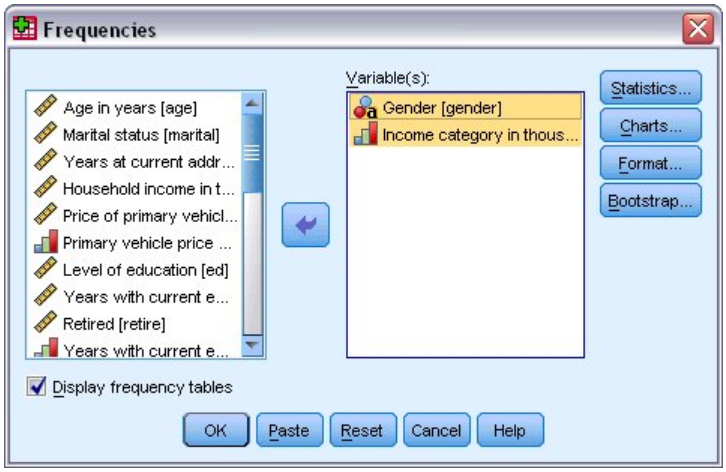


Abbildung 4. Dialogfeld "Häufigkeiten"

Ein Symbol neben jeder Variablen stellt Informationen zum Datentyp und Messniveau bereit.

	Numerisch	**Zeichenfolge**	**Datum**	**Zeit**
Metrisch (stetig)		nicht zutreffend		
Ordinal				
Nominal				

Wenn die Variablenbeschriftung und/oder der Variablenname in der Liste nicht vollständig angezeigt wird, können Sie die vollständige Beschriftung bzw. den vollständigen Namen anzeigen, indem Sie mit dem Cursor darauf zeigen. Der Variablenname *inccat* wird in eckigen Klammern nach der beschreibenden Variablenbeschriftung angezeigt. Die Variablenbeschriftung lautet *Einkommensklassen in Tausend*. Wenn keine Variablenbeschriftung vorhanden ist, wird im Listenfeld nur der Variablenname angezeigt.

Sie können die Größe von Dialogfeldern ebenso ändern wie die von Fenstern, indem Sie auf die äußeren Ränder oder Ecken klicken und daran ziehen. Wenn Sie beispielsweise das Dialogfeld verbreitern, werden auch die Variablenlisten breiter.

Wählen Sie im Dialogfeld die Variablen, die analysiert werden sollen, aus der Quellenvariablenliste auf der linken Seite aus und verschieben Sie sie in die Liste "Variable(n)" auf der rechten Seite. Die Schaltfläche **OK** (mit der die Analyse ausgeführt wird) wird erst aktiviert, wenn mindestens eine Variable in die Liste "Variable(n)" übernommen wurde.

In vielen Dialogfeldern erhalten Sie weitere Informationen, indem Sie mit der rechten Maustaste auf einen beliebigen Variablennamen in der Liste klicken und im Popup-Menü **Variableninformationen** auswählen.

2. Klicken Sie in der Quellenvariablenliste auf *Geschlecht [geschl])* und ziehen Sie die Variable in die Zielliste "Variable(n)".
3. Klicken Sie in der Quellenliste auf *Einkommensklassen in Tausend [eink_kl]* und ziehen Sie die Variable in die Zielliste.

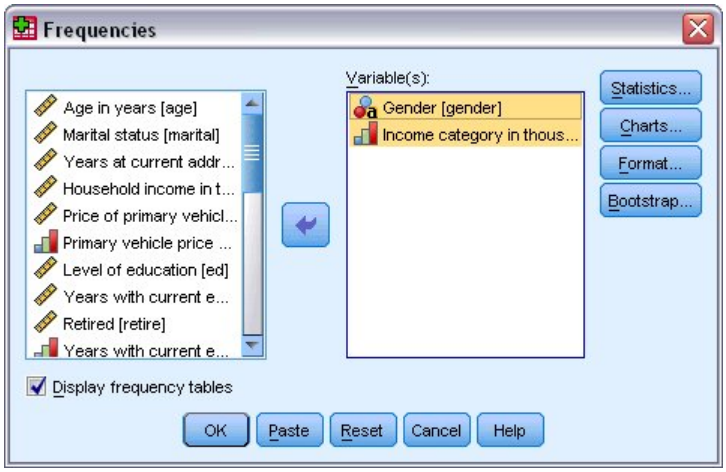


Abbildung 5. Für die Analyse ausgewählte Variablen

4. Klicken Sie auf **OK**, um die Prozedur auszuführen.

Die Ergebnisse werden im Fenster "Viewer" angezeigt.

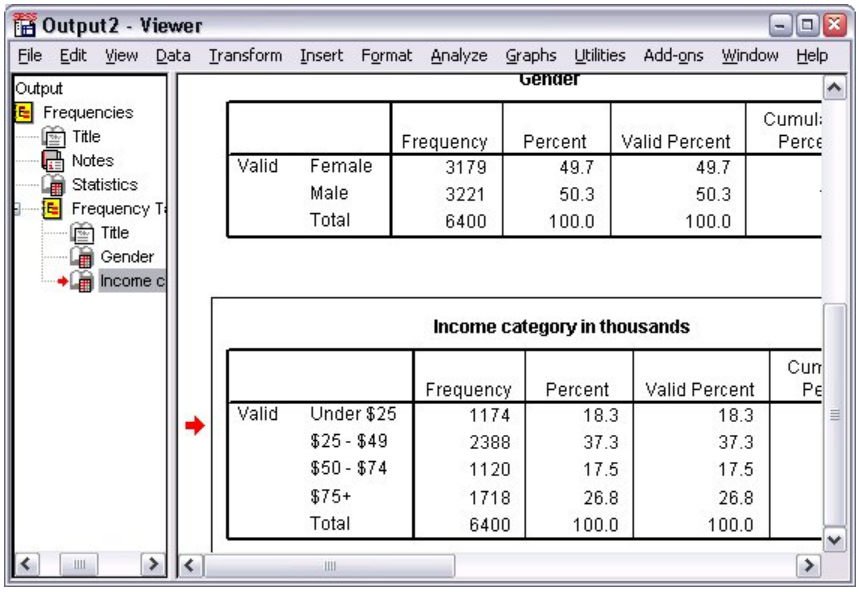


Abbildung 6. Häufigkeitstabelle für Einkommensklassen

Erstellen von Diagrammen

Diagramme können zwar auch mit einigen statistischen Prozeduren erstellt werden, Sie können zum Erstellen von Diagrammen jedoch auch das Menü "Grafik" verwenden.

So können Sie beispielsweise eine Grafik erstellen, in der die Beziehung zwischen dem Besitz von schnurlosen Telefonen und Palm Pilots dargestellt wird.

1. Wählen Sie in den Menüs Folgendes aus:

   Diagramme > **Diagrammerstellung...**

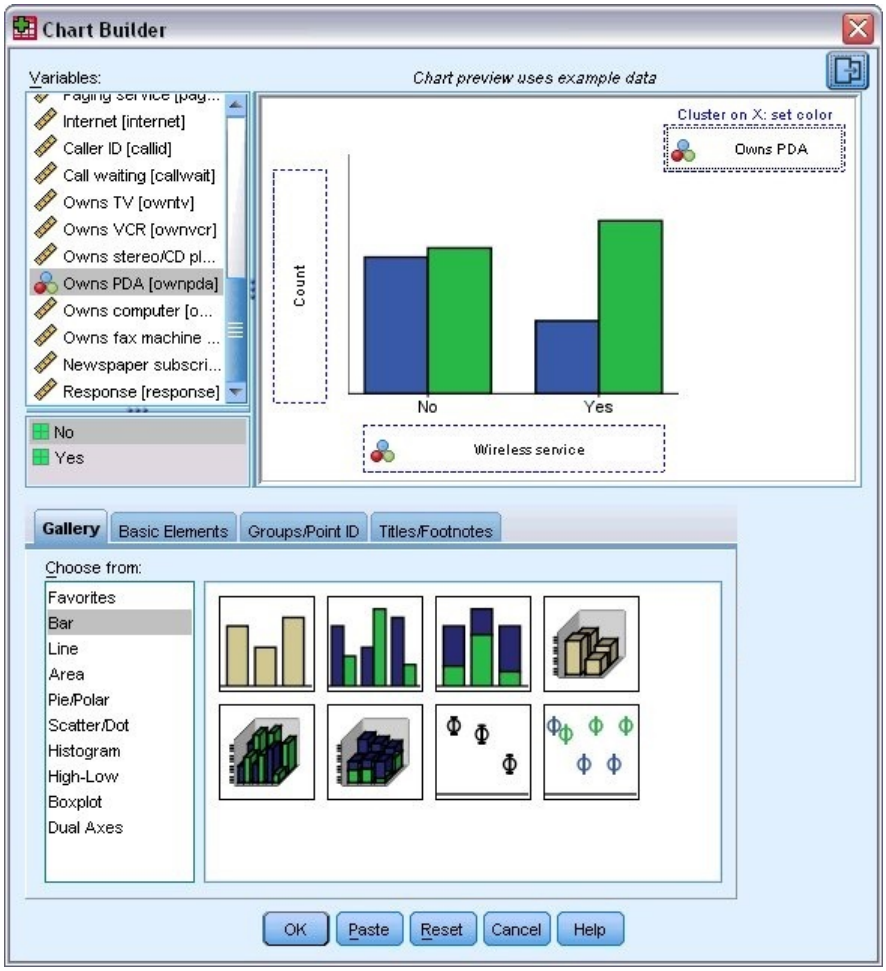


Abbildung 7. Dialogfeld "Diagrammerstellung" mit belegten Ablegebereichen

2. Klicken Sie auf die Registerkarte **Galerie**, falls diese nicht bereits ausgewählt ist.
3. Klicken Sie auf **Balken**, falls dies nicht bereits ausgewählt ist.
4. Ziehen Sie das Symbol "Gruppiertes Balkendiagramm" in den Erstellungsbereich. Dies ist der große Bereich oberhalb der Galerie.
5. Blättern Sie in der Liste "Variablen" abwärts, klicken Sie mit der rechten Maustaste auf *Schnurloses Telefon [kabellos]* und wählen Sie als Messniveau **Nominal** aus.
6. Ziehen Sie die Variable *Schnurloses Telefon [kabellos]* auf die *X*-Achse.
7. Klicken Sie mit der rechten Maustaste auf *Palm Pilot im Haushalt vorhanden [palm]* und wählen Sie als Messniveau **Nominal** aus.
8. Ziehen Sie die Variable *Palm Pilot im Haushalt vorhanden [palm]* auf den Clusterablegebereich, der sich in der rechten oberen Ecke des Erstellungsbereichs befindet.
9. Klicken Sie auf **OK**, um das Diagramm zu erstellen.

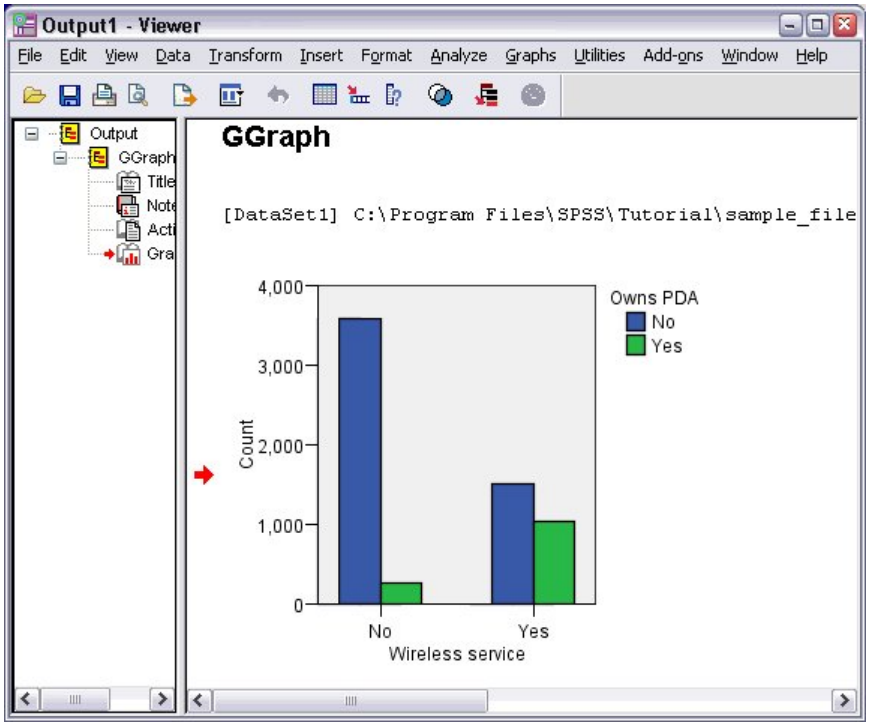


Abbildung 8. Balkendiagramm im Fenster "Viewer"

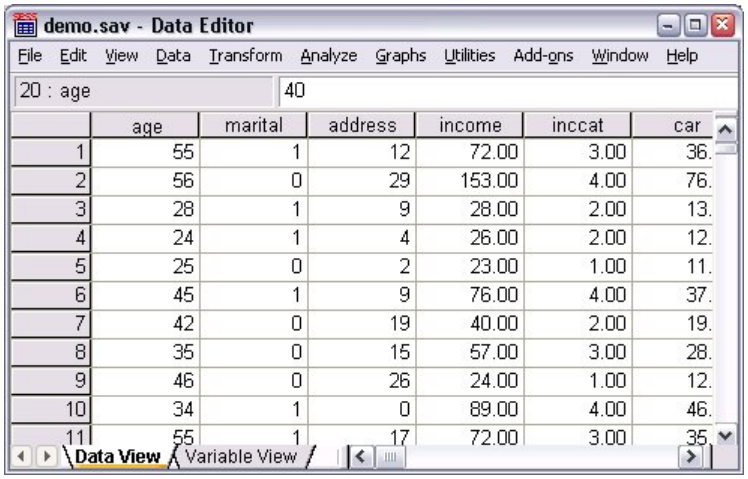
Das Balkendiagramm wird im Viewer angezeigt. Das Diagramm zeigt, dass Benutzer von schnurlosen Telefonen weitaus häufiger Palm Pilots besitzen als Personen ohne schnurloses Telefon.

Diagramme und Tabellen können bearbeitet werden, indem Sie im Inhaltsfenster des Fensters "Viewer" darauf doppelklicken. Außerdem können die Ergebnisse kopiert und in andere Anwendungen eingefügt werden. Diese Themen werden später behandelt.

Kapitel 2. Einlesen von Daten

Daten können entweder direkt eingegeben oder aus einer Reihe unterschiedlicher Quellen importiert werden. In diesem Kapitel wird das Einlesen von Daten erläutert, die in IBM SPSS Statistik -Datendateien, Tabellenkalkulationsanwendungen wie Microsoft Excel, Datenbankanwendungen wie Microsoft Access und Textdateien gespeichert sind.

Grundlegende Struktur von IBM SPSS Statistik -Datendateien



	age	marital	address	income	inccat	car
1	55	1	12	72.00	3.00	36.
2	56	0	29	153.00	4.00	76.
3	28	1	9	28.00	2.00	13.
4	24	1	4	26.00	2.00	12.
5	25	0	2	23.00	1.00	11.
6	45	1	9	76.00	4.00	37.
7	42	0	19	40.00	2.00	19.
8	35	0	15	57.00	3.00	28.
9	46	0	26	24.00	1.00	12.
10	34	1	0	89.00	4.00	46.
11	55	1	17	72.00	3.00	35.

Abbildung 9. Dateneditor

IBM SPSS Statistik -Datendateien sind in Fälle (Zeilen) und Variablen (Spalten) untergliedert. In dieser Datendatei stellen die Fälle die einzelnen Befragten einer Umfrage dar. Die Variablen stellen die Antworten auf die einzelnen Fragen dar, die in der Umfrage gestellt wurden.

Einlesen von IBM SPSS Statistik -Datendateien

IBM SPSS Statistik -Datendateien mit der Dateinamenerweiterung *.sav* enthalten die gespeicherten Daten.

1. Wählen Sie in den Menüs Folgendes aus:

   Datei > **Öffnen** > **Daten...**

2. Wechseln Sie zu der Datei *demo.sav* und öffnen Sie sie. Weitere Informationen finden Sie im Thema Kapitel 10, „Stichprobendateien”, auf Seite 83.

Die Daten werden nun im Dateneditor angezeigt.

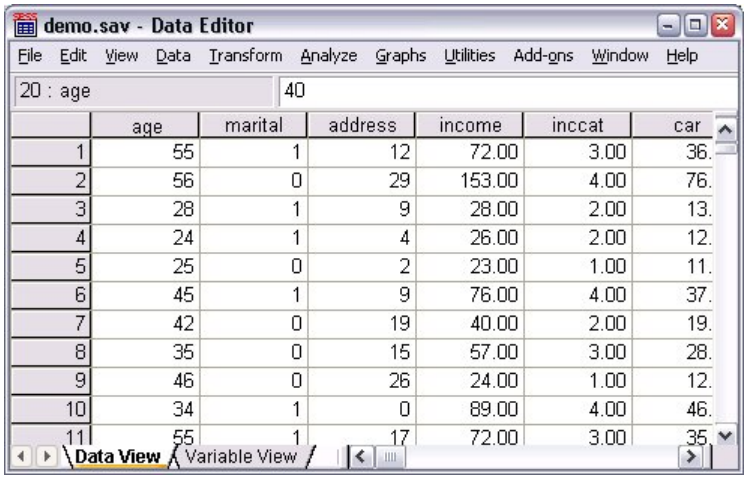


Abbildung 10. Geöffnete Datendatei

Einlesen von Excel-Daten

Statt alle Daten direkt in den Dateneditor einzugeben, können Sie Daten aus Anwendungen wie Microsoft Excel einlesen. Sie können auch die Spaltenüberschriften als Variablennamen einlesen.

1. Wählen Sie in den Menüs Folgendes aus:

   Datei > **Daten importieren** > **Excel**

2. Wechseln Sie in den Ordner `Samples\English` und wählen Sie `demo.xlsx` aus.

   Im Dialogfeld **Excel-Datei lesen** wird eine Vorschau der Datendatei angezeigt. Es wird der Inhalt des ersten Arbeitsblatts in der Datei angezeigt. Wenn die Datei mehrere Arbeitsblätter enthält, können Sie das Arbeitsblatt aus der Liste auswählen.

   Sie können erkennen, dass einige der Zeichenfolgewerte für *Gender* führende Leerzeichen aufweisen. Einige der Werte für *MaritalStatus* werden als Punkt (.) angezeigt.

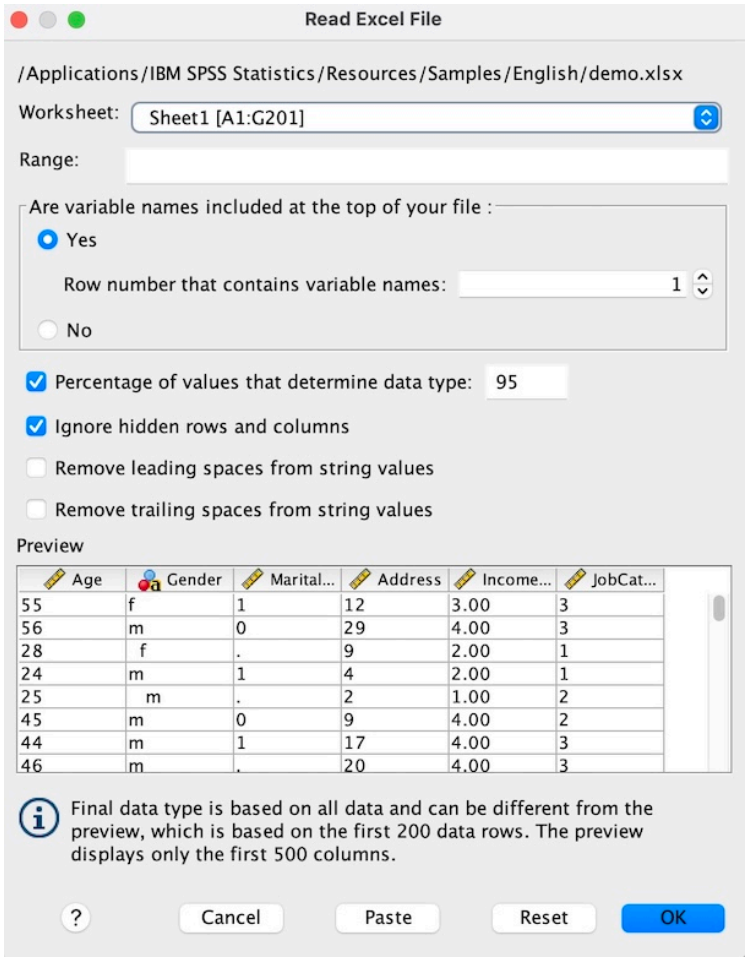
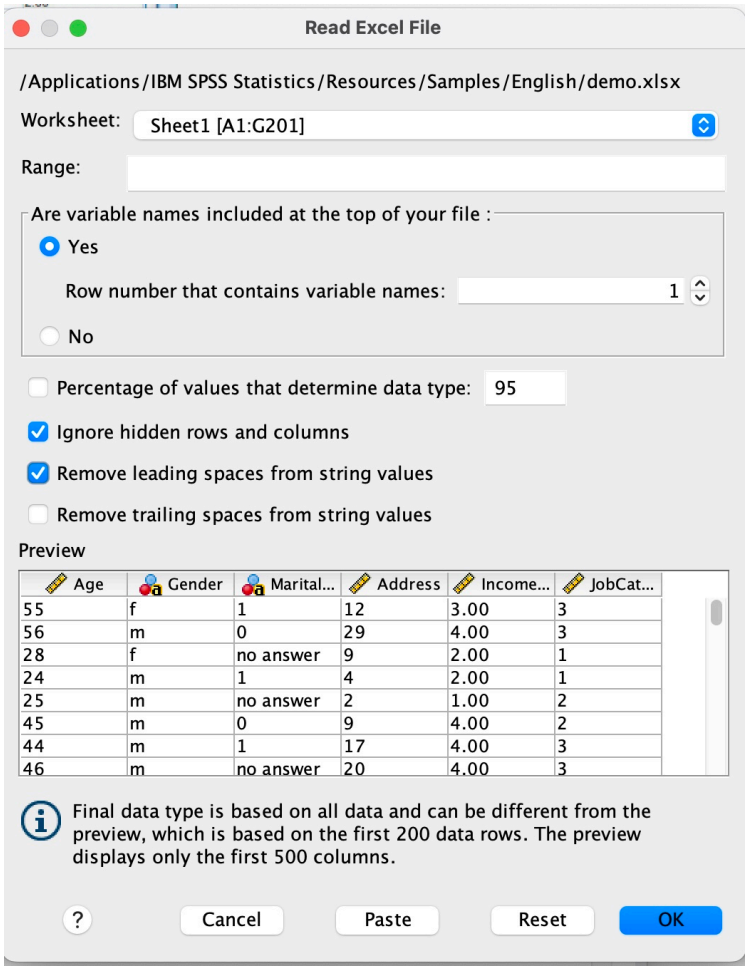


Abbildung 11. Dialogfeld "Excel-Datei lesen"

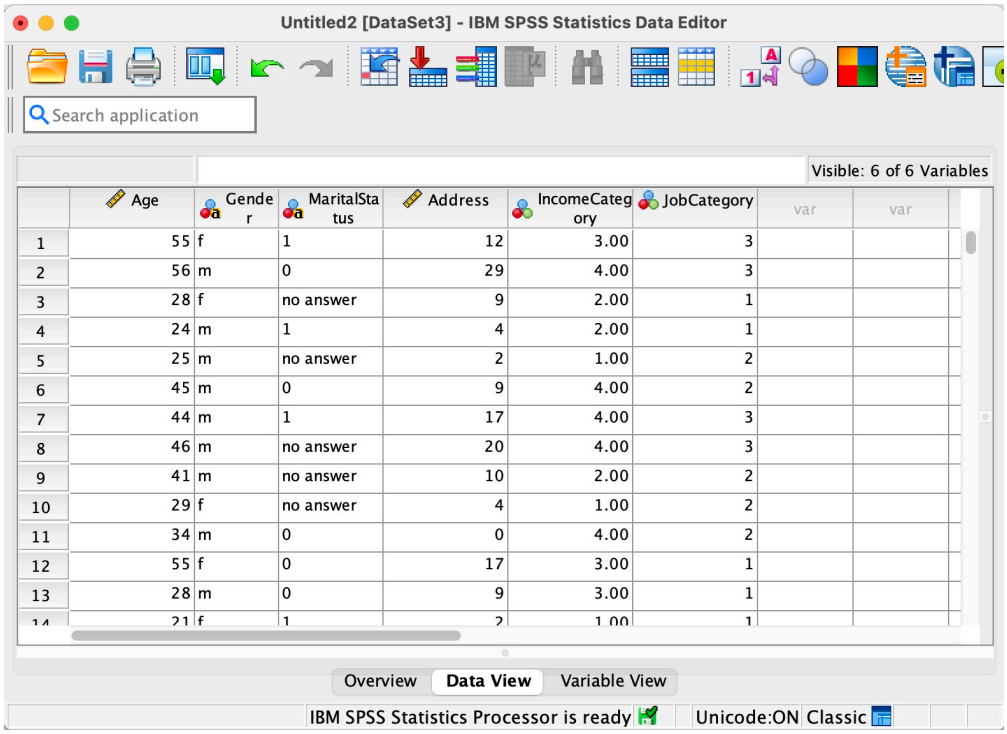
3. Vergewissern Sie sich, dass die Option **Ja** für **Sind Variablennamen am Anfang Ihrer Datei enthalten** ausgewählt ist. Der Standardwert für **Zeilennummern, die Variablennamen enthalten**, ist 1. Wenn die Spaltenüberschriften nicht den Regeln für Variablennamen entsprechen, werden sie in gültige Variablennamen umgewandelt. Die ursprünglichen Spaltenüberschriften werden als Variablenbeschriftungen gespeichert.
4. Wählen Sie **Führende Leerzeichen aus Zeichenfolgewert entfernen** aus.
5. Heben Sie die Auswahl von **Prozentsatz der Werte, die den Datentyp festlegen** auf.



In den systemdefiniert fehlenden Zellen wird jetzt der Zeichenfolgewert "Keine Angabe" angezeigt. Wenn kein Prozentsatz des Wertparameters vorhanden ist und die Spalte eine Kombination aus Datentypen enthält, wird die Variable als Zeichenfolgedatentyp gelesen. Alle Werte werden beibehalten, numerische Werte werden jedoch als Zeichenfolgewerte behandelt.

6. Aktivieren Sie **Prozentsatz der Werte, die den Datentyp festlegen**, damit *heirat* als numerische Variable behandelt wird.
7. Klicken Sie auf **OK**, um die Excel-Datei einzulesen.

Die Daten werden jetzt im Dateneditor angezeigt, wobei die Spaltenüberschriften als Variablennamen verwendet werden. Da die Variablennamen keine Leerzeichen enthalten dürfen, werden die Leerzeichen aus den ursprünglichen Spaltenüberschriften entfernt. Beispielsweise wird die Spaltenüberschrift „Familienstand" in die Variable *Familienstand*konvertiert. Die ursprüngliche Überschrift der Spalte wird als Variablenbeschriftung übernommen.



	Age	Gender	MaritalStatus	Address	IncomeCategory	JobCategory	var	var
1	55	f	1	12	3.00	3		
2	56	m	0	29	4.00	3		
3	28	f	no answer	9	2.00	1		
4	24	m	1	4	2.00	1		
5	25	m	no answer	2	1.00	2		
6	45	m	0	9	4.00	2		
7	44	m	1	17	4.00	3		
8	46	m	no answer	20	4.00	3		
9	41	m	no answer	10	2.00	2		
10	29	f	no answer	4	1.00	2		
11	34	m	0	0	4.00	2		
12	55	f	0	17	3.00	1		
13	28	m	0	9	3.00	1		
14	21	f	1	2	1.00	1		

Abbildung 12. Importierte Excel-Daten

Zugehörige Informationen

„Stichprobendateien” auf Seite 83

Einlesen von Daten aus einer Datenbank

Daten aus Datenbankquellen können mithilfe des Datenbankassistenten problemlos importiert werden. Jede Datenbank, bei der ODBC-Treiber (Open Database Connectivity) verwendet werden, kann nach der Installation der Treiber direkt eingelesen werden. Auf der Installations-CD befinden sich ODBC-Treiber für viele Datenbankformate. Weitere Treiber können von Drittanbietern bezogen werden. In diesem Beispiel wird Microsoft Access, eine der gängigsten Datenbankanwendungen, behandelt.

Hinweis: Dieses Beispiel bezieht sich ausschließlich auf Microsoft Windows. Außerdem ist ein ODBC-Treiber für Access erforderlich. Der ODBC-Treiber für Microsoft Access funktioniert nur mit der 32-Bit-Version von IBM SPSS Statistik . Die Schritte sind auf den anderen Plattformen ähnlich. Es kann jedoch ein ODBC-Treiber eines anderen Anbieters für Access erforderlich sein.

1. Wählen Sie in den Menüs Folgendes aus:

   Datei > **Daten importieren** > **Datenbank** > **Neue Abfrage...**

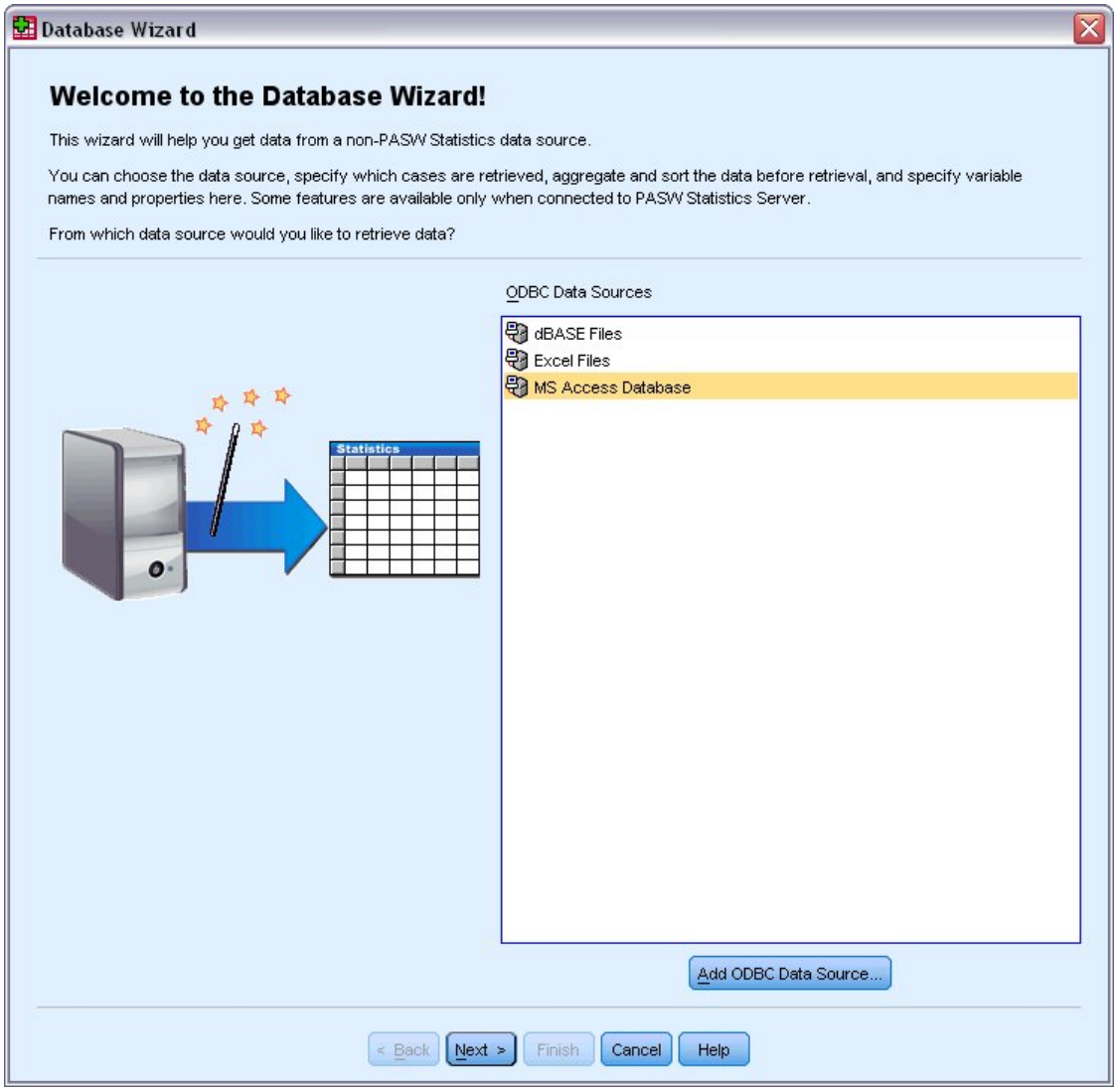


Abbildung 13. Eingangsanzeige des Datenbankassistenten

2. Wählen Sie in der Liste der Datenquellen die Option **MS Access Database** aus und klicken Sie dann auf **Weiter**.

   Hinweis: Abhängig von Ihrer Installation werden auf der linken Seite des Assistenten möglicherweise auch OLE DB-Datenquellen angezeigt (nur bei Windows-Betriebssystemen). In diesem Beispiel wird jedoch die Liste der ODBC-Datenquellen verwendet, die auf der rechten Seite angezeigt wird.

3. Klicken Sie auf **Durchsuchen**, um die zu öffnende Access-Datenbank zu suchen.
4. Öffnen Sie die Datei *demo.mdb*. Weitere Informationen finden Sie im Thema Kapitel 10, „Stichprobendateien", auf Seite 83.
5. Klicken Sie im Anmeldungsdialogfeld auf **OK**.

   Im folgenden Schritt können Sie die zu importierenden Tabellen und Variablen angeben.

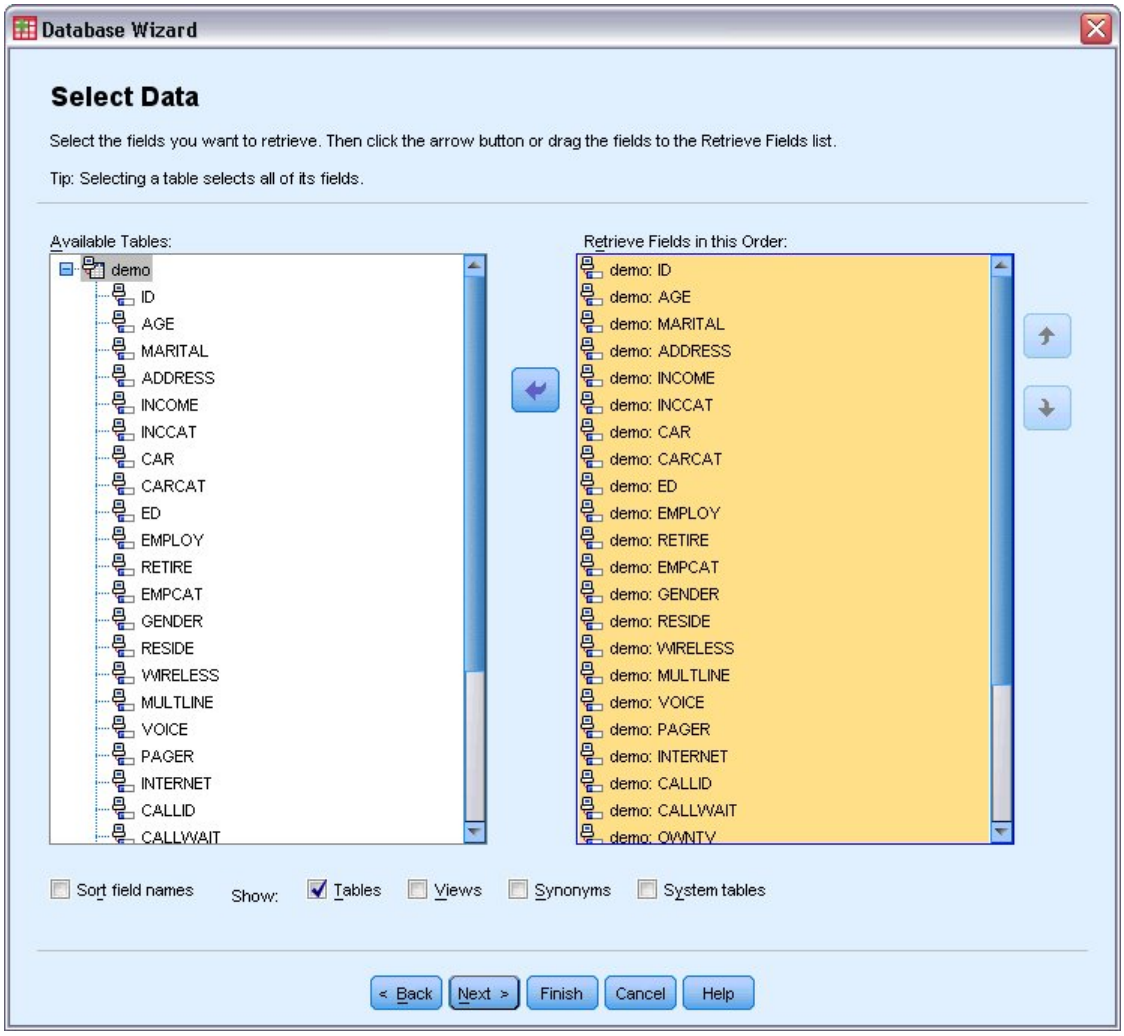


Abbildung 14. Schritt "Daten auswählen"

6. Ziehen Sie die gesamte Tabelle **Demo** in die Liste "Felder in dieser Reihenfolge einlesen".
7. Klicken Sie auf **Weiter**.

   Im folgenden Schritt können Sie auswählen, welche Datensätze (Fälle) importiert werden sollen.

   Wenn Sie nicht alle Fälle importieren möchten, können Sie ein Subset von Fällen (beispielsweise Männer, die älter sind als 30) oder auch eine Zufallsstichprobe von Fällen aus der Datenquelle importieren. Bei großen Datenquellen soll die Anzahl der Fälle möglicherweise auf eine kleine, repräsentative Auswahl begrenzt werden, um die Verarbeitungszeit zu verkürzen.
8. Klicken Sie zum Fortfahren auf **Weiter**.

   Feldnamen werden für die Erstellung von Variablennamen verwendet. Falls erforderlich, werden die Namen in gültige Variablennamen konvertiert. Die ursprünglichen Feldnamen werden als Variablenbeschriftungen übernommen. Die Variablennamen können vor dem Importieren der Datenbank geändert werden.

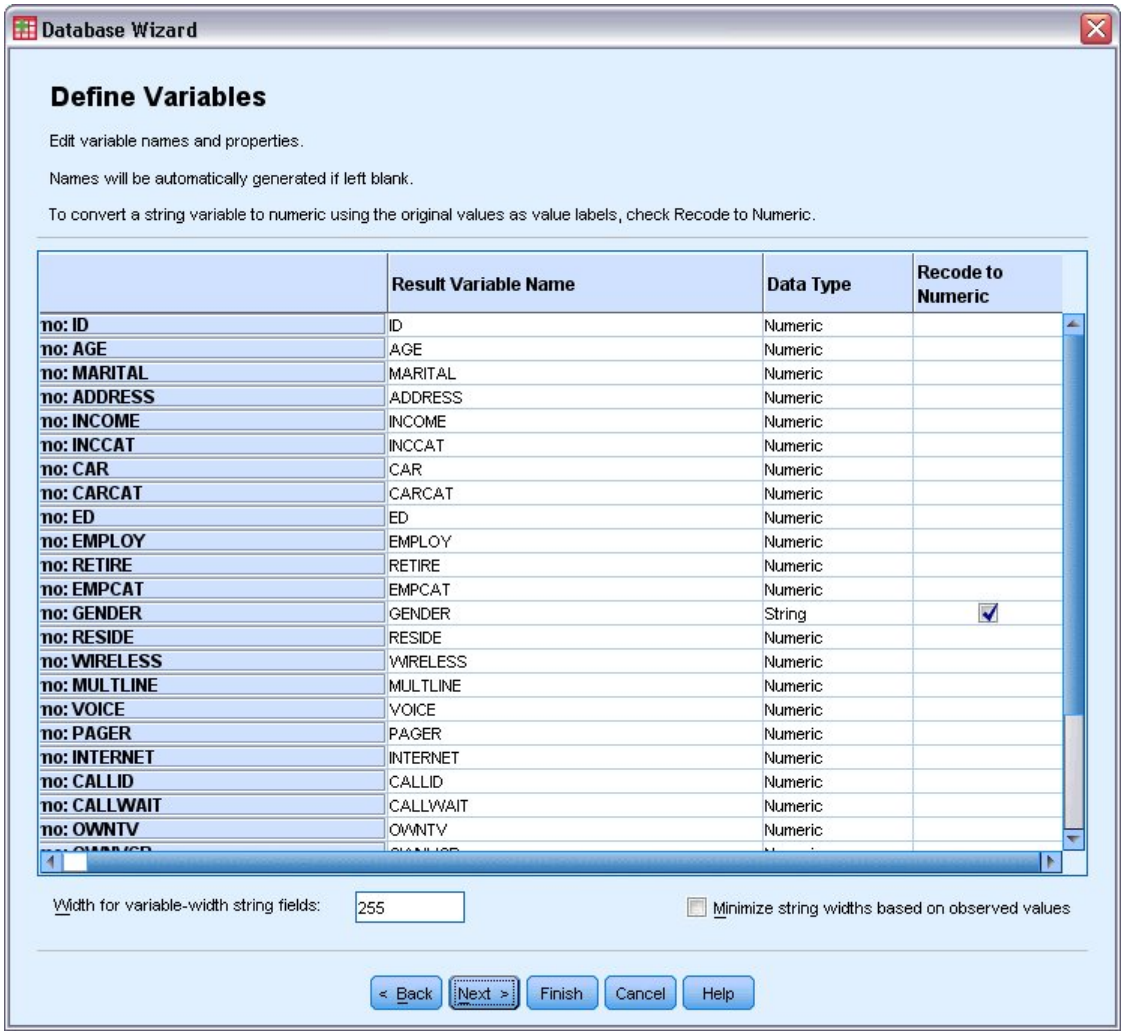


Abbildung 15. Schritt "Variablen definieren"

9. Klicken Sie im Feld "GENDER" auf die Zelle **Als numerisch umcodieren**. Durch diese Option werden Zeichenfolgevariablen in ganzzahlige Variablen umgewandelt und der ursprüngliche Wert wird als Wertbeschriftung für die neue Variable beibehalten.
10. Klicken Sie zum Fortfahren auf **Weiter**.

    Die aus Ihren Angaben im Datenbankassistenten erstellte SQL-Anweisung wird im Schritt "Ergebnisse" angezeigt. Diese Anweisung kann sofort ausgeführt oder für die spätere Verwendung in einer Datei gespeichert werden.
11. Klicken Sie auf **Fertigstellen**, um die Daten zu importieren.

Alle in der Access-Datenbank zum Importieren ausgewählten Daten stehen jetzt im Dateneditor zur Verfügung.

Einlesen von Daten aus einer Textdatei

Textdateien stellen eine weitere gängige Datenquelle dar. Bei vielen Tabellenkalkulationsprogrammen und Datenbanken kann der Inhalt in einer Reihe von Textdateiformaten gespeichert werden. Der Begriff komma- oder tabstoppgetrennte Datei bezieht sich auf das Trennzeichen (Komma oder Tabulator), das die einzelnen Variablen in einer Datenzeile voneinander trennt. In diesem Beispiel sind die Daten tabstoppgetrennt.

1. Wählen Sie in den Menüs Folgendes aus:

   Datei > **Daten importieren** > **Textdaten**

2. Wechseln Sie in den Ordner `Samples\English` und wählen Sie `demo.txt` aus.

   Der Assistent für Textimport unterstützt Sie bei der Definition, wie die angegebene Textdatei interpretiert wird.

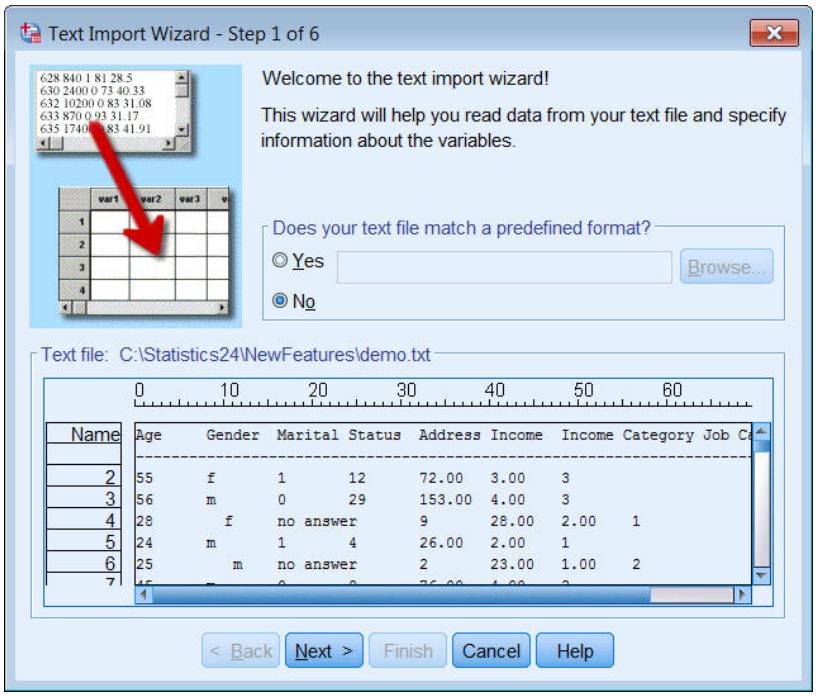


Abbildung 16. Assistent für Textimport: Schritt 1 von 6

3. In Schritt 1 können Sie ein vordefiniertes Format auswählen oder im Assistenten ein neues Format erstellen. Wählen Sie **Nein** aus.
4. Klicken Sie zum Fortfahren auf **Weiter**.

   Wie bereits erwähnt, werden in dieser Datei Tabulatoren als Trennzeichen verwendet. Außerdem werden die Variablennamen in der obersten Zeile dieser Datei definiert.
5. Wählen Sie in Schritt 2 des Assistenten **Mit Trennzeichen** aus, um anzugeben, dass die Daten durch Trennzeichen voneinander getrennt werden.
6. Wählen Sie **Ja** aus, um anzugeben, dass die Datei am Anfang Variablennamen enthält.
7. Klicken Sie zum Fortfahren auf **Weiter**.
8. Geben Sie in Schritt 3 den Wert 2 als Zeilennummer für die Zeile an, in der der erste Fall in den Daten beginnt (da Variablennamen in der ersten Zeile stehen).
9. Behalten Sie die Standardwerte für diesen Schritt bei und klicken Sie zum Fortfahren auf **Weiter**.

   In der Datenvorschau in Schritt 4 können Sie sich vergewissern, dass die Datei richtig eingelesen wird.
10. Wählen Sie **Tabulator** aus und inaktivieren Sie die anderen Trennzeichenoptionen. Standardmäßig ist **Leerzeichen** ausgewählt, weil die Datei Leerzeichen enthält. Für diese Datei sind Leerzeichen Teil der Datenwerte und keine Trennzeichen. Sie müssen die Auswahl von **Leerzeichen** aufheben, damit die Datei richtig eingelesen wird.
11. Wählen Sie **Führende Leerzeichen aus Zeichenfolgewert entfernen** aus. Leerzeichen am Anfang von Zeichenfolgewerten wirken sich auf die Auswertung von Zeichenfolgewerten in Ausdrücken aus. In dieser Datei haben einige Werte für *Gender* führende Leerzeichen, die nicht Teil des Werts sind. Wenn Sie diese Leerzeichen nicht entfernen, wird der Wert " f" als ein anderer Wert als "f" behandelt.

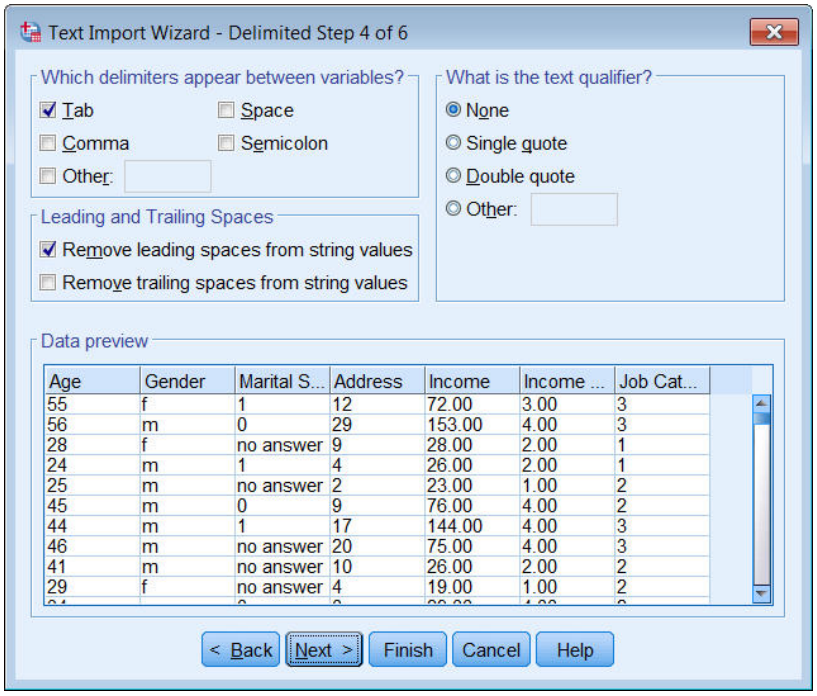


Abbildung 17. Assistent für Textimport: Schritt 4 von 6

12. Klicken Sie zum Fortfahren auf **Weiter**.

    Da die Variablennamen aufgrund der Benennungsregeln geändert werden, können Sie unpassende Namen in Schritt 5 ändern.

    An dieser Stelle können auch Datentypen definiert werden. Sie können zum Beispiel *Einkommen* in ein Dollarwährungsformat ändern.

    So ändern Sie den Datentyp:

13. Wählen Sie in der **Datenvorschau**die Option *Einkommen* aus.
14. Wählen Sie in der Liste "Datenformat" den Eintrag **Dollar** aus.

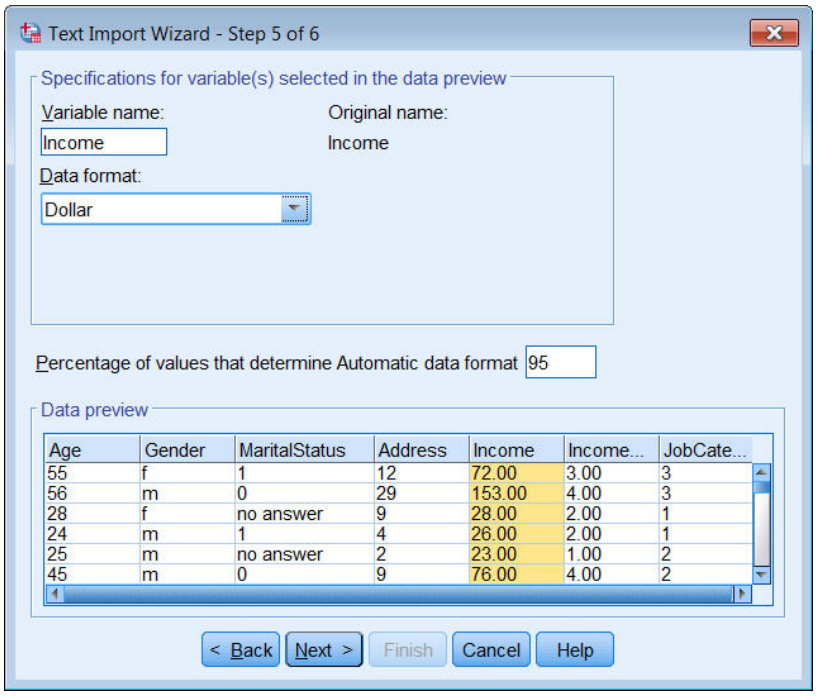


Abbildung 18. Ändern des Datentyps

Die Variable *MaritalStatus* enthält sowohl Zeichenfolgewerte als auch numerische Werte. Weniger als fünf Prozent der Werte sind Zeichenfolgen. Bei der Standardeinstellung von 95 % für **Prozentsatz der Werte, die das automatische Datenformat festlegen** wird die Variable als numerischer Wert behandelt und Zeichenfolgewerte werden als systemdefiniert fehlend festgelegt. Wenn kein Datenformat dem Prozentwert entspricht, wird die Variable als Zeichenfolgevariable behandelt. Wenn Sie die Einstellung in 100 ändern, werden alle Werte beibehalten, aber alle numerischen Werte werden als Zeichenfolgen behandelt.

15. Klicken Sie zum Fortfahren auf **Weiter**.
16. Behalten Sie die Standardeinstellungen im letzten Schritt bei und klicken Sie auf **Fertigstellen**, um die Daten zu importieren.

Kapitel 3. Verwenden des Dateneditors

Der Dateneditor zeigt den Inhalt der aktiven Datendatei an. Die Informationen im Dateneditor bestehen aus Variablen und Fällen.

- In der Datenansicht stellen Spalten Variablen und Zeilen Fälle (Beobachtungen) dar.
- In der Variablenansicht entspricht eine Zeile einer Variablen und eine Spalte einem Attribut, das dieser Variablen zugeordnet ist.

Variablen dienen der Darstellung der verschiedenen Typen von Daten, die Sie zusammengestellt haben. Häufig wird dies mit einer Umfrage verglichen. Die Antwort auf jede Frage einer Umfrage entspricht einer Variablen. Es gibt viele verschiedene Variablentypen, z. B. Zahlen, Zeichenfolgen, Währungen und Datumsangaben.

Eingeben von numerischen Daten

Daten können im Dateneditor eingegeben werden. Dies kann bei kleinen Datendateien und bei kleineren Änderungen in größeren Datendateien nützlich sein.

1. Klicken Sie am unteren Rand des Fensters "Dateneditor" auf die Registerkarte **Variablenansicht**.

   Sie müssen die Variablen definieren, die verwendet werden sollen. In diesem Fall werden nur drei Variablen benötigt: *alter*, *heirat* und *einkomm*.

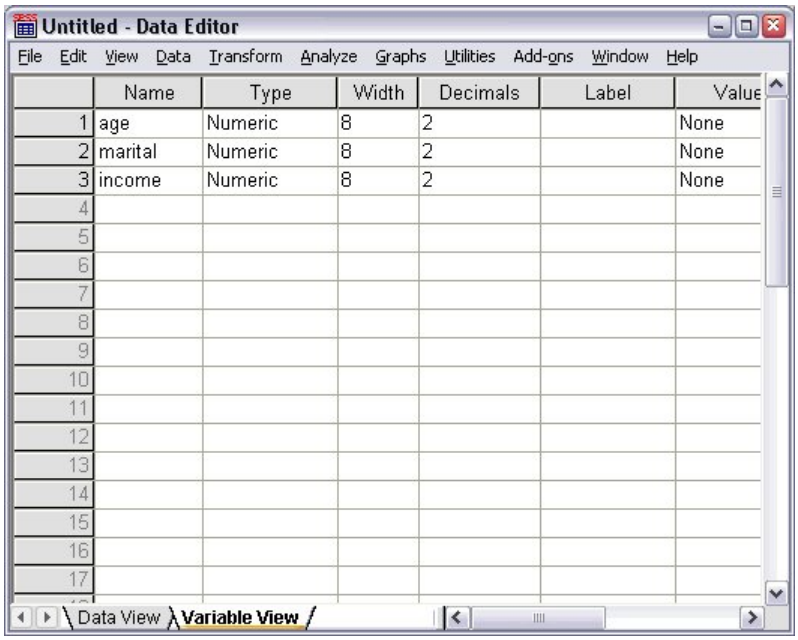


   Abbildung 19. Variablennamen in der Variablenansicht

2. Geben Sie in der ersten Zeile der ersten Spalte `age` ein.
3. Geben Sie in der zweiten Zeile `marital` ein.
4. Geben Sie in der dritten Zeile `income` ein.

   Neue Variablen erhalten automatisch den Datentyp "Numerisch".

   Wenn Sie keine Variablennamen eingeben, werden automatisch eindeutige Namen generiert. Diese Namen sind allerdings nicht aussagekräftig und es wird empfohlen, sie nicht für große Datendateien zu verwenden.
5. Klicken Sie auf die Registerkarte **Datenansicht**, um die Eingabe von Daten fortzusetzen.

Die in der Variablenansicht eingegebenen Namen sind nun die Überschriften der ersten drei Spalten der Datenansicht.

Beginnen Sie mit der Dateneingabe in der ersten Zeile der ersten Spalte.

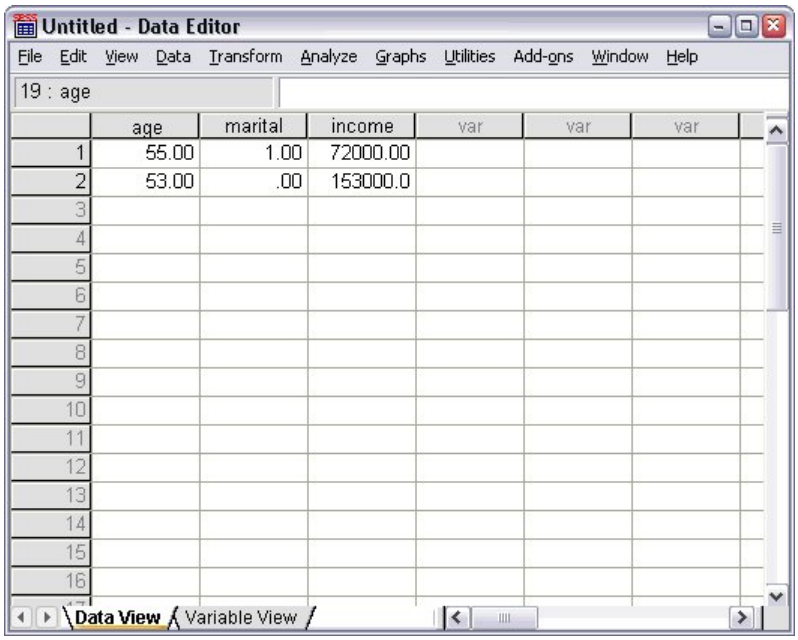


Abbildung 20. In der Datenansicht eingegebene Werte

6. Geben Sie in der Spalte *alter* den Wert 55 ein.
7. Geben Sie in der Spalte *heirat* den Wert 1 ein.
8. Geben Sie in der Spalte *einkomm* den Wert 72000 ein.
9. Positionieren Sie den Cursor in der zweiten Zeile der ersten Spalte und geben Sie die Daten für die nächste Person ein.
10. Geben Sie in der Spalte *alter* den Wert 53 ein.
11. Geben Sie in der Spalte *heirat* den Wert 0 ein.
12. Geben Sie in der Spalte *einkomm* den Wert 153000 ein.

    Derzeit werden die Werte in den Spalten *alter* und *heirat* mit Dezimaltrennzeichen dargestellt, obwohl die Werte als ganze Zahlen gemeint sind. So blenden Sie die Dezimaltrennzeichen bei diesen Variablen aus:
13. Klicken Sie am unteren Rand des Fensters "Dateneditor" auf die Registerkarte **Variablenansicht**.
14. Geben Sie der Spalte *Dezimalstellen* der Zeile *alter* den Wert 0 ein, um die Dezimalstellen auszublenden.
15. Geben Sie der Spalte *Dezimalstellen* der Zeile *heirat* den Wert 0 ein, um die Dezimalstellen auszublenden.

Eingeben von Zeichenfolgedaten

Im Dateneditor können auch nicht numerische Daten wie Textzeichenfolgen eingegeben werden.

1. Klicken Sie am unteren Rand des Fensters "Dateneditor" auf die Registerkarte **Variablenansicht**.
2. Geben Sie in die erste Zelle der ersten leeren Zeile sex als Variablennamen ein.
3. Klicken Sie auf die Zelle *Typ* neben Ihrer Eingabe.
4. Klicken Sie auf die Schaltfläche rechts neben der Zelle *Typ*, um das Dialogfeld "Variablentyp definieren" zu öffnen.
5. Wählen Sie **Zeichenfolge** aus, um den Variablentyp anzugeben.

6. Klicken Sie auf **OK**, um die Auswahl zu speichern und zum Dateneditor zurückzukehren.

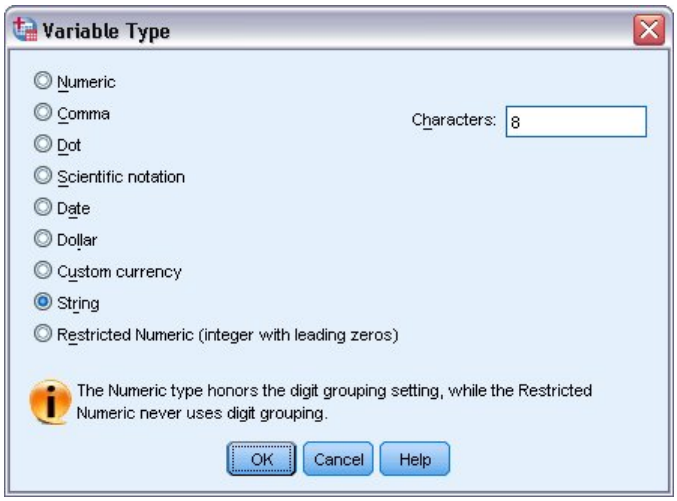


Abbildung 21. Dialogfeld "Variablentyp definieren"

Definieren von Daten

Neben Datentypen können Sie auch aussagekräftige Variablenbeschriftungen und Wertbeschriftungen für Variablennamen und Datenwerte definieren. Diese aussagekräftigen Beschriftungen werden in statistischen Berichten und Diagrammen verwendet.

Hinzufügen von Variablenbeschriftungen

Beschriftungen sollen Beschreibungen von Variablen darstellen. Diese Beschreibungen sind oft längere Versionen von Variablennamen. Beschriftungen können bis zu 255 Byte umfassen. Diese Beschriftungen werden in der Ausgabe verwendet, um die Variablen zu identifizieren.

1. Klicken Sie am unteren Rand des Fensters "Dateneditor" auf die Registerkarte **Variablenansicht**.
2. Geben Sie in der Spalte *"Label"* der Zeile *"Alter"* `Respondent's Age` ein.
3. Geben Sie in der Spalte *"Label"* der Zeile " *Ehestand* " `Marital Status` ein.
4. Geben Sie in der Spalte *Label* der *Einkommenszeile* `Household Income` ein.
5. Geben Sie in der Spalte *"Label"* der Zeile *"Geschlecht"* `Gender` ein.

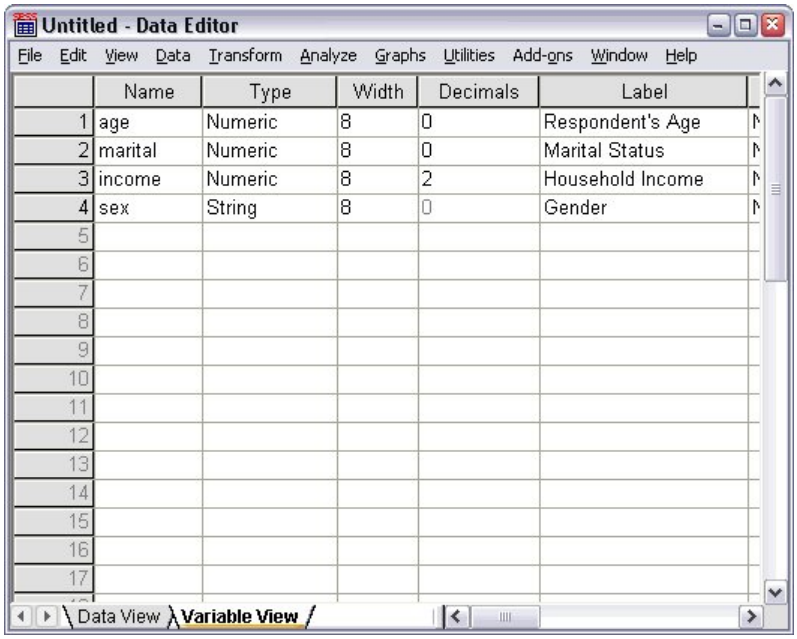


Abbildung 22. In der Variablenansicht eingegebene Variablenbeschriftungen

Ändern des Datentyps und Formats von Variablen

In der Spalte *Typ* wird der aktuelle Datentyp für jede Variable angezeigt. Meist werden die Datentypen "Numerisch" und "Zeichenfolge" verwendet, aber es werden viele weitere Formate unterstützt. In der aktuellen Datendatei ist die Variable *einkomm* als numerischer Typ definiert.

1. Klicken Sie auf die Zelle *Typ* in der Zeile *einkomm* und klicken Sie dann auf die Schaltfläche auf der rechten Seite der Zelle, um das Dialogfeld "Variablentyp definieren" zu öffnen.
2. Wählen Sie **Dollar** aus.

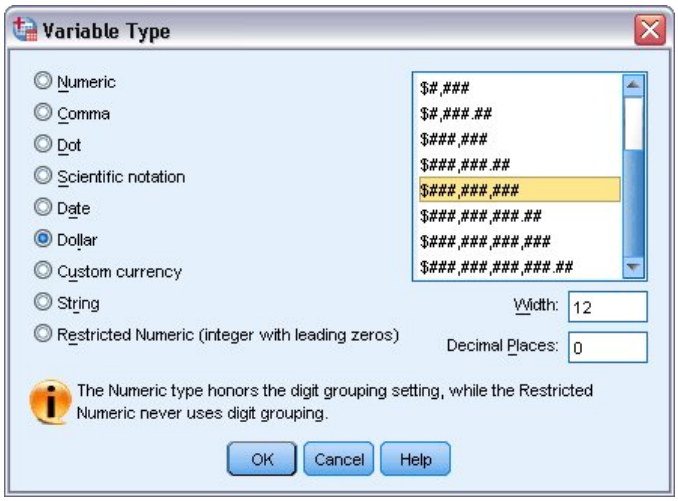


Abbildung 23. Dialogfeld "Variablentyp definieren"

Die Formatierungsoptionen für den derzeit ausgewählten Datentyp werden angezeigt.

3. Wählen Sie für das Währungsformat in diesem Beispiel **$###,###,###** aus.
4. Klicken Sie auf **OK**, um die Änderungen zu speichern.

Hinzufügen von Wertbeschriftungen

Mit Wertbeschriftungen können Variablenwerte einer Zeichenfolgebeschriftung zugeordnet werden. In diesem Beispiel sind für die Variable *heirat* zwei Werte zulässig. 0 bedeutet, dass die Person ledig ist, und 1 bedeutet, dass sie verheiratet ist.

1. Klicken Sie auf die Zelle *Wertbeschriftungen* in der Zeile *heirat* und klicken Sie dann auf die Schaltfläche auf der rechten Seite der Zelle, um das Dialogfeld "Wertbeschriftungen" zu öffnen.

   Wert ist der eigentliche numerische Wert.

   Wertbeschriftung ist die Zeichenfolgebeschriftung, die dem angegebenen numerischen Wert zugeordnet wird.
2. Geben Sie `0` im Feld "Wert" ein.
3. Geben Sie `Single` in das Feld Bezeichnung ein.
4. Klicken Sie auf **Hinzufügen**, um diese Beschriftung der Liste hinzuzufügen.

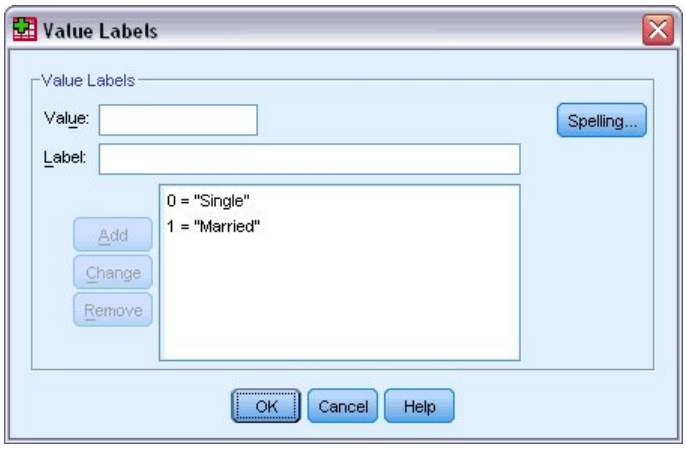


Abbildung 24. Dialogfeld "Wertbeschriftungen"

5. Geben Sie 1 in das Feld Wert und `Married` in das Feld Bezeichnung ein.
6. Klicken Sie auf **Hinzufügen** und anschließend auf **OK**, um die Änderungen zu speichern und zum Dateneditor zurückzukehren.

   Diese Beschriftungen können auch in der Datenansicht angezeigt werden, wodurch die Daten u. U. besser lesbar werden.
7. Klicken Sie am unteren Rand des Fensters "Dateneditor" auf die Registerkarte **Datenansicht**.
8. Wählen Sie in den Menüs Folgendes aus:

Anzeigen > **Wertbeschriftungen**

Die Beschriftungen werden nun in einer Liste angezeigt, wenn Sie im Dateneditor Werte eingeben. Dies hat den Vorteil, dass zulässige Antworten vorgeschlagen werden und die Antwort aussagekräftiger ist.

Wenn das Menüelement "Wertbeschriftungen" bereits aktiviert wurde (neben dem Menüelement wird ein Haken angezeigt) und Sie erneut **Wertbeschriftungen** auswählen, wird die Anzeige von Wertbeschriftungen *inaktiviert*.

Umgang mit fehlenden Daten

Fehlende oder ungültige Werte stellen ein häufig auftretendes Problem dar, das nicht vernachlässigt werden darf. Die Teilnehmer an einer Umfrage beantworten eine Frage vielleicht absichtlich nicht oder sie wissen keine Antwort oder die Form der Antwort ist unerwartet. Wenn Sie keine Maßnahmen zum Filtern oder Identifizieren dieser Daten treffen, ist das Ergebnis der Analyse möglicherweise ungenau.

Bei numerischen Daten werden leere Datenfelder oder Felder mit ungültiger Eingabe in systemdefiniert fehlende Felder konvertiert. Dies wird durch einen einzelnen Punkt gekennzeichnet.

Für die Analyse kann es wichtig sein zu wissen, warum ein Wert fehlt. Es könnte beispielsweise wichtig sein, zwischen Personen zu unterscheiden, die eine Frage nicht beantworten wollten, und solchen, die nicht geantwortet haben, weil die Frage sie nicht betraf.

Fehlende Werte für eine numerische Variable

1. Klicken Sie am unteren Rand des Fensters "Dateneditor" auf die Registerkarte **Variablenansicht**.
2. Klicken Sie auf die Zelle *Fehlende Werte* in der Zeile *alter* und klicken Sie dann auf die Schaltfläche auf der rechten Seite der Zelle, um das Dialogfeld "Fehlende Werte definieren" zu öffnen.

   In diesem Dialogfeld können Sie bis zu drei verschiedene fehlende Werte oder einen Wertebereich und zusätzlich einen diskreten Wert angeben.

   Abbildung 25. Dialogfeld "Fehlende Werte"
3. Wählen Sie **Einzelne fehlende Werte** aus.
4. Geben Sie im ersten Textfeld 999 ein und lassen Sie die beiden Textfelder anderen leer.
5. Klicken Sie auf **OK**, um die Änderungen zu speichern und zum Dateneditor zurückzukehren.

   Nachdem nun der Wert für fehlende Daten definiert wurde, kann diesem Wert eine Beschriftung zugeordnet werden.
6. Klicken Sie auf die Zelle *Wertbeschriftungen* in der Zeile *alter* und klicken Sie dann auf die Schaltfläche auf der rechten Seite der Zelle, um das Dialogfeld "Wertbeschriftungen" zu öffnen.
7. Geben Sie 999 in das Feld "Wert" ein.
8. Geben Sie `No Response` in das Feld Bezeichnung ein.
9. Klicken Sie auf **Hinzufügen**, um diese Beschriftung der Datendatei hinzuzufügen.
10. Klicken Sie auf **OK**, um die Änderungen zu speichern und zum Dateneditor zurückzukehren.

Fehlende Werte für eine Zeichenfolgevariable

Fehlende Werte für Zeichenfolgevariablen werden ähnlich wie solche für numerische Werte behandelt. Im Gegensatz zu numerischen Variablen werden Leerfelder bei Zeichenfolgevariablen jedoch nicht als systemdefiniert fehlend betrachtet. Sie werden stattdessen als leere Zeichenfolge interpretiert.

1. Klicken Sie am unteren Rand des Fensters "Dateneditor" auf die Registerkarte **Variablenansicht**.
2. Klicken Sie auf die Zelle *Fehlende Werte* in der Zeile *geschl* und klicken Sie dann auf die Schaltfläche auf der rechten Seite der Zelle, um das Dialogfeld "Fehlende Werte definieren" zu öffnen.
3. Wählen Sie **Einzelne fehlende Werte** aus.
4. Geben Sie NR in das erste Textfeld ein.

   Bei fehlenden Werten für Zeichenfolgevariablen wird zwischen Groß- und Kleinschreibung unterschieden. Der Wert *ka* wird daher nicht als fehlender Wert behandelt.
5. Klicken Sie auf **OK**, um die Änderungen zu speichern und zum Dateneditor zurückzukehren.

   Jetzt können Sie eine Beschriftung für den fehlenden Wert definieren.

6. Klicken Sie auf die Zelle *Werte* in der Zeile *geschl* und klicken Sie dann auf die Schaltfläche auf der rechten Seite der Zelle, um das Dialogfeld "Wertbeschriftungen" zu öffnen.
7. Geben Sie NR in das Feld Wert ein.
8. Geben Sie No Response in das Feld Bezeichnung ein.
9. Klicken Sie auf **Hinzufügen**, um diese Beschriftung zu Ihrem Projekt hinzuzufügen.
10. Klicken Sie auf **OK**, um die Änderungen zu speichern und zum Dateneditor zurückzukehren.

Kapitel 4. Untersuchen der Auswertungsstatistik auf einzelne Variablen

In diesem Abschnitt werden einfache Auswertungsmaße und der Einfluss des Messniveaus einer Variablen auf den zu verwendenden Statistiktyp behandelt. Hierfür wird die Datendatei *demo.sav* verwendet. Weitere Informationen finden Sie im Thema Kapitel 10, „Stichprobendateien", auf Seite 83.

Messniveau

Je nach Messniveau sind für verschiedene Datentypen unterschiedliche Auswertungsmaße geeignet:

Kategorial. Daten mit einer begrenzten Anzahl von eindeutigen Werten oder Kategorien (beispielsweise Geschlecht oder Familienstand). Auch als **qualitative Daten** bezeichnet. Bei kategorialen Variablen kann es sich um Zeichenfolgevariablen (alphanumerisch) oder um numerische Variablen handeln, bei denen numerische Codes zum Darstellen der Kategorien verwendet werden (beispielsweise 0 = *nicht verheiratet* und 1 = *verheiratet*). Es gibt zwei grundlegende Arten von kategorialen Daten:

- **Nominal.** Kategoriale Daten, bei denen die Kategorien keine natürliche Reihenfolge aufweisen. So ist beispielsweise die Berufskategorie *Vertrieb* weder höher noch niedriger als die Kategorie *Marketing* oder *Forschung*.
- **Ordinal.** Kategoriale Daten, bei denen die Kategorien eine sinnvolle Reihenfolge aufweisen, aber keine Distanz zwischen den Kategorien bestimmt werden kann. So liegt bei den Werten *Hoch*, *Mittel* und *Niedrig* zwar beispielsweise eine Reihenfolge vor, eine Distanz zwischen den Werten kann jedoch nicht berechnet werden.

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

Auswertungsmaße für kategoriale Daten

Bei kategorialen Daten ist das typischste Auswertungsmaß die Zahl bzw. der Prozentsatz an Fällen in den einzelnen Kategorien. Der **Modalwert** ist die Kategorie mit der größten Anzahl von Fällen. Wenn eine große Anzahl von Kategorien vorliegt, kann bei ordinalen Daten auch der **Median** (der Wert, ober- und unterhalb dessen die Hälfte aller Fälle angesiedelt sind) ein hilfreiches Auswertungsmaß sein.

Mithilfe der Prozedur "Häufigkeiten" werden Häufigkeitstabellen erstellt, die für jeden bei einer Variablen beobachteten Wert sowohl die Zahl als auch den Prozentwert von Fällen anzeigen.

1. Wählen Sie in den Menüs Folgendes aus:

   Analysieren > **Deskriptive Statistiken** > **Häufigkeiten...**

   Hinweis: Für diese Funktion ist die Option "Statistics Base" erforderlich.
2. Wählen Sie *Palm Pilot im Haushalt vorhanden [palm]* und *Fernseher im Haushalt vorhanden [fernseh]* aus und verschieben Sie diese in die Liste "Variable(n)".

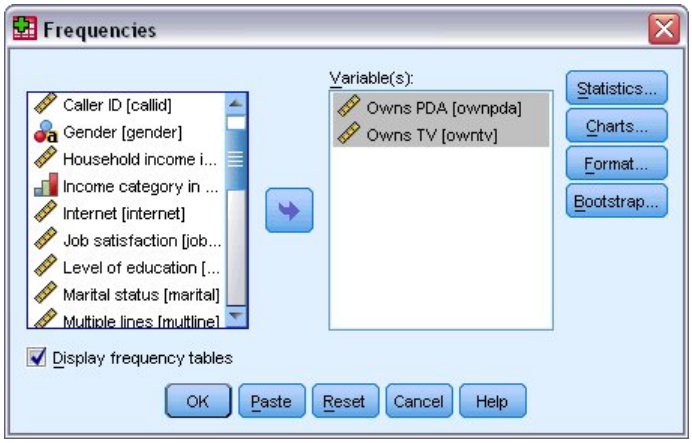


Abbildung 26. Für die Analyse ausgewählte kategoriale Variablen

3. Klicken Sie auf **OK**, um die Prozedur auszuführen.

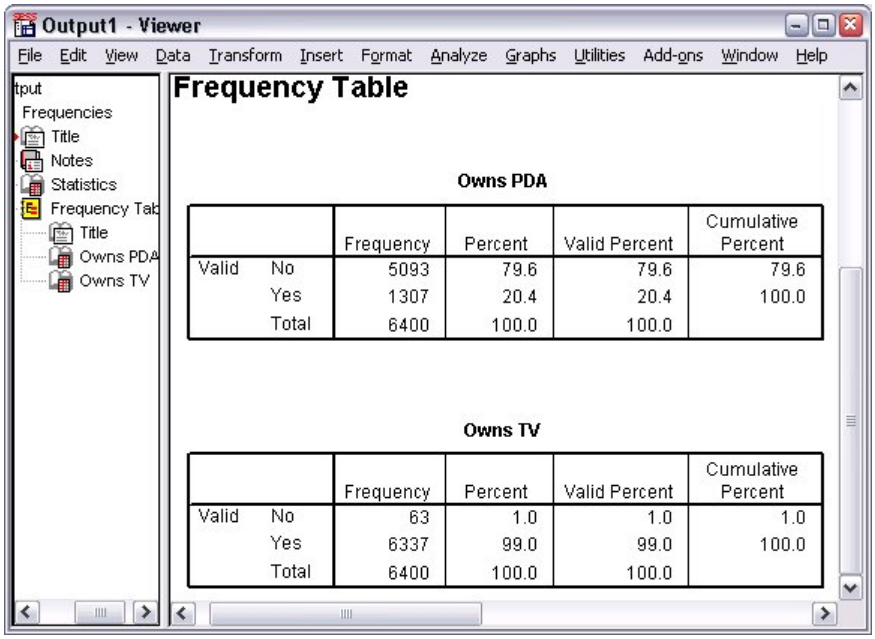


Abbildung 27. Häufigkeitstabellen

Die Häufigkeitstabellen werden im Fenster "Viewer" angezeigt. Die Häufigkeitstabellen besagen, dass nur 20,4 % der Personen Palm Pilots, jedoch fast alle (99,0 %) einen Fernseher besitzen. Dies mag zwar an sich keine besonders interessante Erkenntnis sein, es könnte jedoch durchaus interessant sein, mehr über die kleine Gruppe von Personen zu erfahren, die keinen Fernseher besitzt.

Diagramme für kategoriale Daten

Die in einer Häufigkeitstabelle enthaltenen Informationen können mithilfe eines Balken- oder Kreisdiagramms grafisch dargestellt werden.

1. Öffnen Sie das Dialogfeld "Häufigkeiten" erneut. (Die beiden Variablen sollten noch ausgewählt sein.)

   Sie können auf der Symbolleiste auf die Schaltfläche "Zuletzt verwendete Dialogfelder" klicken, um schnell zu den zuletzt verwendeten Prozeduren zurückzukehren.

Abbildung 28. Schaltfläche "Zuletzt verwendete Dialogfelder"

2. Klicken Sie auf **Diagramme**.
3. Wählen Sie die Option **Balkendiagramme** aus und klicken Sie anschließend auf **Weiter**.
4. Klicken Sie im Hauptdialogfeld auf **OK**, um die Prozedur auszuführen.

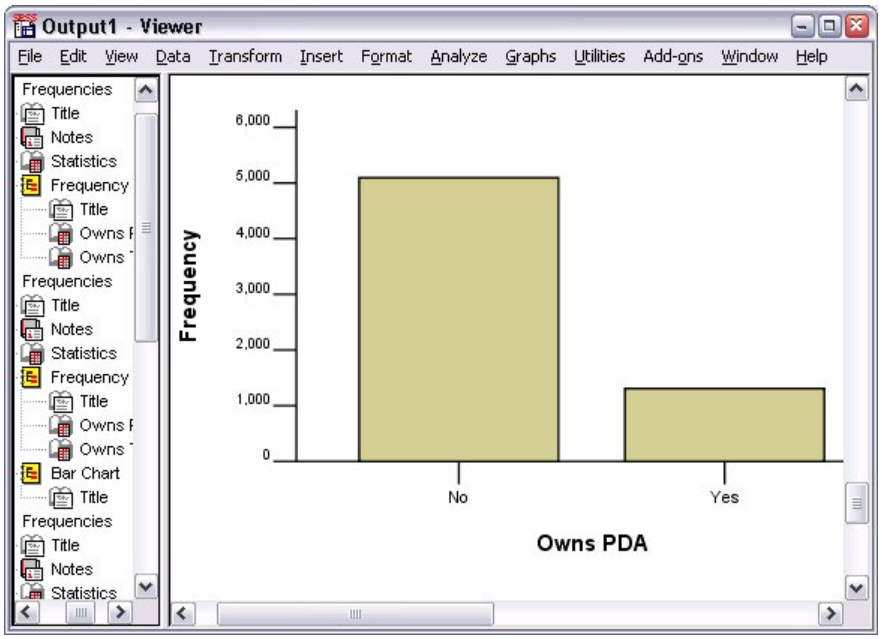


Abbildung 29. Balkendiagramm

Zusätzlich zu den Häufigkeitstabellen werden dieselben Informationen jetzt in Form von Balkendiagrammen dargestellt, wodurch nun leicht zu erkennen ist, dass die meisten Personen keinen Palm Pilot, jedoch fast alle einen Fernseher besitzen.

Auswertungsmaße für metrische Variablen

Für metrische Variablen gibt es viele Auswertungsmaße, wie z. B.:

- **Messungen von Lagemaßen.** Die am häufigsten verwendeten Lagemaße (zentrale Tendenz) sind der **Mittelwert** (arithmetisches Mittel) und der **Median** (Wert, ober- und unterhalb dessen die Hälfte aller Fälle angesiedelt sind).
- **Messungen der Streuung.** Zu den Statistiken, welche die Streubreite oder die Variation in den Daten messen, gehören die Standardabweichung, das Minimum und das Maximum.

1. Öffnen Sie das Dialogfeld "Häufigkeiten" erneut.
2. Klicken Sie auf **Zurücksetzen**, um frühere Einstellungen zu löschen.
3. Wählen Sie den Eintrag *Haushaltseinkommen in Tausend [einkomm]* aus und verschieben Sie ihn in die Liste "Variable(n)".
4. Klicken Sie auf **Statistiken**.
5. Wählen Sie **Mittelwert**, **Median**, **Std. abweichung**, **Minimum** und **Maximum**.
6. Klicken Sie auf **Weiter**.
7. Inaktivieren Sie (durch Entfernen der Auswahlmarkierung) **Häufigkeitstabellen anzeigen** im Hauptdialogfeld. (Häufigkeitstabellen sind bei metrischen Variablen in der Regel nicht besonders nützlich, da fast genauso viele unterschiedliche Werte wie Fälle in der Datendatei vorliegen können.)
8. Klicken Sie auf **OK**, um die Prozedur auszuführen.

Die Tabelle für Häufigkeitsstatistik wird im Fenster "Viewer" angezeigt.

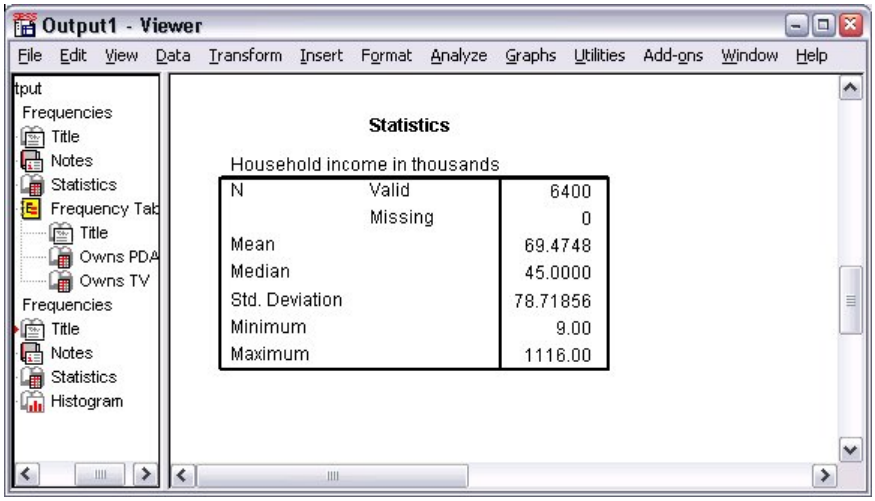


Abbildung 30. Tabelle für Häufigkeitsstatistik

In diesem Beispiel liegt eine große Differenz zwischen dem Mittelwert und dem Median vor. Der Mittelwert übersteigt den Median um fast 25.000. Dies bedeutet, dass die Werte nicht normalverteilt sind. Anhand eines Histogramms kann die Verteilung anschaulich überprüft werden.

Histogramme für metrische Variablen

1. Öffnen Sie das Dialogfeld "Häufigkeiten" erneut.
2. Klicken Sie auf **Diagramme**.
3. Wählen Sie die Option **Histogramm** aus und klicken Sie auf **Mit Normalverteilungskurve**.
4. Klicken Sie auf **Weiter** und danach im Hauptdialogfeld auf **OK**, um die Prozedur auszuführen.

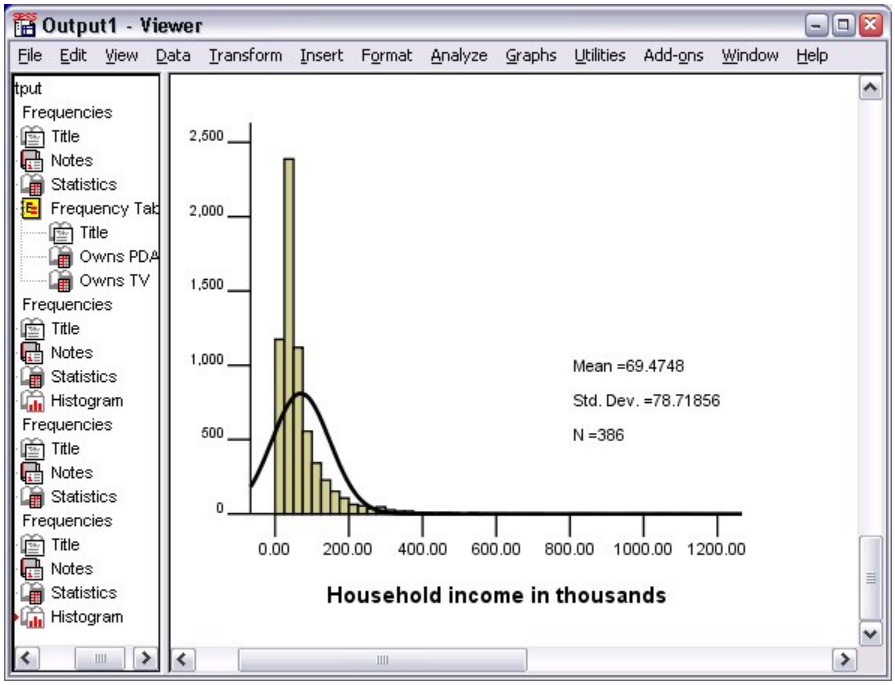


Abbildung 31. Histogramm

Die Mehrheit der Fälle ballt sich am unteren Ende der Skala, wobei die meisten unter 100.000 liegen. Ein paar Fälle liegen jedoch auch im Bereich um 500.000 und darüber (zu wenige, als dass sie ohne Änderung des Histogramms sichtbar wären). Die hohen Werte bei einigen wenigen Fällen haben große

Auswirkungen auf den Mittelwert, aber nur eine geringe oder gar keine Auswirkung auf den Median. Daher ist der Median in diesem Beispiel ein besserer Indikator für das Lagemaß (zentrale Tendenz).

Kapitel 5. Erstellen und Bearbeiten von Diagrammen

Sie können eine Vielzahl von Diagrammtypen erstellen und bearbeiten. In diesem Kapitel wird die Erstellung und Bearbeitung von Balkendiagrammen behandelt Die behandelten Prinzipien können auf alle Diagrammtypen übertragen werden.

Grundlagen der Diagrammerstellung

Um die Grundlagen der Diagrammerstellung zu demonstrieren, wird ein Balkendiagramm des Durchschnittseinkommens für die verschiedenen Stufen der Zufriedenheit mit der Arbeit erstellt. In diesem Beispiel wird die Datendatei *demo.sav* verwendet. Weitere Informationen finden Sie im Thema Kapitel 10, „Stichprobendateien", auf Seite 83.

1. Wählen Sie in den Menüs Folgendes aus:

Diagramme > **Diagrammerstellung...**

Das Dialogfeld "Diagrammerstellung" ist ein interaktives Fenster, in dem Sie beim Erstellen in der Vorschau verfolgen können, wie das Diagramm aussieht.

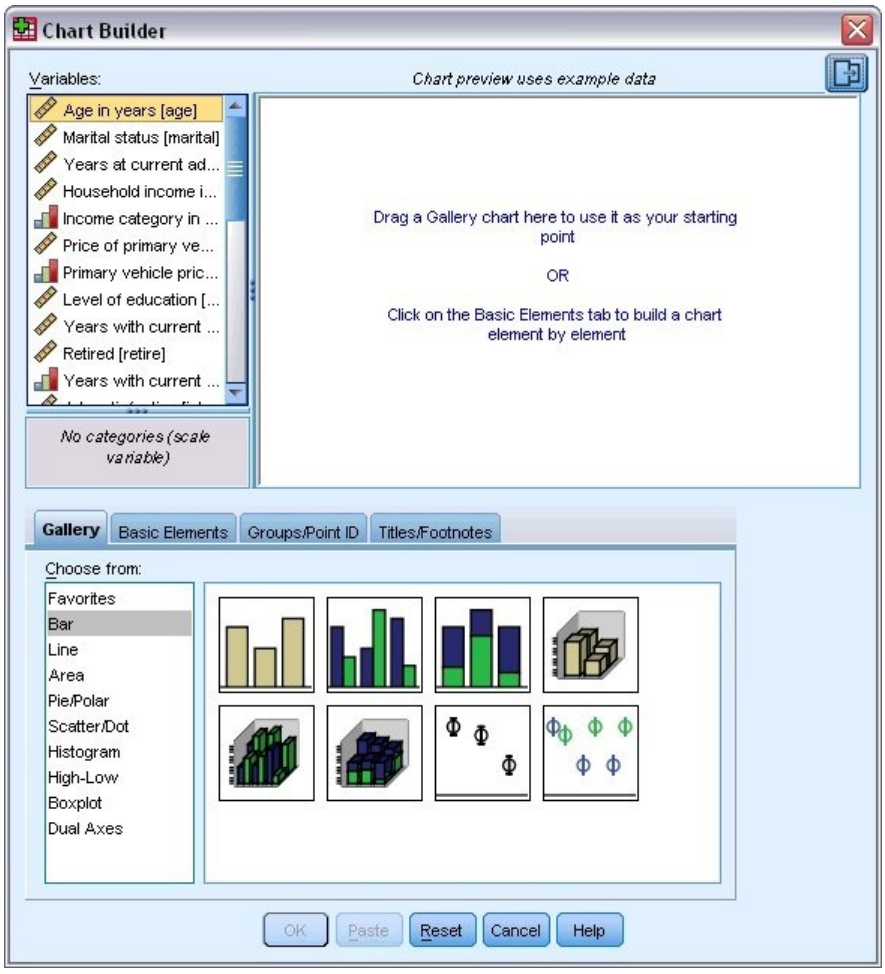


Abbildung 32. Dialogfeld "Diagrammerstellung"

Verwenden der Galerie für die Diagrammerstellung

1. Klicken Sie auf die Registerkarte **Galerie**, falls diese nicht bereits ausgewählt ist.

Die Galerie enthält zahlreiche verschiedene vordefinierte Diagramme, die nach Diagrammtypen angeordnet sind. Auf der Registerkarte "Grundelemente" stehen Ihnen auch Grundelemente wie Achsen und Grafikelemente zur Verfügung, mit denen Sie eigene Diagramme von Grund auf erstellen können. Die Galerie erleichtert Ihnen diese Arbeit jedoch.

2. Klicken Sie auf **Balken**, falls dies nicht bereits ausgewählt ist.

   Im Dialogfeld werden Symbole für die Balkendiagramme angezeigt, die in der Galerie verfügbar sind. Anhand der Bilder können Sie den jeweiligen Diagrammtyp erkennen. Wenn Sie weitere Informationen benötigen, können Sie mit dem Mauszeiger auf ein Symbol zeigen. Ihnen wird dann eine QuickInfo mit einer Beschreibung des Diagramms angezeigt.

3. Ziehen Sie das Symbol für das einfache Balkendiagramm in den Erstellungsbereich. Dies ist der große Bereich oberhalb der Galerie. Im Dialogfeld "Diagrammerstellung" wird eine Vorschau des Diagramms im Erstellungsbereich angezeigt. Beachten Sie, dass das Diagramm dabei nicht auf der Grundlage Ihrer Daten gezeichnet wird. Hierfür werden nur Beispieldaten verwendet.

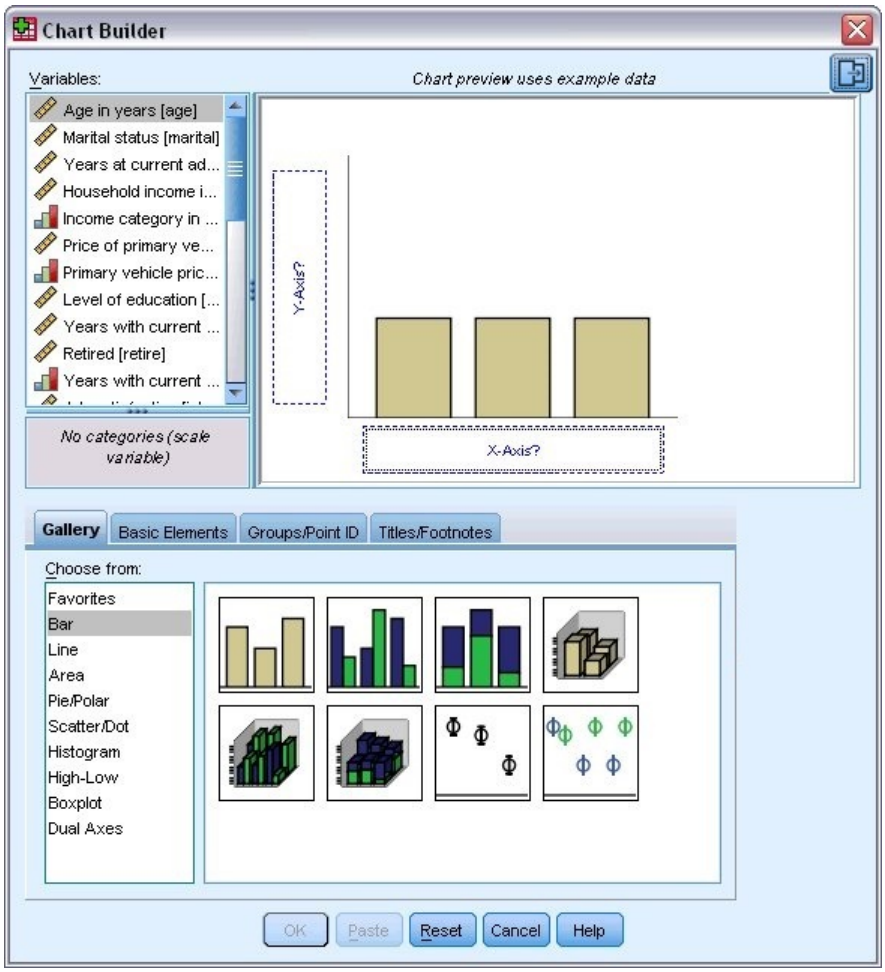


Abbildung 33. Balkendiagramm im Erstellungsbereich für die Diagrammerstellung

Definieren von Variablen und Statistiken

Das im Erstellungsbereich angezeigte Diagramm ist noch nicht vollständig, da es weder Variablen noch Statistiken enthält, die die Größe der Balken steuern und angeben, welche Variablenkategorie den einzelnen Balken entspricht. Sie können kein Diagramm ohne Variablen und Statistiken anlegen. Sie fügen Variablen hinzu, indem Sie diese aus der Variablenliste ziehen, die sich links vom Erstellungsbereich befindet.

Für die Diagrammerstellung ist das Messniveau einer Variablen wichtig. Sie möchten auf der *X*-Achse die Variable *Zufriedenheit mit der Arbeit* verwenden. Das Symbol (es ähnelt einem Lineal) neben der Variablen gibt jedoch an, dass das Messniveau als metrisch definiert ist. Sie müssen für ein fehlerfreies Diagramm

daher ein kategoriales Messniveau verwenden. Es ist nun nicht nötig, wieder in die Variablenansicht zu wechseln und das Messniveau zu ändern. Sie können das Messniveau vorübergehen im Dialogfeld "Diagrammerstellung" ändern.

1. Klicken Sie in der Liste "Variablen" mit der rechten Maustaste auf *Zufriedenheit mit der Arbeit* und wählen Sie **Ordinal** aus. "Ordinal" eignet sich als Messniveau, da die Kategorien in *Zufriedenheit mit der Arbeit* nach dem Zufriedenheitsgrad bewertet werden können. Beachten Sie, dass sich das Symbol nach der Änderung des Messniveaus ändert.
2. Ziehen Sie *Zufriedenheit mit der Arbeit* nun aus der Liste "Variablen" auf den Ablegebereich für die *X*-Achse.

   Der Ablegebereich für die *Y*-Achse ist standardmäßig auf die Statistik *Anzahl* festgelegt. Sie können jedoch problemlos eine andere Statistik wie Prozentsatz oder Mittelwert festlegen. Im vorliegenden Beispiel wird keine dieser Statistiken verwendet, dennoch wird das Verfahren für den Fall erläutert, dass Sie diese Statistik zu einem späteren Zeitpunkt ändern möchten.
3. Klicken Sie auf die Registerkarte **Elementeigenschaften** in der Seitenleiste der Diagrammerstellung. (Wird die Seitenleiste nicht angezeigt, klicken Sie auf die Schaltfläche in der rechten oberen Ecke der Diagrammerstellung, um die Seitenleiste anzuzeigen.)

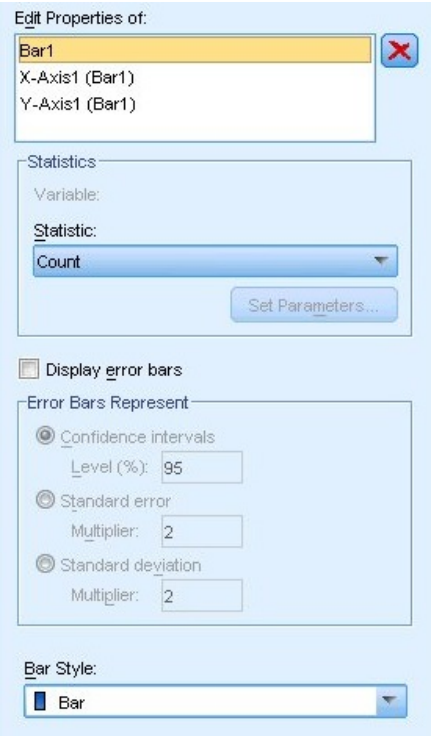


Abbildung 34. Elementeigenschaften

Mit "Elementeigenschaften" können Sie die Eigenschaften der verschiedenen Diagrammelemente ändern. Dabei handelt es sich um die Grafikelemente, z. B. die Balken im Balkendiagramm, und die Achsen des Diagramms. Wählen Sie in der Liste "Eigenschaften bearbeiten von" ein Element aus, um die zugehörigen Eigenschaften zu ändern. Beachten Sie auch das rote *X* rechts neben der Liste. Mit dieser Schaltfläche kann ein Grafikelement aus dem Erstellungsbereich gelöscht werden. Hier ist **Balken1** ausgewählt, daher beziehen sich die angezeigten Eigenschaften auf Grafikelemente, in diesem Fall das Grafikelement für den Balken.

In der Dropdown-Liste "Statistik" werden die verfügbaren spezifischen Statistiken angezeigt. Diese stehen in der Regel für alle Diagrammtypen zur Verfügung. Beachten Sie, dass es für einige Statistiken erforderlich ist, dass der Ablegebereich für die *Y*-Achse eine Variable enthält.

4. Ziehen Sie *Haushaltseinkommen in Tausend* aus der Liste *Variablen* auf den Ablegebereich für die Y-Achse. Da die *Y*-Achse eine metrische und die *X*-Achse eine kategoriale Variable enthält (ordinal ist ein Typ des kategorialen Messniveaus), ist der Ablegebereich der *Y*-Achse in der Standardeinstellung

auf die Statistik *Mittelwert* festgelegt. Dies sind die benötigten Variablen und Statistiken, Sie müssen die Elementeigenschaften daher nicht ändern.

Hinzufügen von Text

Sie können Titel und Fußnoten zum Diagramm hinzufügen.

1. Klicken Sie auf die Registerkarte **Titel/Fußnoten**.
2. Wählen Sie **Titel 1** aus.

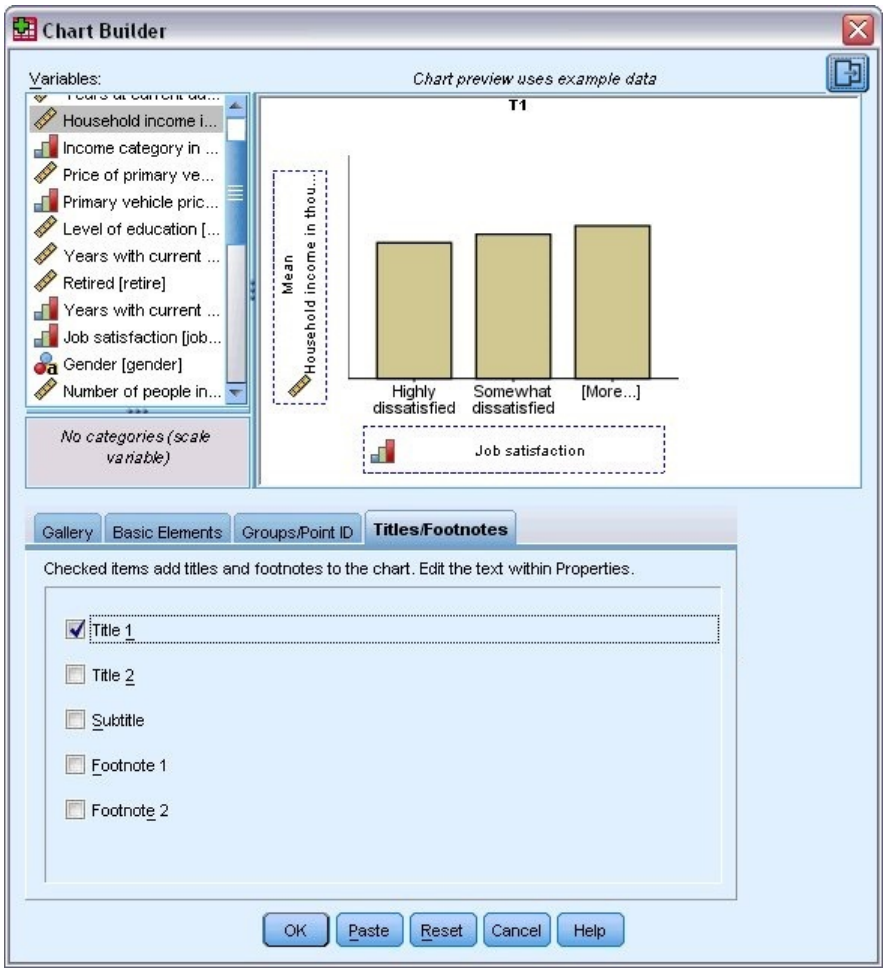


Abbildung 35. "Titel 1" im Erstellungsbereich

Der Titel wird mit der Beschriftung **T1** im Erstellungsbereich angezeigt.

3. Wählen Sie auf der Registerkarte **Elementeigenschaften** in der Liste "Eigenschaften bearbeiten von" den Eintrag **Titel 1** aus.
4. Geben Sie in das Textfeld Inhalt `Income by Job Satisfaction` ein. Dieser Text wird dann als Titel angezeigt.

Erstellen des Diagramms

1. Klicken Sie auf **OK**, um das Balkendiagramm zu erstellen.

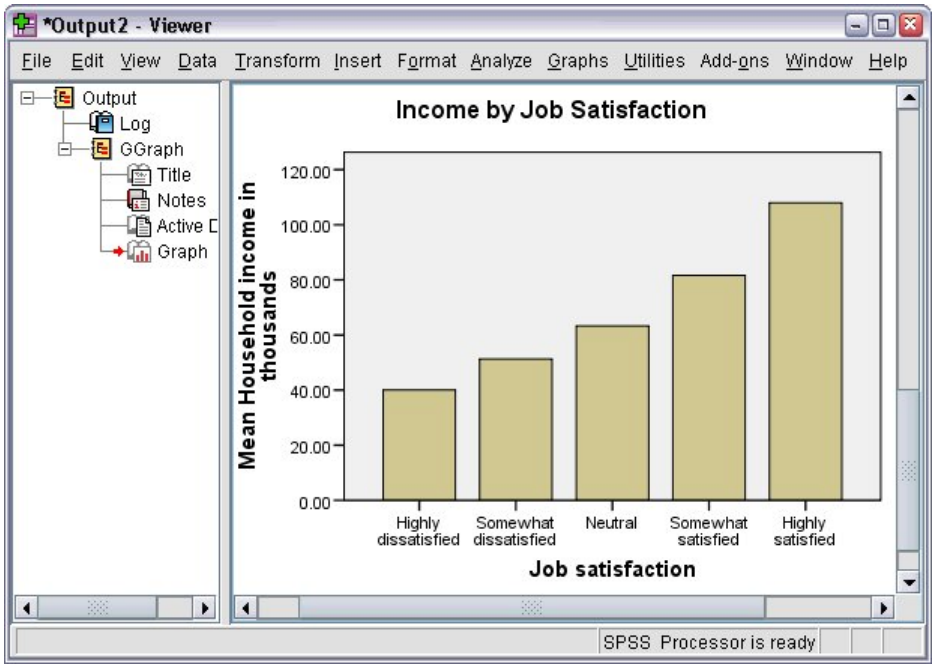


Abbildung 36. Balkendiagramm

Das Balkendiagramm zeigt, dass Befragte, die zufriedener mit ihrer Arbeit sind, in der Regel über ein höheres Haushaltseinkommen verfügen.

Kapitel 6. Arbeiten mit Ausgaben

Die Ergebnisse aus der Ausführung statistischer Prozeduren werden im Viewer angezeigt. Bei der Ausgabe kann es sich um Statistiktabellen, Diagramme, Grafiken oder Text handeln, je nachdem, welche Optionen beim Ausführen der Prozedur ausgewählt wurden. In diesem Abschnitt werden die Dateien *viewertut.spv* und *demo.sav* verwendet. Weitere Informationen finden Sie in Kapitel 10, „Stichprobendateien”, auf Seite 83.

Arbeiten mit dem Viewer

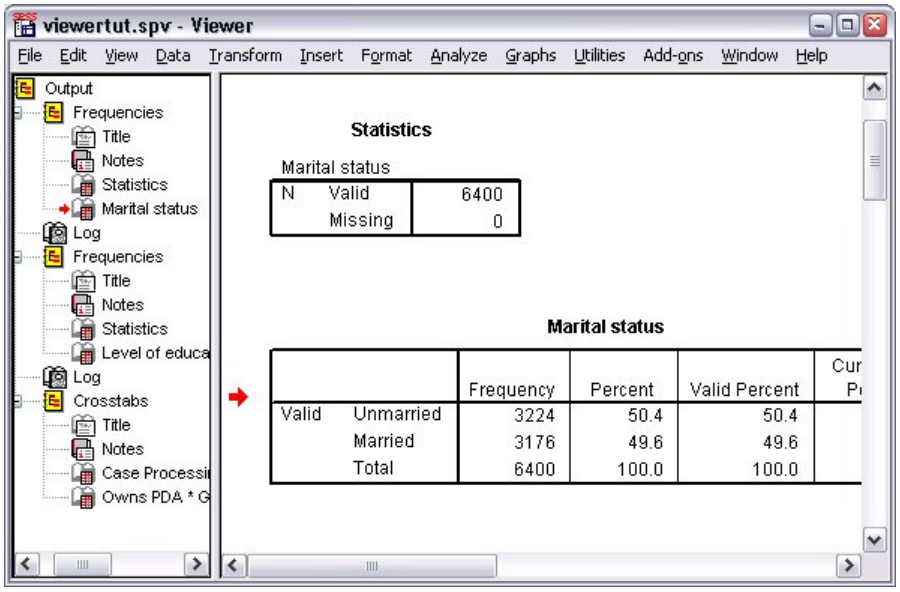


Abbildung 37. Viewer

Das Fenster "Viewer" ist in zwei Fensterbereiche unterteilt. Das **Gliederungsfenster** enthält eine Gliederung aller im Viewer enthaltenen Daten. Das **Inhaltsfenster** enthält Statistiktabellen, Diagramme und Textausgaben.

Mithilfe der Bildlaufleisten können Sie vertikal und horizontal durch den Inhalt des Fensters navigieren. Wenn Sie im Gliederungsfenster auf ein Element klicken, wird es im Inhaltsfenster angezeigt.

1. Klicken Sie auf den rechten Rahmen des Gliederungsfensters und ziehen Sie ihn auf die gewünschte Breite.

   Ein Symbol in Form eines offenen Buches im Gliederungsfenster weist darauf hin, dass das betreffende Element derzeit im Viewer angezeigt wird, obwohl es sich möglicherweise im gegenwärtig nicht sichtbaren Bereich des Inhaltsfensters befindet.
2. Wenn Sie eine Tabelle oder ein Diagramm ausblenden möchten, können Sie im Gliederungsfenster auf das entsprechende Buchsymbol doppelklicken.

   Das Symbol mit dem offenen Buch ändert sich in ein geschlossenes Buch, woran Sie erkennen, dass die zugeordneten Daten ausgeblendet sind.
3. Wenn die ausgeblendete Ausgabe wieder angezeigt werden soll, doppelklicken Sie auf das Symbol mit dem geschlossenen Buch.

   Außerdem können Sie die gesamte Ausgabe aus einer bestimmten statistischen Prozedur oder die gesamte Ausgabe im Viewer ausblenden.

4. Klicken Sie links neben der Prozedur, deren Ergebnisse ausgeblendet werden sollen, auf die Schaltfläche mit dem Minuszeichen (–) oder klicken Sie auf die Schaltfläche neben dem obersten Element der Gliederung, um die gesamte Ausgabe auszublenden.

   Die Gliederung wird reduziert, um optisch darauf hinzuweisen, dass die Ergebnisse nicht mehr angezeigt werden.

   Sie können die Anzeigereihenfolge für die Ausgabe auch ändern.
5. Klicken Sie im Gliederungsfenster auf die Elemente, die Sie verschieben möchten.
6. Ziehen Sie die ausgewählten Elemente an die neue Position in der Gliederung.

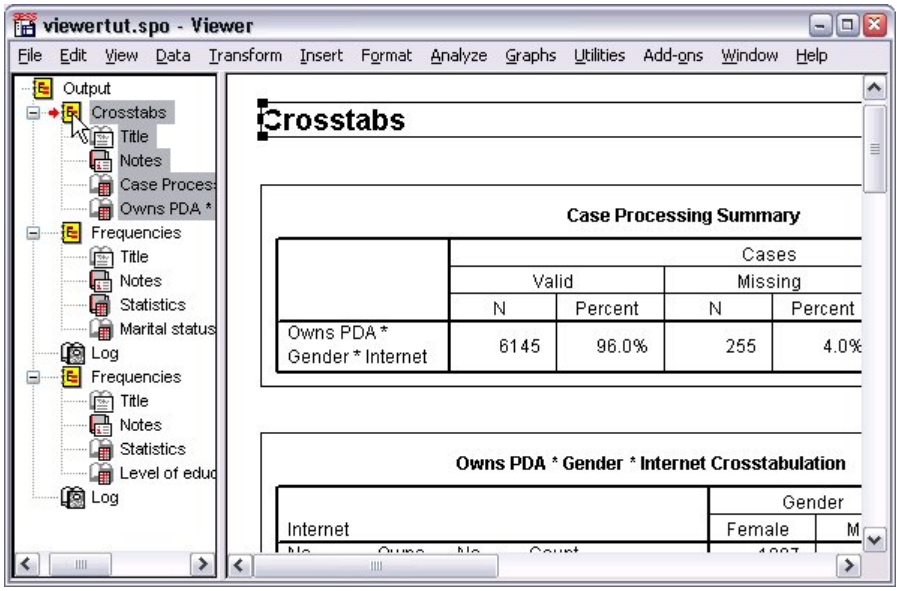


Abbildung 38. Neu geordnete Ausgabe im Viewer

Sie können Ausgabeelemente auch verschieben, indem Sie auf diese klicken und sie in das Inhaltsfenster ziehen.

Verwenden des Editors für Pivot-Tabellen

Die Ergebnisse der meisten statistischen Prozeduren werden in **Pivot-Tabellen** angezeigt.

Aufrufen von Ausgabedefinitionen

In der Ausgabe werden zahlreiche statistische Begriffe angezeigt. Auf die Definitionen dieser Begriffe können Sie direkt im Viewer zugreifen.

1. Doppelklicken Sie auf die Tabelle *Palm Pilot im Haushalt vorhanden * Geschlecht * Internetanschluss*.
2. Klicken Sie mit der rechten Maustaste auf *Expected Count* und wählen Sie **What's This?** aus dem Pop-up-Menü.

Die Definition wird in einem Popup-Fenster angezeigt.

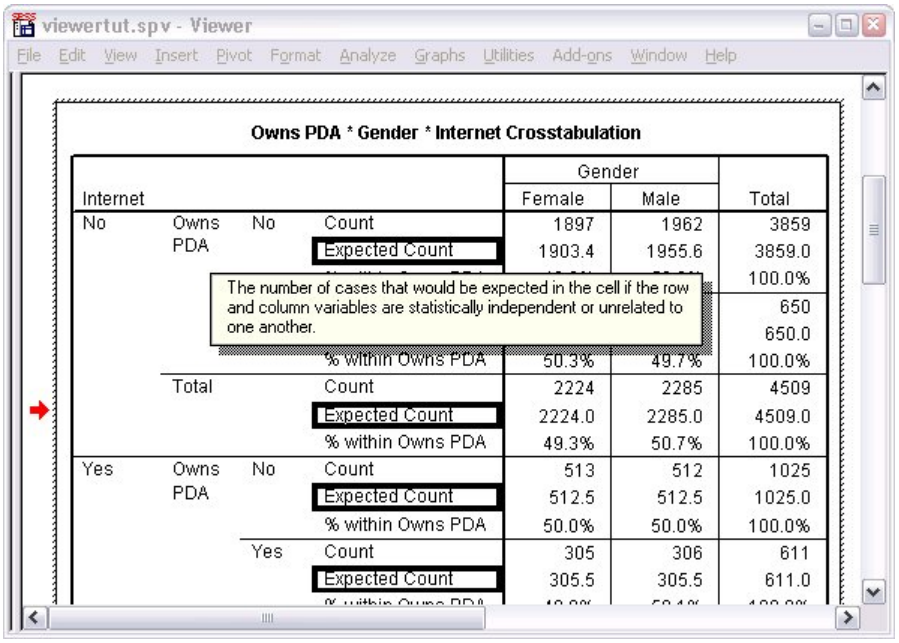


Abbildung 39. Popup-Definition

Pivot-Tabellen

Möglicherweise werden die Daten in den erstellten Standardtabellen unübersichtlich angezeigt. In Pivot-Tabellen können Sie Zeilen und Spalten transponieren (die Tabelle "kippen"), die Reihenfolge der Daten in einer Tabelle ändern und die Tabelle auf vielfältige Art und Weise bearbeiten. Durch Transponieren von Zeilen und Spalten wird z. B. aus einer kurzen, breiten Tabelle eine lange, schmale Tabelle. Das Ändern des Tabellenlayouts wirkt sich nicht auf die Ergebnisse aus. Es bietet nur eine Möglichkeit, Informationen auf andere bzw. angemessenere Weise darzustellen.

1. Falls sie noch nicht aktiviert ist, doppelklicken Sie auf die Tabelle *Palm Pilot im Haushalt vorhanden * Geschlecht * Internetanschluss*, um sie zu aktivieren.
2. Wenn das Fenster "Pivot-Leisten" ausgeblendet ist, wählen Sie die folgenden Menübefehle aus:

   Pivot > **Schwenkbare Fächer**

   Mithilfe von Pivot-Leisten können Daten zwischen Spalten, Daten und Schichten verschoben werden.

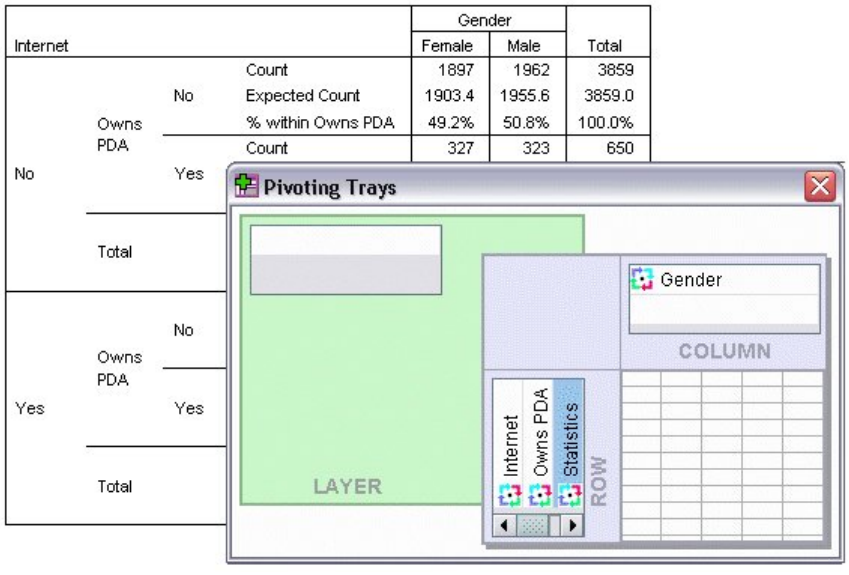


Abbildung 40. Pivot-Leisten

3. Ziehen Sie das Element *Statistik* aus der Zeilendimension in die Spaltendimension, unterhalb von *Geschlecht*. Die Tabelle wird unmittelbar neu konfiguriert und die Änderungen werden angezeigt.

   Die Reihenfolge der Elemente auf der Pivot-Leiste entspricht der Reihenfolge der Elemente in der Tabelle.

4. Ziehen Sie das Element *Palm Pilot im Haushalt vorhanden* vor das Element *Internetanschluss* in der Zeilendimension und legen Sie es dort ab, um die Reihenfolge dieser beiden Zeilen zu vertauschen.

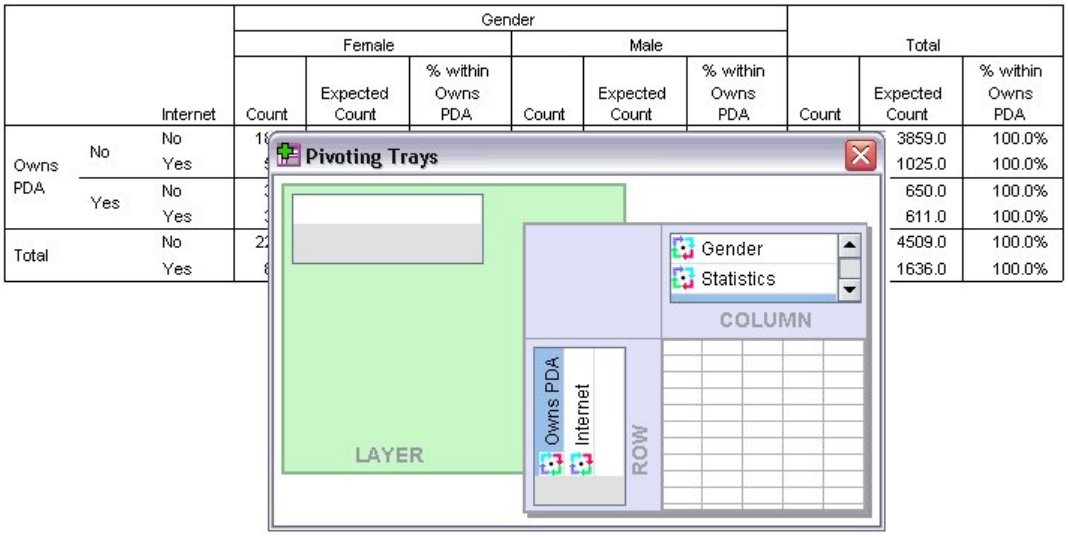


Abbildung 41. Vertauschen von Zeilen

Erstellen und Anzeigen von Schichten

Schichten können den Umgang mit großen Tabellen mit verschachtelten Kategorien von Informationen erleichtern. Durch das Erstellen von Schichten wird die Tabelle übersichtlicher und damit leichter verständlich.

1. Ziehen Sie das Element *Geschlecht* aus der Spaltendimension in die Schichtdimension.

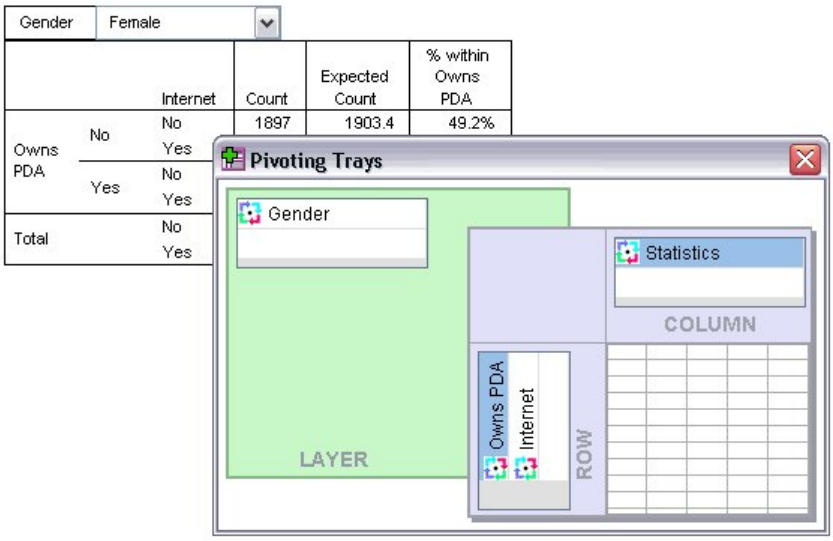


Abbildung 42. Pivot-Symbol "Geschlecht" in der Schichtdimension

Um eine andere Schicht anzuzeigen, wählen Sie eine Kategorie aus der Dropdown-Liste in der Tabelle aus.

Tabellen bearbeiten

Wenn Sie keine benutzerdefinierte Tabellenvorlage erstellt haben, werden Pivot-Tabellen mit Standardformatierung erstellt. Die Formatierung des gesamten Textes in einer Tabelle kann geändert werden. Sie können folgende Formatierungen ändern: Schriftartname, Schriftgröße, Schriftschnitt (fett oder kursiv) und Farbe.

1. Doppelklicken Sie auf die Tabelle *Schulabschluss*.
2. Wenn die Formatierungssymbolleiste ausgeblendet ist, wählen Sie die folgenden Menübefehle aus:

   Anzeigen > **Symbolleiste**
3. Klicken Sie auf den Titel *Schulabschluss*.
4. Wählen Sie in der Liste der Schriftgrößen **12** aus.
5. Um die Farbe des Titeltexts zu ändern, klicken Sie auf das Tool für die Textfarbe und wählen Sie eine neue Farbe aus.

Level of education

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Did not complete high school	1390	21.7	21.7	21.7
	High school degree	1936	30.3	30.3	52.0
	Some college	1360	21.3	21.3	73.2
	College degree	1355	21.2	21.2	94.4
	Post-undergraduate degree	359	5.6	5.6	100.0
	Total	6400	100.0	100.0	

Abbildung 43. Neu formatierter Titeltext in der Pivot-Tabelle

Sie können auch den Inhalt der Tabellen und Beschriftungen bearbeiten. Sie können beispielsweise den Titel dieser Tabelle ändern.

6. Doppelklicken Sie auf den Titel.
7. Geben Sie `Education Level` für das neue Etikett ein.

Hinweis: Wenn die Werte in einer Tabelle geändert werden, werden Gesamtergebnisse und andere Statistiken nicht neu berechnet.

Ausblenden von Zeilen und Spalten

Einige der Daten, die in einer Tabelle angezeigt werden, sind möglicherweise nicht nützlich oder können die Tabelle unnötig komplizieren. Sie können ganze Zeilen oder Spalten ausblenden, ohne dass Daten verlorengehen.

1. Falls sie noch nicht aktiviert ist, doppelklicken Sie auf die Tabelle *Ausbildungsniveau*, um sie zu aktivieren.
2. Klicken Sie auf die Spaltenbeschriftung *Gültige Prozente*, um diese Spalte auszuwählen.
3. Wählen Sie im Menü **Bearbeiten** oder im Popup-Menü mit der rechten Maustaste Folgendes aus:

   Auswählen > **Daten-und Beschriftungszellen**
4. Wählen Sie im Menü **Ansicht** die Option **Ausblenden** aus oder wählen Sie im Kontextmenü die Option **Kategorie ausblenden**aus.

   Die Spalte wird nun ausgeblendet, jedoch nicht gelöscht.

Education Level

		Frequency	Percent	Cumulative Percent
Valid	Did not complete high school	1390	21.7	21.7
	High school degree	1936	30.3	52.0
	Some college	1360	21.3	73.2
	College degree	1355	21.2	94.4
	Post-undergraduate degree	359	5.6	100.0
	Total	6400	100.0	

Abbildung 44. Spalte "Gültige Prozente" in der Tabelle ausgeblendet

5. Um die Spalte erneut anzuzeigen, doppelklicken Sie auf dieselbe Tabelle.
6. Klicken Sie auf **Anzeigen** und wählen Sie **Alle Kategorien anzeigen**aus.

   Die Zeilen können auch auf dieselbe Weise wie Spalten ausgeblendet und erneut angezeigt werden.

Ändern der Anzeigeformate für Daten

Sie können das Anzeigeformat für Daten in Pivot-Tabellen ganz einfach ändern.

1. Falls sie noch nicht aktiviert ist, doppelklicken Sie auf die Tabelle *Ausbildungsniveau*, um sie zu aktivieren.
2. Klicken Sie auf die Spaltenbeschriftung *Prozent*, um diese Spalte auszuwählen.
3. Wählen Sie im Menü "Bearbeiten" bzw. im Popup-Menü folgende Optionen aus:

   Auswählen > **Datenzellen**
4. Wählen Sie im Menü "Format" oder im Kontextmenü die Option **Zelleneigenschaften** aus.
5. Klicken Sie auf die Registerkarte **Formatwert**.
6. Geben Sie im Feld "Dezimalstellen" den Wert `0` ein, damit alle Dezimaltrennzeichen in dieser Spalte ausgeblendet werden.

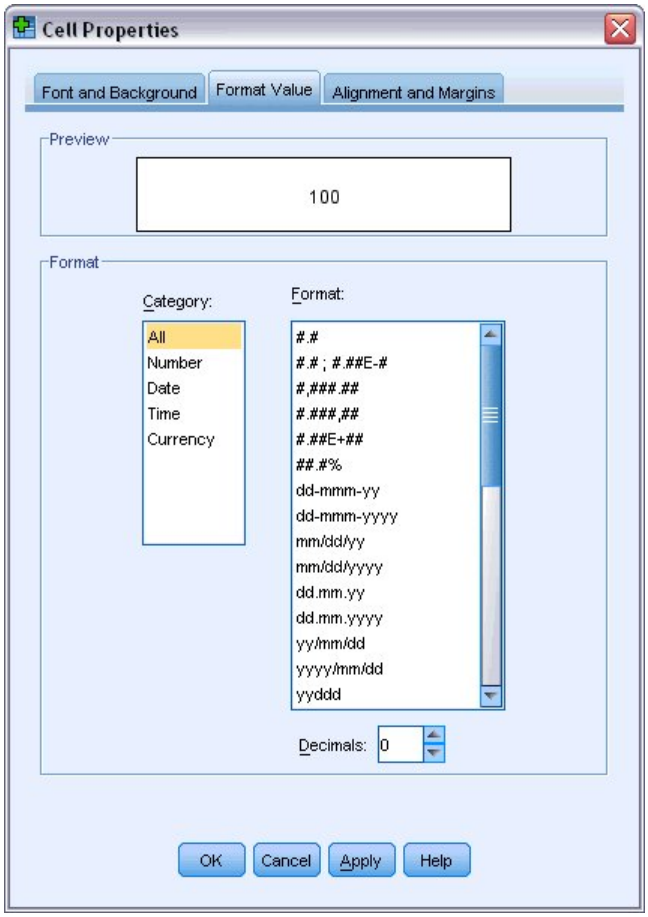


Abbildung 45. "Zelleneigenschaften", Registerkarte "Formatwert"

In diesem Dialogfeld können auch Datentyp und Format geändert werden.

7. Wählen Sie in der Liste "Kategorie" einfach den gewünschten Typ und anschließend in der Liste "Format" das Format für diesen Typ aus.
8. Klicken Sie auf **OK** oder **Zuweisen**, um Ihre Änderungen zu übernehmen.

Education Level

		Frequency	Percent	Cumulative Percent
Valid	Did not complete high school	1390	22	21.7
	High school degree	1936	30	52.0
	Some college	1360	21	73.2
	College degree	1355	21	94.4
	Post-undergraduate degree	359	6	100.0
	Total	6400	100	

Abbildung 46. Spalte "Prozent" mit ausgeblendeten Dezimalstellen

Die Dezimalstellen in der Spalte *Prozent* sind jetzt ausgeblendet.

TableLooks

Das Format Ihrer Tabellen ist der entscheidende Faktor für die Präsentation übersichtlicher, prägnanter und aussagekräftiger Ergebnisse. Wenn eine Tabelle unübersichtlich ist, sind die darin enthaltenen Informationen möglicherweise schwer verständlich.

Verwenden von vordefinierten Formaten

1. Doppelklicken Sie auf die Tabelle *Verheiratet*.

2. Wählen Sie in den Menüs Folgendes aus:

   Format > **TableLooks...**

   Im Dialogfeld "Tabellenvorlagen" wird eine Reihe vordefinierter Stile angezeigt. Wenn Sie einen Stil in der Liste markieren, wird im Vorschaufenster rechts daneben eine Vorschau angezeigt.

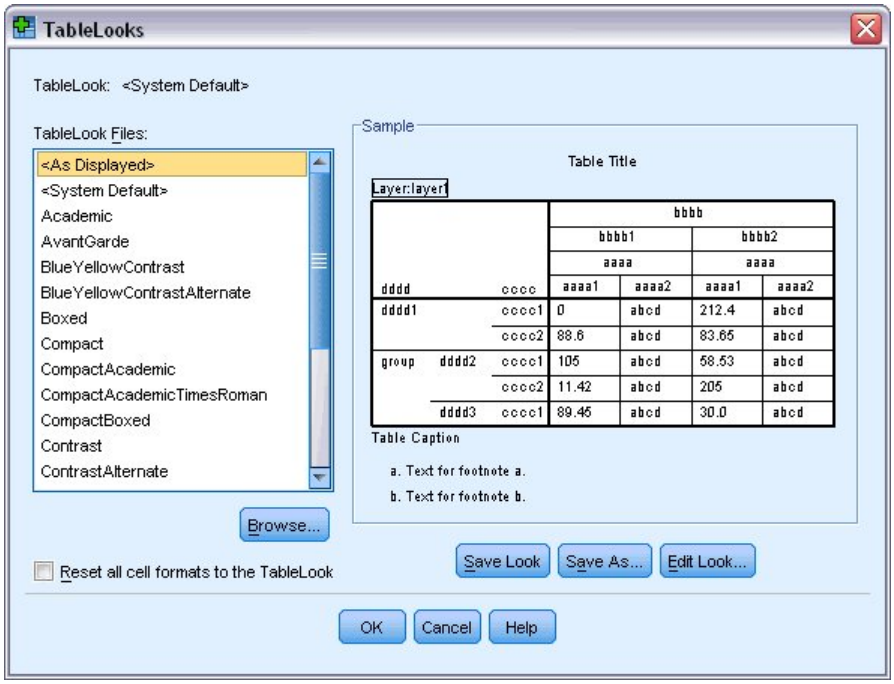


Abbildung 47. Dialogfeld "Tabellenvorlagen"

   Sie können den Stil unverändert übernehmen oder diesen nach Ihren Bedürfnissen anpassen.

3. Wenn Sie einen vorhandenen Stil verwenden möchten, wählen Sie ihn aus und klicken Sie auf **OK**.

Anpassen von Tabellenvorlagen

Sie können das Format an Ihre Bedürfnisse anpassen. Fast alle Aspekte einer Tabelle von der Vordergrundfarbe bis hin zur Rahmenart können angepasst werden.

1. Doppelklicken Sie auf die Tabelle *Verheiratet*.
2. Wählen Sie in den Menüs Folgendes aus:

   Format > **TableLooks...**

3. Wählen Sie den Stil aus, der dem gewünschten Format am ehesten entspricht, und klicken Sie auf **Vorlage bearbeiten**.
4. Wenn Sie auf die Registerkarte **Zellenformate** klicken, werden die Formatierungsoptionen angezeigt.

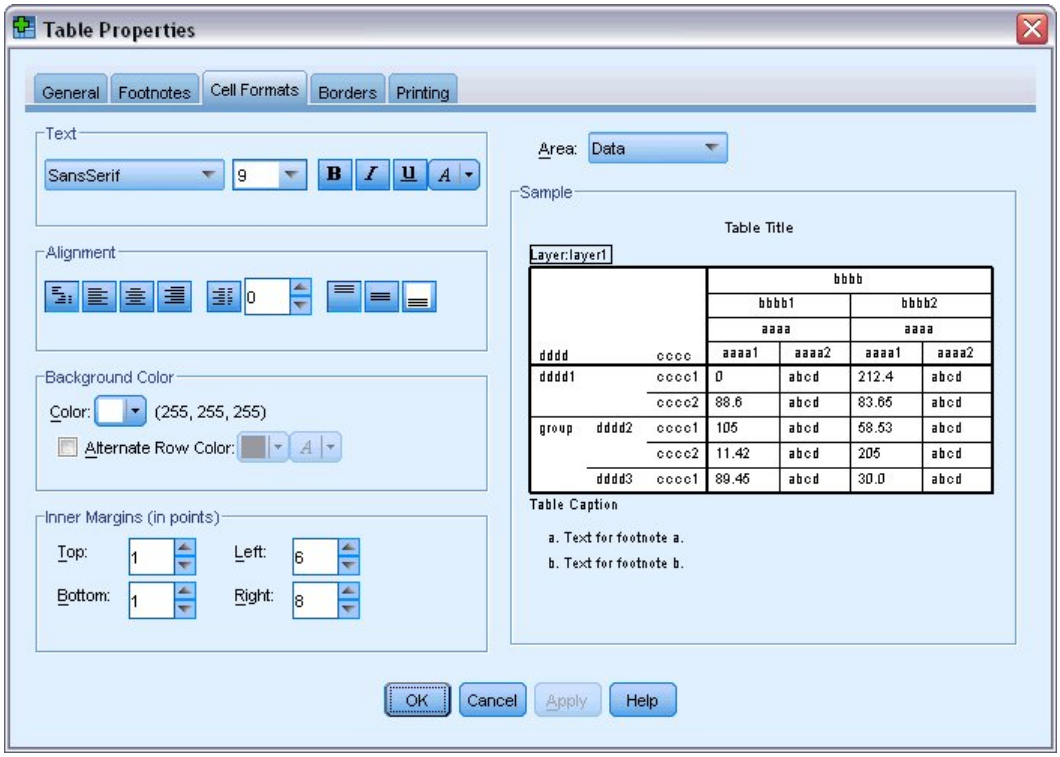


Abbildung 48. Dialogfeld "Tabelleneigenschaften"

Zu den Formatierungsoptionen gehören: Name der Schriftart, Schriftgröße, Schriftschnitt und Farbe. Weitere Optionen sind Ausrichtung, Farben für Text und Hintergrund und Randgrößen.

Im Vorschaufenster rechts können Sie sehen, wie sich die Änderungen an der Formatierung auf Ihre Tabelle auswirken. Jeder Bereich der Tabelle kann auf andere Weise formatiert werden. Der Titel soll sich von den angezeigten Daten vermutlich unterscheiden. Um einen Tabellenbereich für die Bearbeitung auszuwählen, können Sie entweder dessen Namen in der Dropdown-Liste "Bereich" auswählen oder im Vorschaufenster auf den zu ändernden Bereich klicken.

5. Wählen Sie in der Dropdown-Liste "Bereich" die Option **Daten** aus.
6. Wählen Sie aus der Dropdown-Palette "Hintergrund" einen neue Farbe aus.
7. Wählen Sie anschließend eine neue Textfarbe aus.

   Der neue Stil wird im Vorschaufenster angezeigt.

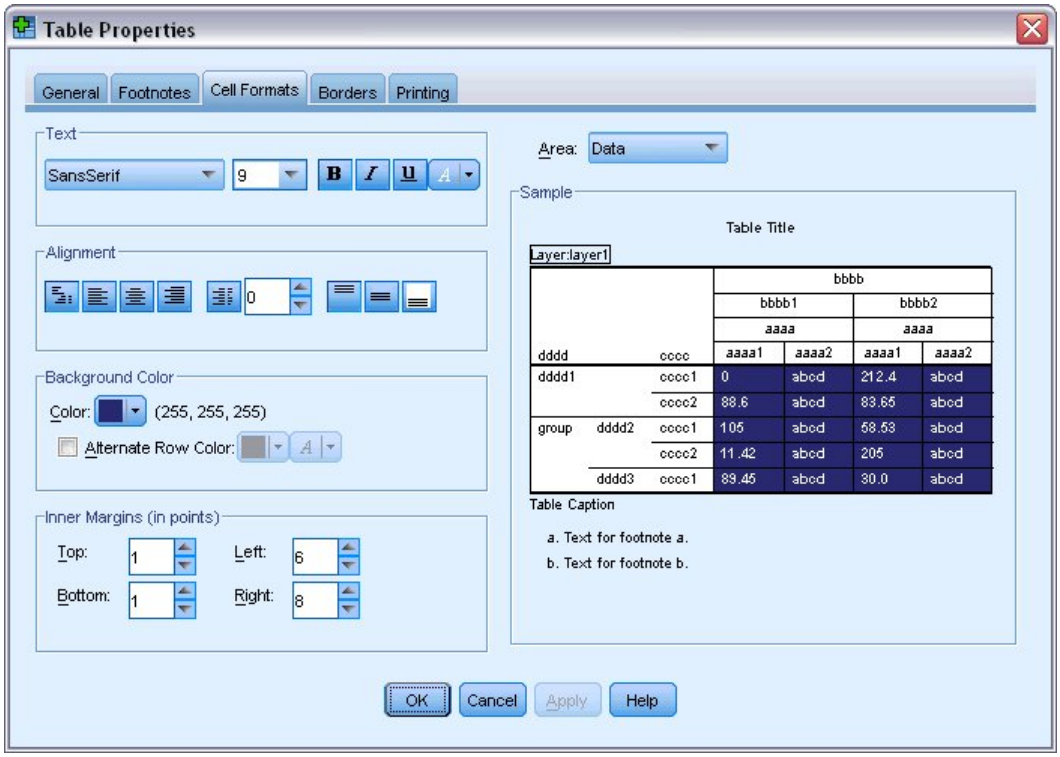


Abbildung 49. Ändern der Zellenformate von Tabellen

8. Klicken Sie auf **OK**, um zum Dialogfeld "Tabellenvorlagen" zurückzukehren.

   Sie können den neuen Stil speichern, damit Sie ihn für neue Tabellen wiederverwenden können.
9. Klicken Sie auf **Speichern unter**.
10. Wechseln Sie in das Zielverzeichnis und geben Sie im Textfeld "Dateiname" einen Namen für den neuen Stil ein.
11. Klicken Sie auf **Speichern**.
12. Klicken Sie auf **OK**, um die Änderungen zu übernehmen und zum Viewer zurückzukehren.

Die Tabelle enthält jetzt die angegebenen benutzerdefinierten Formatierungen.

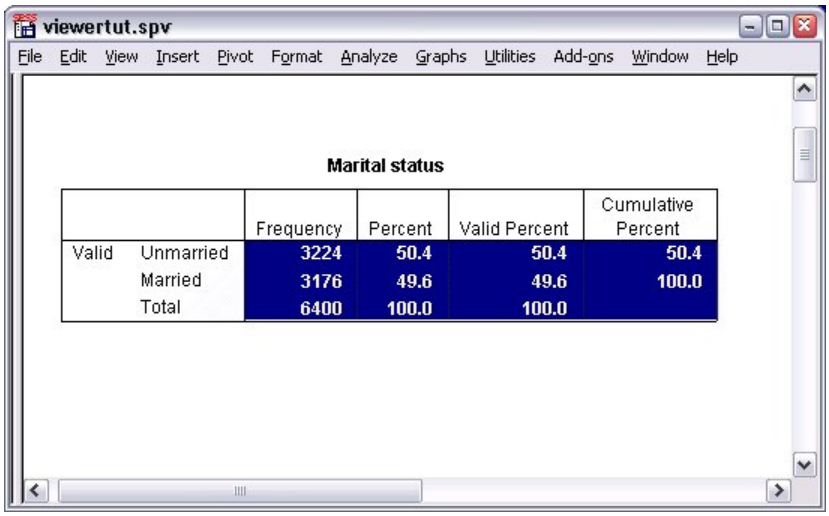


Abbildung 50. Benutzerdefinierte Tabellenvorlage

Ändern der Standardtabellenformate

Auch wenn Sie die Möglichkeit haben, das Format einer Tabelle nach ihrer Erstellung zu ändern, kann es effizienter sein, das Standardtabellenformat zu ändern, sodass Sie nicht bei jeder Erstellung einer Tabelle das Format ändern müssen.

Wenn Sie die Standardtabellenvorlage für Pivot-Tabellen ändern möchten, wählen Sie die folgenden Menübefehle aus:

Bearbeiten > **Optionen...**

1. Klicken Sie im Dialogfeld "Optionen" auf die Registerkarte **Pivot-Tabellen**.

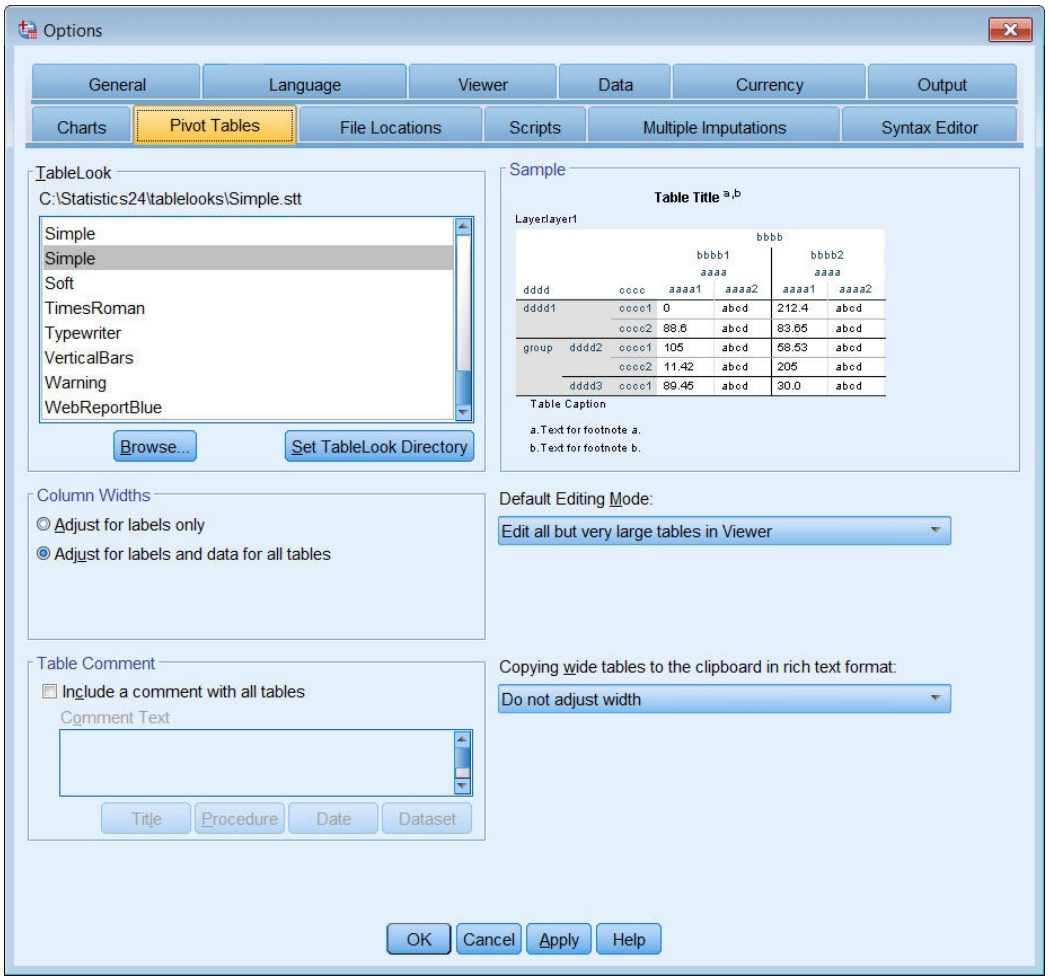


Abbildung 51. Dialogfeld "Optionen"

2. Wählen Sie für die Tabellenvorlage den Stil aus, den Sie für alle neuen Tabellen verwenden möchten.

   Im Vorschaufenster rechts können Sie eine Vorschau der verschiedenen Tabellenvorlagen anzeigen.
3. Klicken Sie auf **OK**, um die Einstellungen zu speichern und das Dialogfeld zu schließen.

Alle Tabellen, die nach der Änderung der Standardtabellenvorlage erstellt werden, entsprechen automatisch den neuen Formatierungsvorschriften.

Ändern der anfänglichen Einstellungen für die Anzeige

Zu den ursprünglichen Anzeigeeinstellungen gehören die Ausrichtung der Objekte im Viewer, die Breite des Fensters "Viewer" und die Angabe, ob Objekte standardmäßig ein- oder ausgeblendet werden. So ändern Sie diese Einstellungen:

1. Wählen Sie in den Menüs Folgendes aus:

Bearbeiten > **Optionen...**

2. Klicken Sie auf die Registerkarte **Viewer**.

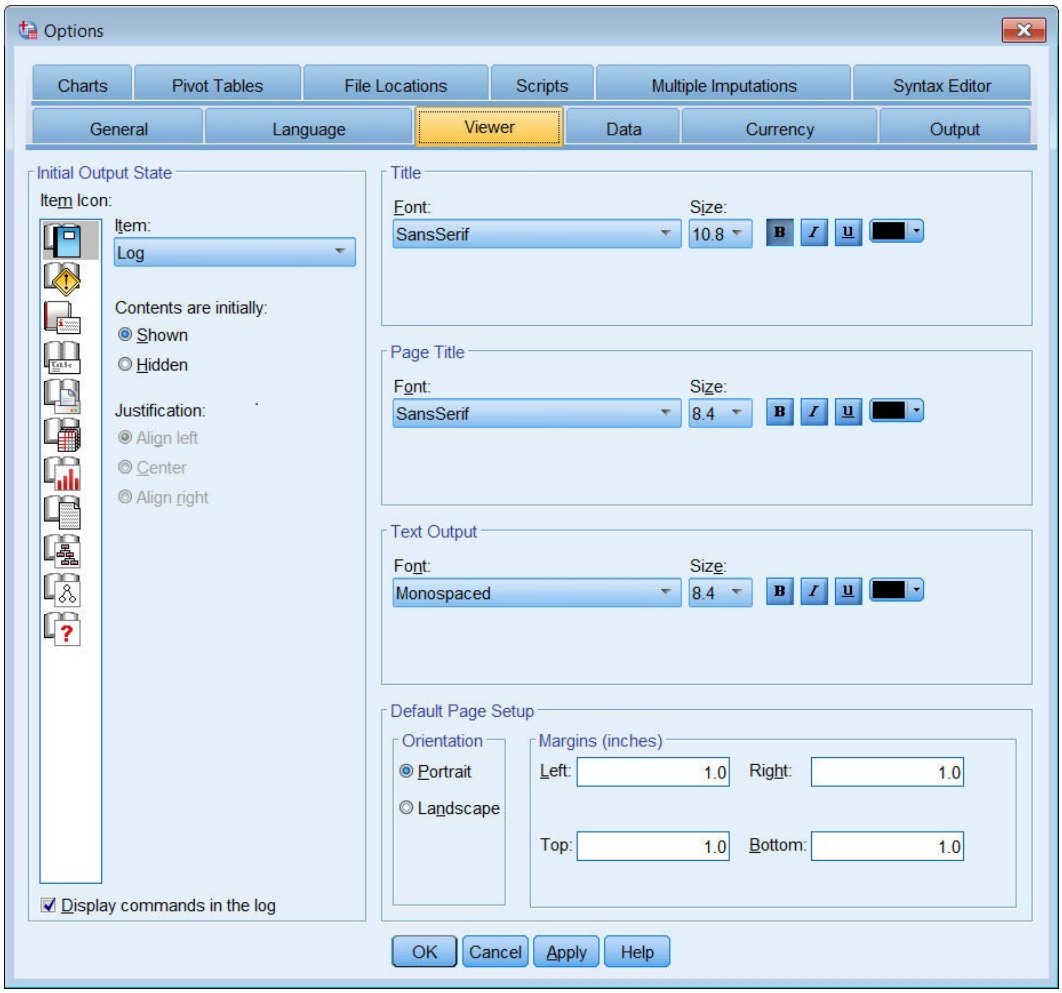


Abbildung 52. Viewer-Optionen

Die Einstellungen werden für jedes Objekt einzeln festgelegt. Sie können beispielsweise die Anzeige von Diagrammen anpassen, ohne damit die Form zu ändern, in der die Tabellen angezeigt werden. Wählen Sie einfach das Objekt aus, das Sie anpassen möchten, und nehmen Sie die Änderungen vor.

3. Klicken Sie auf das Symbol **Titel**, um dessen Einstellungen anzeigen zu lassen.
4. Klicken Sie auf **Zentriert**, wenn alle Titel im Viewer horizontal zentriert angezeigt werden sollen.

   Elemente wie Protokolle und Warnungen, die die Ausgabe häufig unübersichtlich machen, können ebenfalls ausgeblendet werden. Wenn Sie auf ein Symbol doppelklicken, werden die Anzeigeeigenschaften des entsprechenden Objekts automatisch geändert.
5. Wenn Sie auf das Symbol **Warnungen** doppelklicken, werden Warnungen in der Ausgabe ausgeblendet.
6. Klicken Sie auf **OK**, um die Änderungen zu speichern und das Dialogfeld zu schließen.

Anzeige von Variablen- und Wertbeschriftungen

Das Anzeigen der Variablen- und Wertbeschriftungen ist meist wirkungsvoller als das Anzeigen des Variablennamens oder des eigentlichen Datenwerts. In manchen Fällen kann es jedoch günstig sein, sowohl die Namen als auch die Beschriftungen anzuzeigen.

1. Wählen Sie in den Menüs Folgendes aus:

   Bearbeiten > **Optionen...**

2. Klicken Sie auf die Registerkarte **Beschriftung der Ausgabe**.

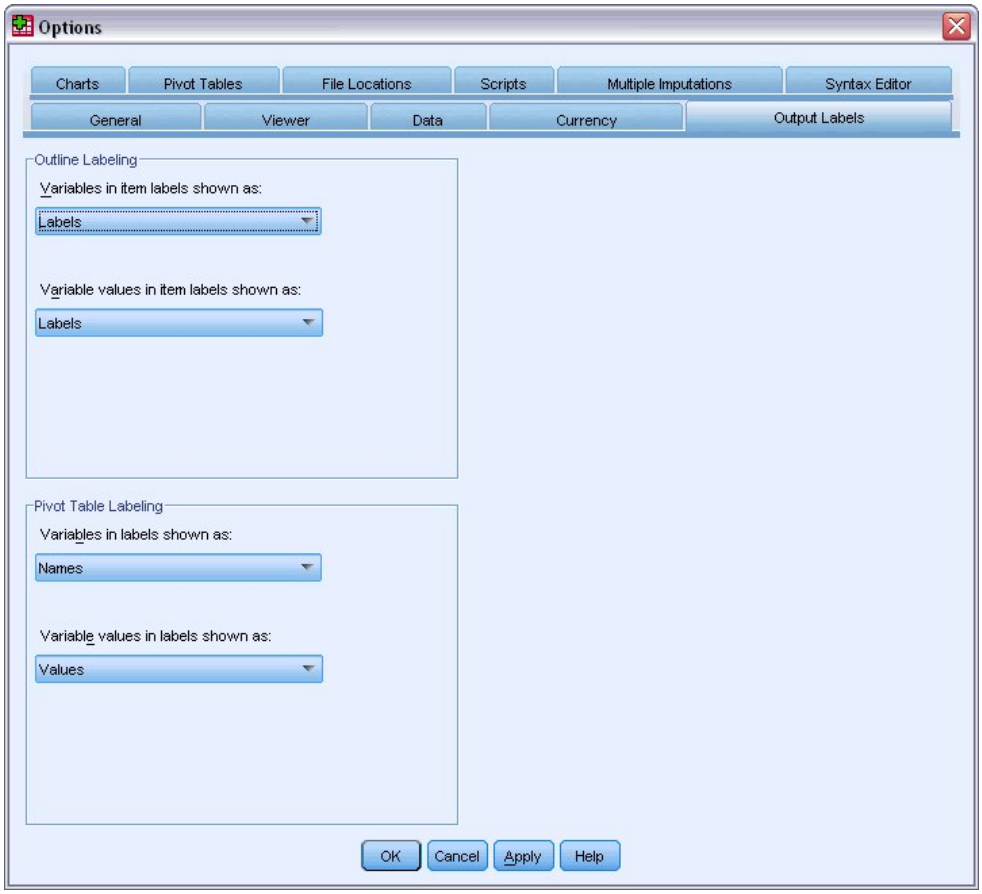


Abbildung 53. Einstellungen für die Beschriftung der Pivot-Tabellen

Sie können unterschiedliche Einstellungen für das Gliederungs- und das Inhaltsfenster angeben. So lassen Sie beispielsweise im Gliederungsfenster Beschriftungen und im Inhaltsfenster Variablennamen und Datenwerte anzeigen:

3. Wählen Sie im Gruppenfeld "Beschriftung für Pivot-Tabellen" aus der Dropdown-Liste "Variablen in Beschriftungen anzeigen als" die Option **Namen** aus, um Variablennamen anstelle von Beschriftungen anzuzeigen.
4. Wählen Sie anschließend aus der Dropdown-Liste "Variablenwerte in Beschriftungen anzeigen als:" die Option **Werte** aus, um Datenwerte anstelle von Beschriftungen anzuzeigen.

In den weiteren in dieser Sitzung erstellten Tabellen werden diese Änderungen berücksichtigt.

marital

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	3224	50.4	50.4	50.4
	1	3176	49.6	49.6	100.0
	Total	6400	100.0	100.0	

Abbildung 54. Anzeigen von Variablennamen und -werten

Verwenden von Ergebnissen in anderen Anwendungen

Die Ergebnisse können in vielen Anwendungen verwendet werden. Sie könnten zum Beispiel eine Tabelle oder ein Diagramm in eine Präsentation oder einen Bericht einfügen.

Die folgenden Beispiele beziehen sich auf Microsoft Word, sind aber möglicherweise auch auf andere Textverarbeitungsprogramme anwendbar.

Einfügen von Ergebnissen als Tabellen in Word

Sie können Pivot-Tabellen in Word als echte Word-Tabellen einfügen. Alle Tabellenattribute wie Schriftgrößen und Farben werden beibehalten. Da die Tabelle im Word-Tabellenformat eingefügt wird, können Sie sie in Word genau wie jede andere Tabelle bearbeiten.

1. Klicken Sie im Viewer auf eine Pivot-Tabelle, um sie auszuwählen.
2. Wählen Sie in den Menüs Folgendes aus:

   Bearbeiten > **Kopieren**
3. Öffnen Sie das Textverarbeitungsprogramm.
4. Wählen Sie in den Menüs der Textverarbeitung Folgendes aus:

   Bearbeiten > **Spezial einfügen...**
5. Wählen Sie im Dialogfeld "Inhalte einfügen" die Option **Formatierten Text (RTF)** aus.
6. Klicken Sie auf **OK**, um die Ergebnisse in das aktuelle Dokument einzufügen.

Die Tabelle wird nun im Dokument angezeigt. Sie können das Format anpassen, die Daten bearbeiten und die Größe der Tabelle nach Ihren Wünschen ändern.

Einfügen von Ergebnissen als Text

Pivot-Tabellen können in andere Anwendungen als einfacher Text kopiert werden. Bei diesem Verfahren werden zwar keine Formatierungen übernommen, Sie können die Tabellendaten jedoch nach dem Einfügen in die Zielanwendung bearbeiten.

1. Klicken Sie im Viewer auf eine Pivot-Tabelle, um sie auszuwählen.
2. Wählen Sie in den Menüs Folgendes aus:

   Bearbeiten > **Kopieren**
3. Öffnen Sie das Textverarbeitungsprogramm.
4. Wählen Sie in den Menüs der Textverarbeitung Folgendes aus:

   Bearbeiten > **Spezial einfügen...**
5. Wählen Sie im Dialogfeld "Inhalte einfügen" die Option **Unformatierter Text** aus.
6. Klicken Sie auf **OK**, um die Ergebnisse in das aktuelle Dokument einzufügen.

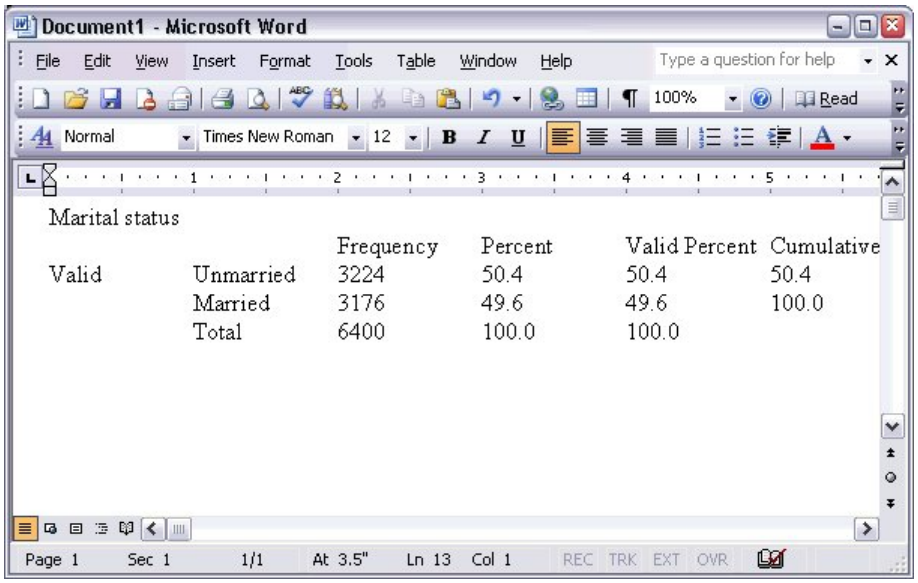


Abbildung 55. In Word angezeigte Pivot-Tabelle

Die Spalten der Tabelle sind durch Tabulatoren getrennt. Sie können die Spaltenbreiten ändern, indem Sie die Tabstops in der Textverarbeitung entsprechend korrigieren.

Exportieren von Ergebnissen in Microsoft Word-, PowerPoint- und Excel-Dateien

Sie können Ergebnisse in eine Microsoft Word-, PowerPoint- oder Excel-Datei exportieren. Sie können ausgewählte Elemente oder auch alle Elemente im Viewer exportieren. In diesem Abschnitt werden die Dateien *msouttut.spv* und *demo.sav* verwendet. Weitere Informationen finden Sie in Kapitel 10, „Stichprobendateien", auf Seite 83.

Hinweis: Der Export nach PowerPoint ist nur unter Windows-Betriebssystemen und nicht in der Studentenversion verfügbar.

Im Gliederungsfenster des Viewers können Sie bestimmte Elemente für den Export auswählen oder alle Elemente bzw. alle sichtbaren Elemente exportieren.

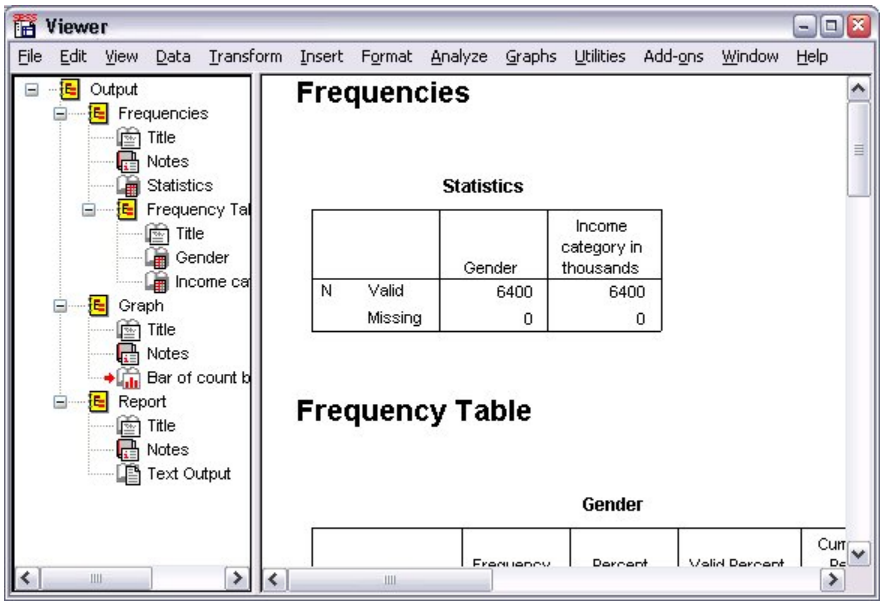


Abbildung 56. Viewer

1. Wählen Sie in den Menüs des Viewers Folgendes aus:

   Datei > **Exportieren...**

   Anstatt alle Objekte im Viewer zu exportieren, können Sie auch nur sichtbare Objekte (geöffnete Bücher im Gliederungsfenster) exportieren oder nur die Objekte, die Sie im Gliederungsfenster ausgewählt haben. Wenn Sie im Gliederungsfenster keine Elemente ausgewählt haben, steht Ihnen die Option zum Export ausgewählter Objekte nicht zur Verfügung.

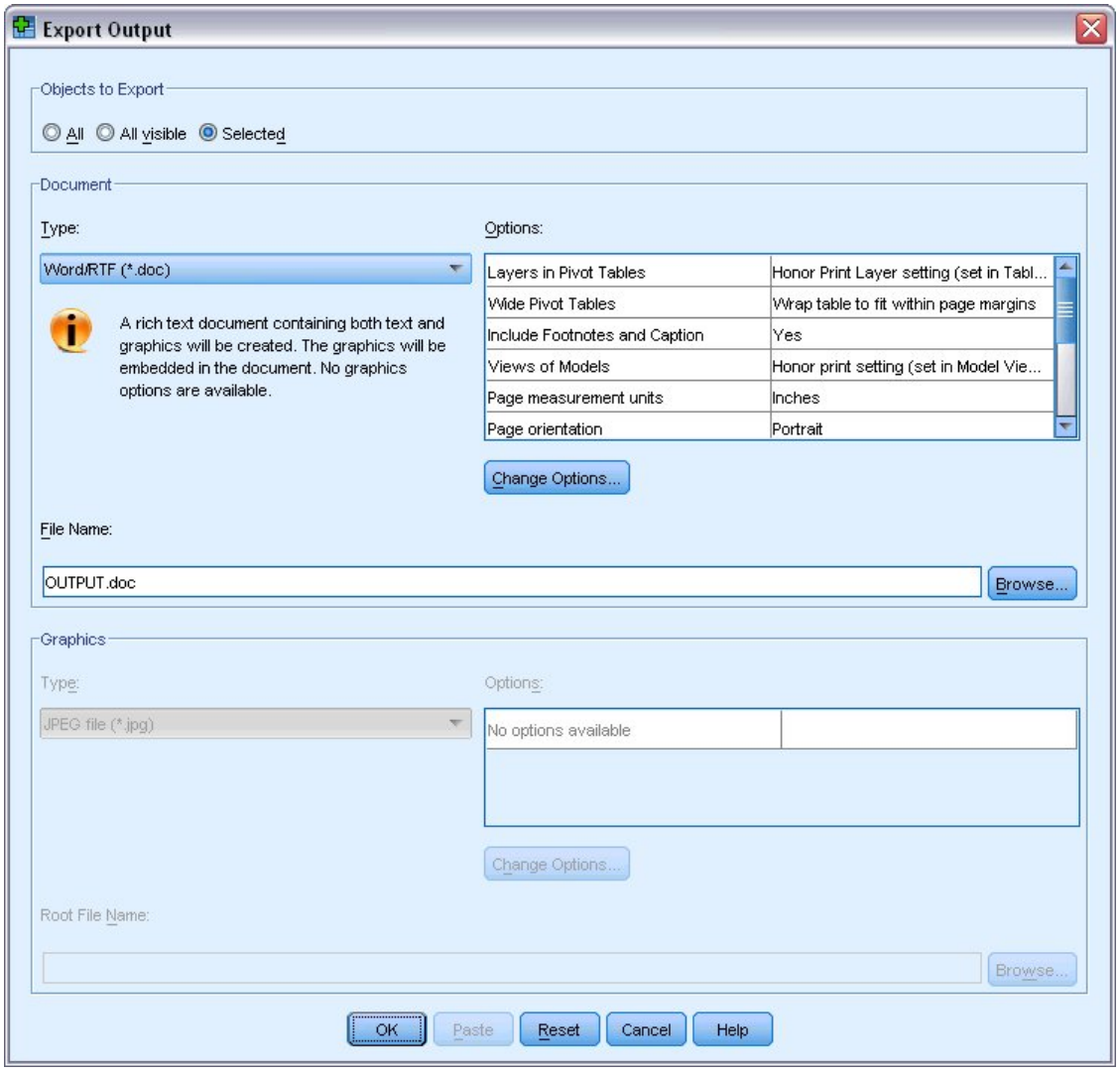


Abbildung 57. Dialogfeld "Ausgabe exportieren"

2. Wählen Sie in der Gruppe "Zu exportierende Objekte" die Option **Alle** aus.
3. Wählen Sie in der Dropdown-Liste "Typ" die Option **Word/RTF-Datei (*.doc)** aus.
4. Klicken Sie auf **OK**, um die Word-Datei zu erstellen.

Wenn Sie die resultierende Datei in Word öffnen, können Sie sehen, wie die Ergebnisse exportiert wurden. Anmerkungen, die keine sichtbaren Objekte sind, werden in Word angezeigt, da Sie ausgewählt haben, dass alle Objekte exportiert werden sollen.

Aus Pivot-Tabellen werden Word-Tabellen. Dabei bleiben alle Formatierungen der ursprünglichen Pivot-Tabelle erhalten, einschließlich Schriftarten, Farben, Rahmen usw.

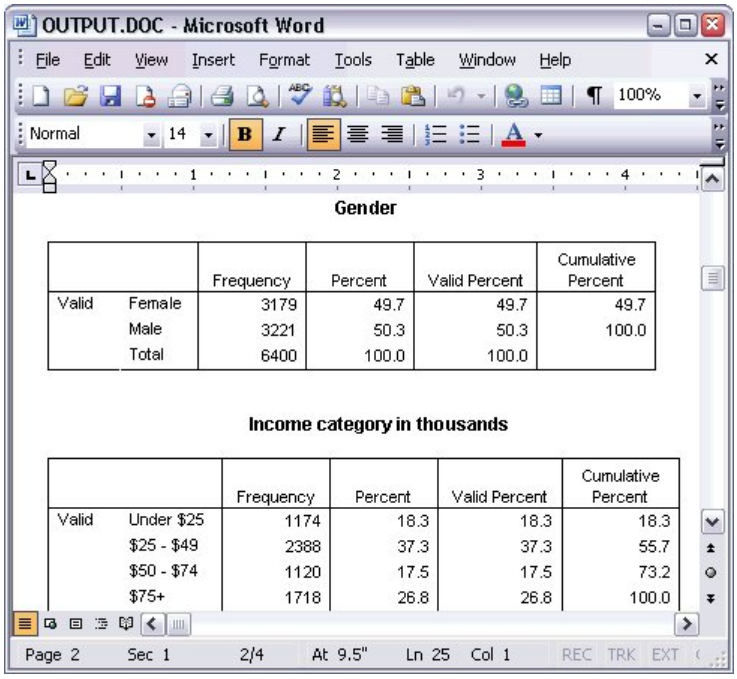


Abbildung 58. Pivot-Tabellen in Word

Diagramme werden als Grafiken in das Word-Dokument aufgenommen.

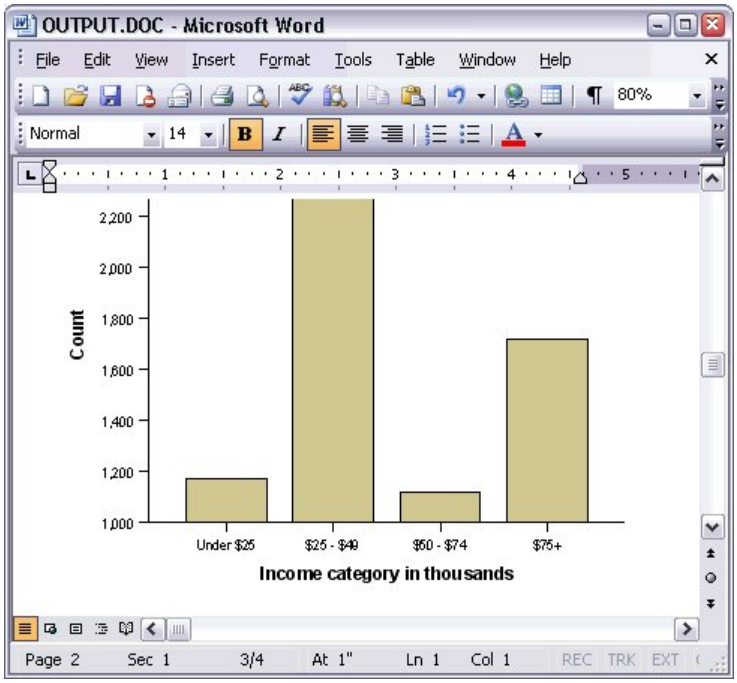


Abbildung 59. Diagramme in Word

Die Textausgabe wird in derselben Schriftart angezeigt, die auch für das Textobjekt im Viewer verwendet wird. Zur richtigen Ausrichtung sollte für Textausgaben ein Zeichensatz mit fester Zeichenbreite (festem Abstand) verwendet werden.

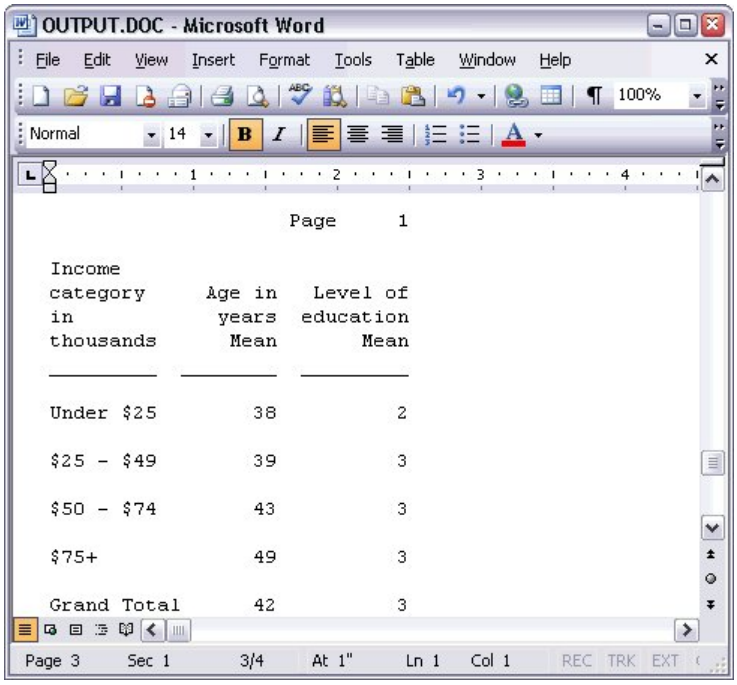


Abbildung 60. Textausgabe in Word

Beim Exportieren in eine PowerPoint-Datei werden die einzelnen exportierten Elemente jeweils auf einer separaten Folie platziert. Aus Pivot-Tabellen werden Word-Tabellen in PowerPoint. Dabei bleiben alle Formatierungen der ursprünglichen Pivot-Tabelle erhalten, einschließlich Schriftarten, Farben, Rahmen usw.

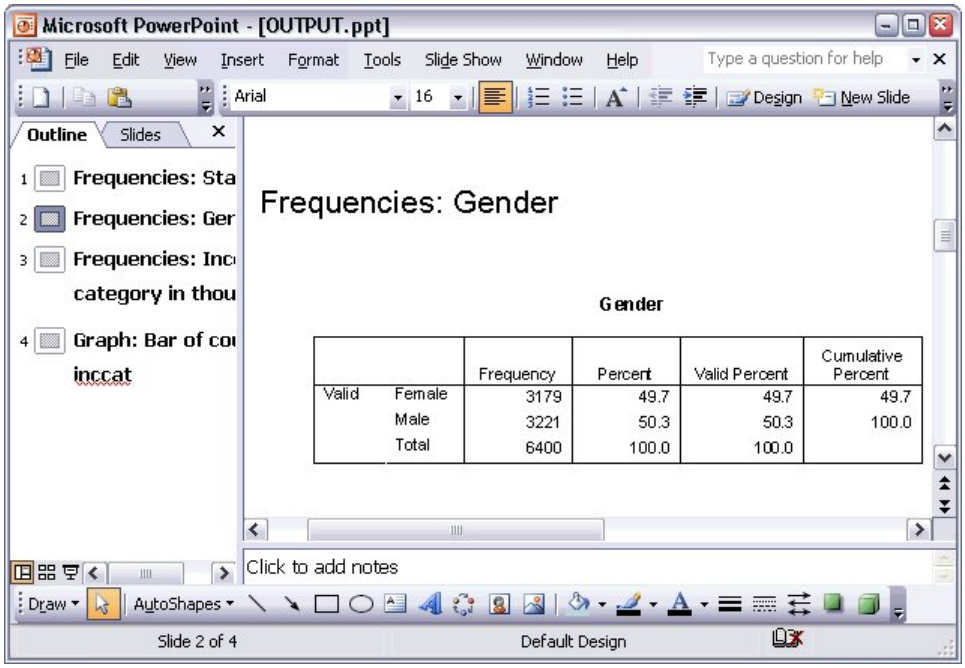


Abbildung 61. Pivot-Tabellen in PowerPoint

Die für den Export in PowerPoint ausgewählten Diagramme werden in die PowerPoint-Datei eingebettet.

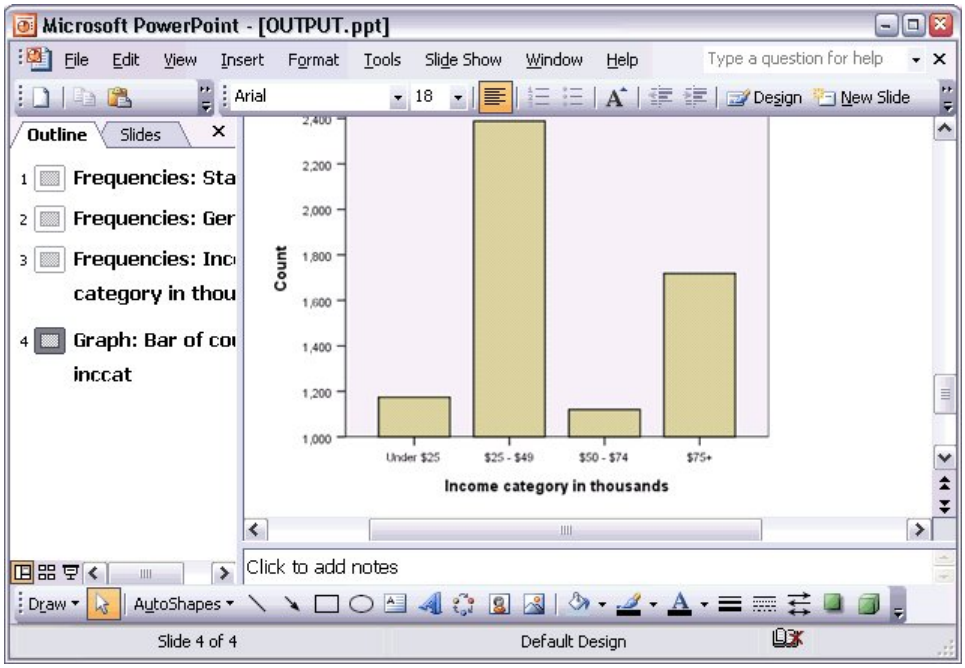


Abbildung 62. Diagramme in PowerPoint

Hinweis: Der Export nach PowerPoint ist nur unter Windows-Betriebssystemen und nicht in der Studentenversion verfügbar.

Beim Export in eine Excel-Datei werden die Ergebnisse in anderer Form exportiert.

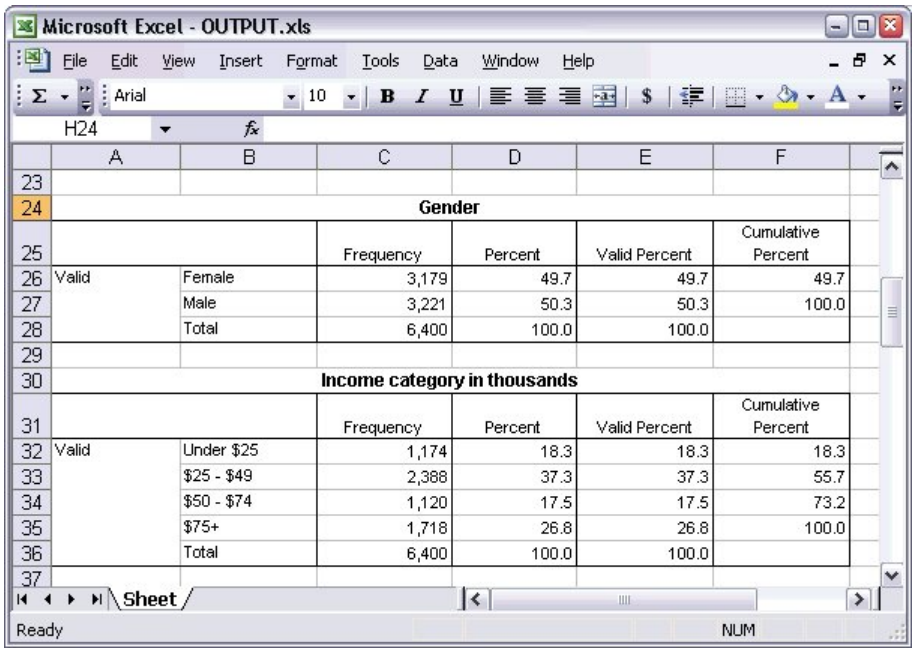


Abbildung 63. Output.xls in Excel

Die Zeilen, Spalten und Zellen von Pivot-Tabellen werden zu Excel-Zeilen, -Spalten und -Zellen.

Jede Zeile in der Textausgabe entspricht einer Zeile in der Excel-Datei, wobei der gesamte Inhalt der Zeile in einer einzelnen Zelle enthalten ist.

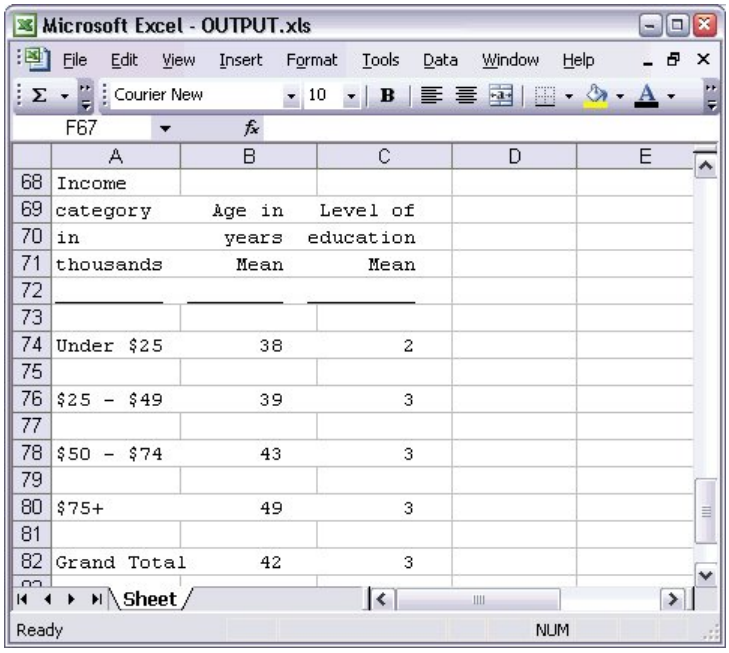


Abbildung 64. Textausgabe in Excel

Exportieren von Ergebnissen als PDF

Sie können alle Elemente oder ausgewählte Elemente im Viewer in eine PDF-Datei (Portable Document Format) exportieren.

1. Wählen Sie in den Menüs im Fenster "Viewer", das die Ergebnisse enthält, die Sie als PDF exportieren möchten, Folgendes aus:

   Datei > **Exportieren...**

2. Wählen Sie im Dialogfeld "Ausgabe exportieren" aus der Dropdown-Liste "Exportformat - Dateityp" die Option **Portable Document Format** aus.

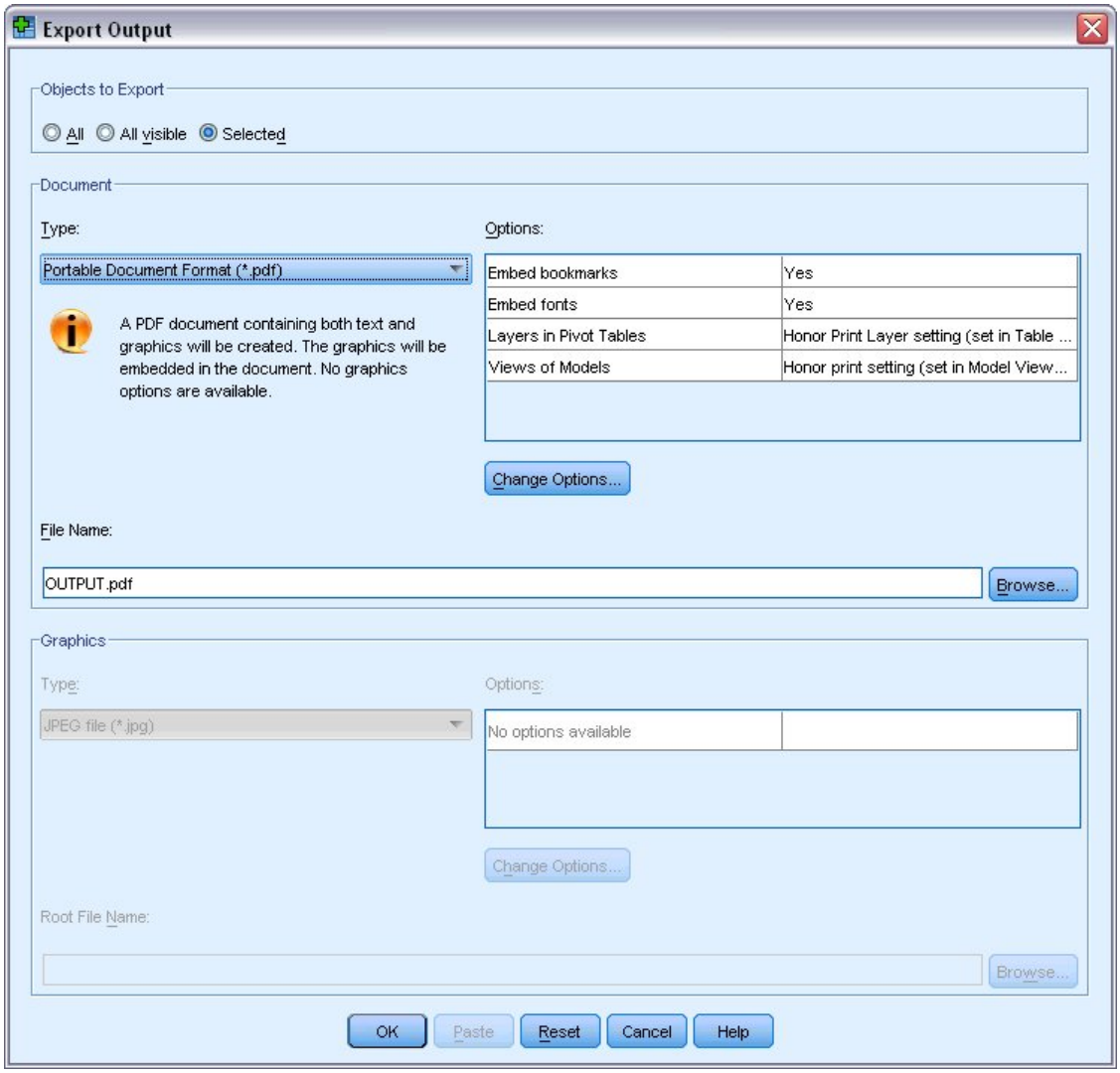


Abbildung 65. Dialogfeld "Ausgabe exportieren"

- Das Gliederungsfenster des Viewer-Dokuments wird in der PDF-Datei in Lesezeichen konvertiert, um die Navigation zu erleichtern.
- In PDF-Dokumenten werden Seitengröße, Ausrichtung, Ränder Inhalt und Anzeige von Kopf- und Fußzeilen sowie die Größe des gedruckten Diagramms über die Optionen für die Seiteneinrichtung (Menü "Datei", "Seite einrichten" im Fenster "Viewer") gesteuert.
- Die Auflösung (DPI) des PDF-Dokuments ist die aktuelle Auflösungseinstellung für den Standarddrucker bzw. den aktuell ausgewählten Drucker (kann über "Seite einrichten" geändert werden). Die maximale Auflösung beträgt 1200 DPI. Wenn eine höhere Druckerauflösung eingestellt ist, wird für das PDF-Dokument eine Auflösung von 1200 DPI verwendet. *Hinweis*: Dokumente mit höherer Auflösung können beim Drucken auf Druckern mit niedrigerer Auflösung zu schlechten Ergebnissen führen.

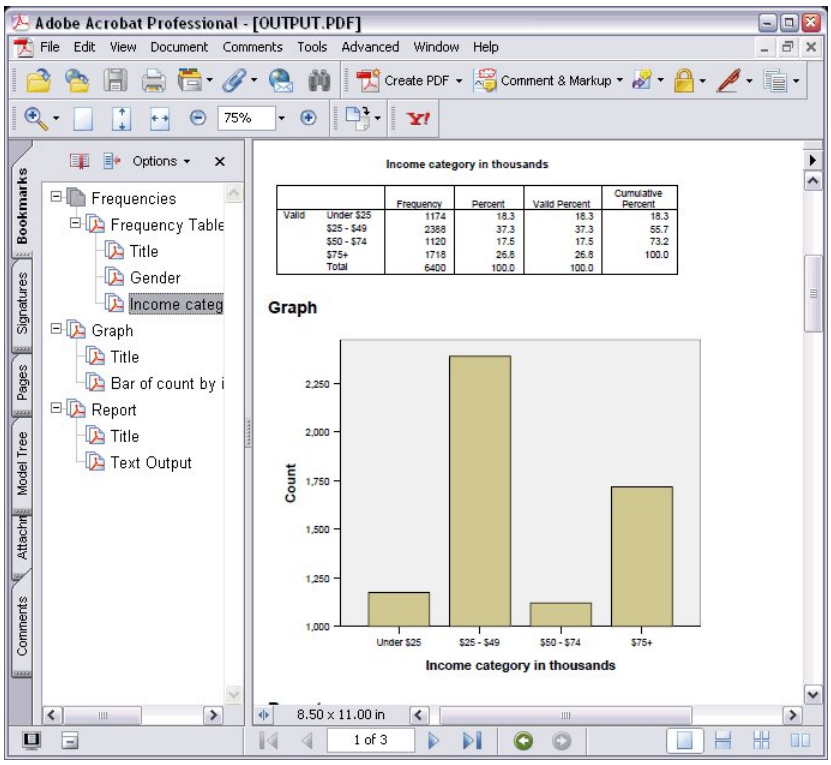


Abbildung 66. PDF-Datei mit Lesezeichen

Exportieren von Ergebnissen als HTML

Sie können Ergebnisse auch als HTML (Hypertext Markup Language) exportieren. Beim Speichern als HTML werden alle nicht grafischen Ausgaben in eine einzelne HTML-Datei exportiert.

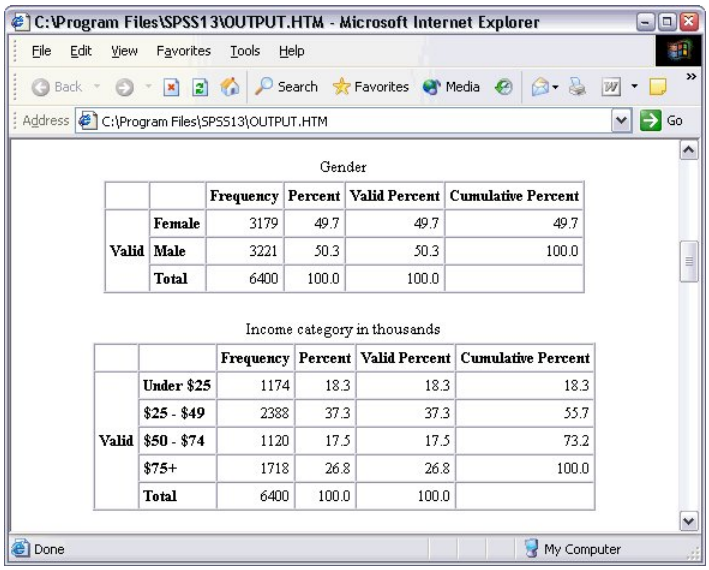


Abbildung 67. Output.htm im Web-Browser

Beim Exportieren als HTML können auch Diagramme exportiert werden, diese werden jedoch in eine separate Datei exportiert.

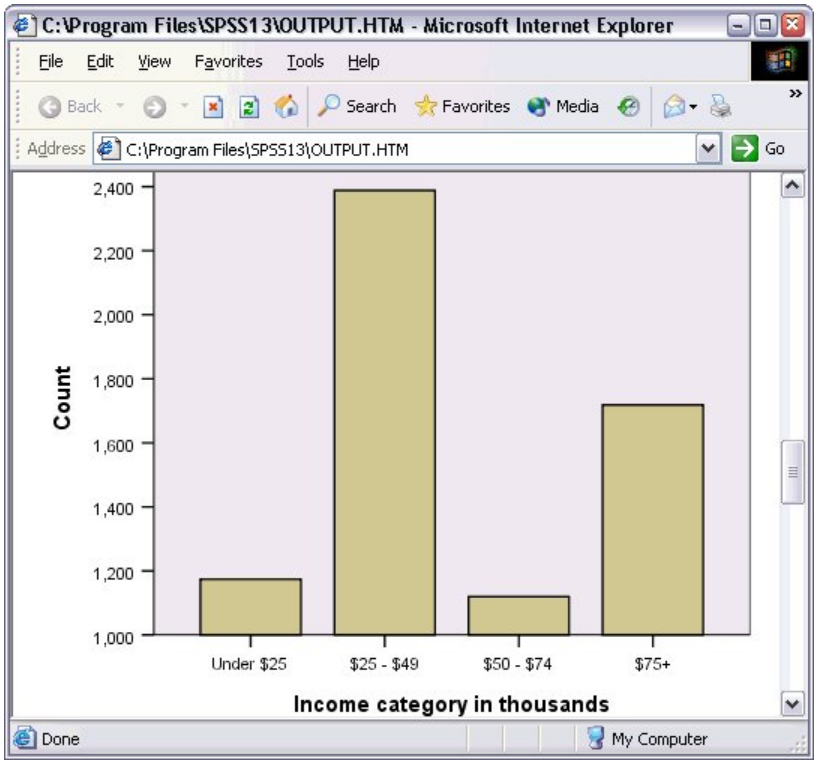


Abbildung 68. Diagramm in HTML

Jedes Diagramm wird als separate Datei in einem Format gespeichert, das Sie angeben. In die HTML-Datei werden Verweise auf diese Grafikdateien eingefügt. Darüber hinaus ist eine Option vorhanden, mit der alle (oder ausgewählte) Diagramme in separate Grafikdateien exportiert werden können.

Kapitel 7. Arbeiten mit Syntax

Sie können viele häufig verwendete Tasks mithilfe der leistungsstarken Befehlssprache speichern und automatisieren. Sie bietet außerdem einige Funktionen, die nicht über die Menüs und Dialogfelder zur Verfügung stehen. Die meisten Befehle können über die Menüs und Dialogfelder aufgerufen werden. Einige Befehle und Optionen stehen jedoch ausschließlich über die Befehlssprache zur Verfügung. Mit der Befehlssprache verfügen Sie außerdem über die Möglichkeit, Jobs in einer Syntaxdatei zu speichern. Sie können eine Analyse dann zu einem späteren Zeitpunkt wiederholen.

Eine Befehlssyntaxdatei ist einfach eine Textdatei, die IBM SPSS Statistik -Syntaxbefehle enthält. Sie können ein Syntaxfenster öffnen und direkt Befehle eingeben. Oftmals ist es aber einfacher, wenn Sie sich der Dialogfelder bedienen.

Für die Beispiele dieses Kapitels wird die Datendatei *demo.sav* verwendet. Weitere Informationen finden Sie im Thema Kapitel 10, „Stichprobendateien", auf Seite 83.

Hinweis: Die Befehlssyntax ist in der Studentenversion nicht verfügbar.

Übernehmen von Befehlssyntax

Am einfachsten kann Syntax über die Schaltfläche "Einfügen" erstellt werden, die in den meisten Dialogfeldern vorhanden ist.

1. Öffnen Sie die Datendatei *demo.sav*. Weitere Informationen finden Sie im Thema Kapitel 10, „Stichprobendateien", auf Seite 83.
2. Wählen Sie in den Menüs Folgendes aus:

   Analysieren > **Deskriptive Statistiken** > **Häufigkeiten...**
3. Wählen Sie den Eintrag *Verheiratet [heirat]* aus und verschieben Sie ihn in die Liste "Variable(n)".
4. Klicken Sie auf **Diagramme**.
5. Wählen Sie im Dialogfeld "Grafiken" die Option **Balkendiagramme** aus.
6. Wählen Sie im Gruppenfeld "Diagrammwerte" die Option **Prozente** aus.
7. Klicken Sie auf **Weiter**. Klicken Sie auf **Einfügen**, um die Syntax, die anhand der Angaben in den Dialogfeldern erstellt wurde, in den Syntaxeditor zu kopieren.

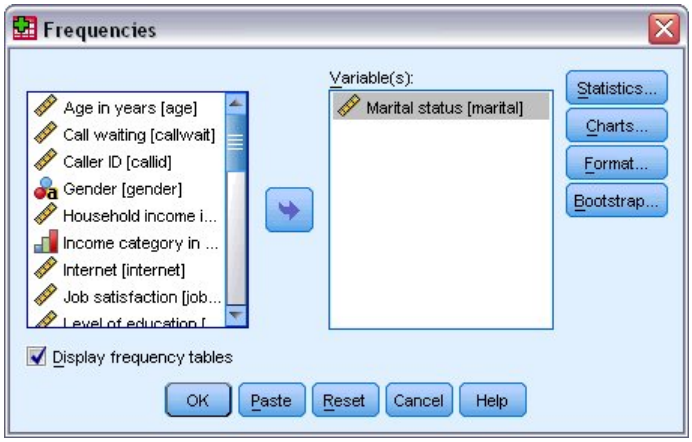


Abbildung 69. Dialogfeld "Häufigkeiten"

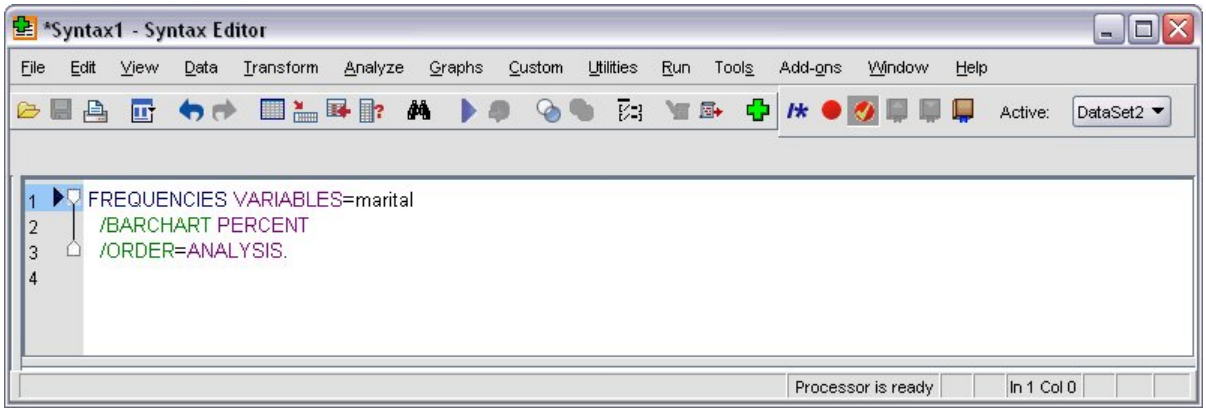


Abbildung 70. Häufigkeitssyntax

8. Wenn Sie die gerade angezeigte Syntax ausführen möchten, wählen Sie die folgenden Menübefehle aus:

   Ausführen > **Auswahl**

Bearbeiten von Befehlssyntax

Die Befehlssyntax kann im Syntaxfenster bearbeitet werden. Sie können den Unterbefehl /BARCHART beispielsweise so bearbeiten, dass Häufigkeiten anstelle von Prozentwerten angezeigt werden. (Ein Unterbefehl wird durch einen Schrägstrich gekennzeichnet.) Wenn Sie das Schlüsselwort zur Anzeige der Häufigkeiten kennen, können Sie es direkt eingeben. Wenn Sie es nicht kennen, können Sie eine Liste der verfügbaren Schlüsselwörter für den Unterbefehl abrufen, indem Sie den Cursor hinter den Namen des Unterbefehls setzen und Strg+Leertaste drücken. Damit wird die Steuerung zur automatischen Vervollständigung für den Unterbefehl angezeigt.

Löschen Sie das Schlüsselwort PERCENT aus dem Unterbefehl BARCHART.

Drücken Sie Strg+Leertaste.

1. Klicken Sie auf das Element mit der Beschriftung **FREQ** für Häufigkeiten (Frequencies). Wenn Sie in der Steuerung zur automatischen Vervollständigung auf ein Element klicken, wird dieses an der aktuellen Cursorposition eingefügt.

   Standardmäßig bietet Ihnen die Steuerung zur automatischen Vervollständigung eine Liste verfügbarer Terme, während Sie schreiben. Beispiel: Sie möchten zusätzlich zum Balkendiagramm auch ein Kreisdiagramm anzeigen. Das Kreisdiagramm wird mit einem separaten Unterbefehl angegeben.
2. Drücken Sie nach dem Schlüsselwort **FREQ** die Eingabetaste und geben Sie einen Schrägstrich ein, um den Beginn eines Unterbefehls zu markieren.

Der Syntaxeditor stellt Ihnen eine Liste von Unterbefehlen für den aktuellen Befehl zur Auswahl.

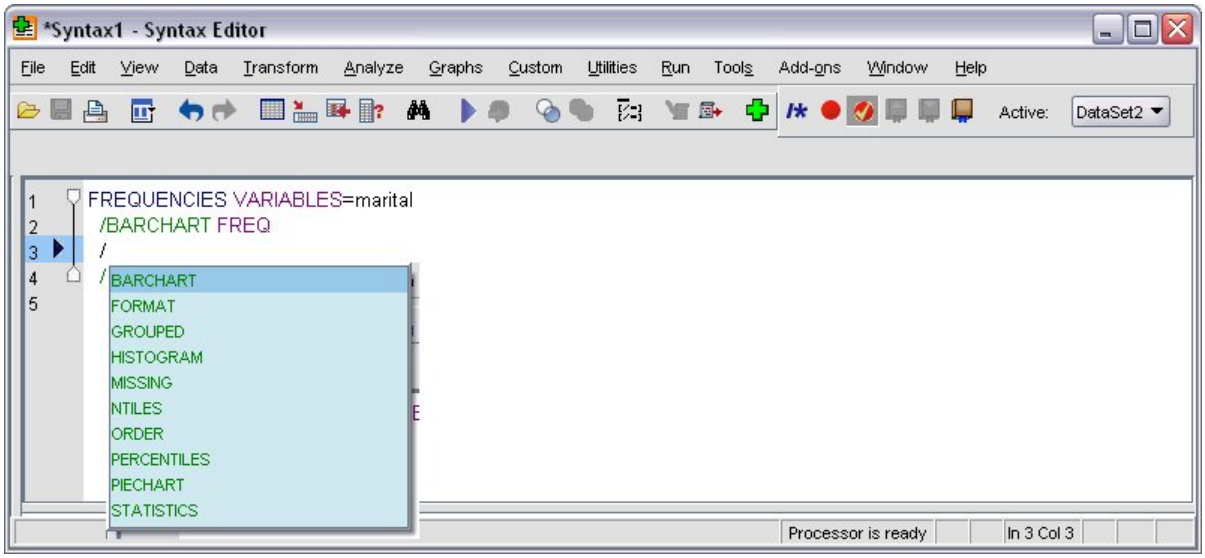


Abbildung 71. Steuerung zur automatischen Vervollständigung mit Unterbefehlen

Drücken Sie für ausführlichere Hilfe zum aktuellen Befehl die Taste F1. Damit gelangen Sie direkt zur Befehlssyntaxreferenz für den aktuellen Befehl.

Sie haben eventuell bemerkt, dass Text im Syntaxfenster farbig angezeigt wird. Mit der Farbcodierung identifizieren Sie nicht erkannte Terme schnell und einfach, da nur erkannte Terme farbig sind. Beispiel: Ihnen unterläuft ein Tippfehler und Sie schreiben den Unterbefehl FORMAT als FRMAT. Unterbefehle sind standardmäßig grün, aber der Text FRMAT erscheint ohne Farbe, da er nicht erkannt wird.

Öffnen und Ausführen einer Syntaxdatei

1. Zum Öffnen einer gespeicherten Syntaxdatei wählen Sie in den Menüs Folgendes aus:

   Datei > **Öffnen** > **Syntax...**

   Ein Standarddialogfeld zum Öffnen von Dateien wird angezeigt.
2. Wählen Sie eine Syntaxdatei aus. Wenn keine Syntaxdateien angezeigt werden, müssen Sie sicherstellen, dass **Syntax (*.sps)** als anzuzeigender Dateityp ausgewählt ist.
3. Klicken Sie auf **Öffnen**.
4. Verwenden Sie das Menü "Ausführen" im Syntaxeditor, um die Befehle auszuführen.

Wenn die Befehle auf eine bestimmte Datendatei angewendet werden sollen, muss die betreffende Datendatei vor der Ausführung der Befehle geöffnet werden oder Sie müssen ein Befehl zum Öffnen der Datendatei in die Syntax einbeziehen. Diesen Befehlstyp können Sie aus den Dialogfeldern übernehmen, mit denen Datendateien geöffnet werden.

Verwenden von Haltepunkten

Mit Haltepunkten können Sie die Ausführung der Befehlssyntax an angegebenen Punkten in der Syntax unterbrechen und die Ausführung fortsetzen, wenn Sie bereit sind. So können Sie die Ausgabe oder Daten an einem Zwischenpunkt in einem Syntaxjob sehen oder Befehlssyntax ausführen, die Informationen über den aktuellen Status von Daten anzeigt, z. B. FREQUENCIES. Haltepunkte können nur auf Befehlsebene gesetzt werden, nicht in bestimmten Zeilen innerhalb eines Befehls.

So fügen Sie einen Haltepunkt in einen Befehl ein:

1. Klicken Sie in einen beliebigen Bereich links neben dem Text, der mit dem Befehl verbunden ist.

   Der Haltepunkt wird unabhängig von der genauen Stelle, an die Sie geklickt haben, im Bereich links neben dem Befehlstext und in der Zeile des Befehlsnamens als roter Kreis dargestellt.

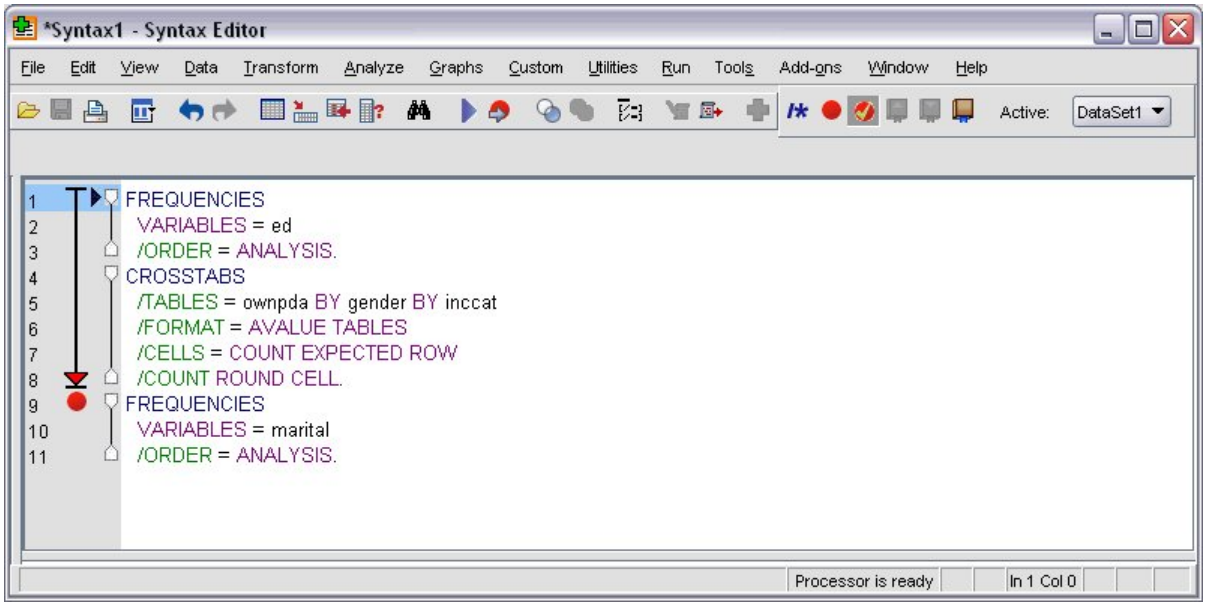


Abbildung 72. Ausführung am Haltepunkt gestoppt

Wenn Sie eine Befehlssyntax mit Haltepunkten ausführen, wird die Ausführung vor jedem Befehl unterbrochen, der einen Haltepunkt enthält.

Der Abwärtspfeil links neben dem Befehlstext zeigt den Fortschritt der Syntaxausführung. Er umfasst den Bereich vom ersten ausgeführten Befehl bis zum zuletzt ausgeführten Befehl und ist besonders nützlich bei der Ausführung von Befehlssyntax mit Haltepunkten.

So setzen Sie die Ausführung nach einem Haltepunkt fort:

2. Wählen Sie in den Menüs des Syntaxeditors Folgendes aus:

   Ausführen > **Weiter**

Kapitel 8. Ändern von Datenwerten

Daten sind nicht immer von Anfang an optimal für Analysen oder Berichte strukturiert. Angenommen, Sie möchten folgende Vorgänge durchführen:

- Erstellen einer kategorialen Variablen aus einer metrischen Variablen.
- Kombinieren mehrerer Ergebniskategorien in einer einzelnen Kategorie.
- Erstellen einer neuen Variablen, welche die berechnete Differenz zwischen zwei vorhandenen Variablen darstellt.
- Berechnen des Zeitabstands zwischen zwei Datumsangabe.

In diesem Kapitel wird die Datendatei *demo.sav* verwendet. Weitere Informationen finden Sie im Thema Kapitel 10, „Stichprobendateien”, auf Seite 83.

Erstellen einer kategorialen Variablen aus einer metrischen Variablen

Mehrere kategoriale Variablen in der Datendatei *demo.sav* sind eigentlich von metrischen Variablen in dieser Datendatei abgeleitet. Die Variable *eink_kl* ergibt sich einfach dadurch, dass *einkomm* in vier Kategorien unterteilt wird. Bei dieser kategorialen Variablen stehen die ganzzahligen Werte 1 – 4 für folgende Einkommensklassen (in Tausend): "unter $25", "$25 - $49", "$50 – $74" und "$75 oder höher".

So erstellen Sie die kategoriale Variable *eink_kl*:

1. Wählen Sie in den Menüs im Fenster "Dateneditor" Folgendes aus:

   Transformation > **Visuelle Klassierung...**

   Im ersten Dialogfeld von "Visuelle Klassierung" wählen Sie die metrischen und/oder ordinalen Variablen aus, für die neue, klassierte Variablen erstellt werden sollen. **Klassierung** bedeutet, dass zwei oder mehrere nebeneinander liegende Werte zusammengefasst und in dieselbe Kategorie eingeordnet werden.

   Da die Funktion "Visuelle Klassierung" die tatsächlichen Werte in der Datendatei verwendet, um Ihnen eine sinnvolle Klassierung zu erleichtern, muss sie die Datendatei zuerst lesen. Da dies einige Zeit in Anspruch nehmen kann, wenn Ihre Datendatei eine große Anzahl von Fällen enthält, können Sie in diesem Anfangsdialogfeld auch die Anzahl der zu lesenden (durchsuchenden) Fälle begrenzen. Bei der verwendeten Beispieldatendatei ist dies nicht erforderlich. Obwohl sie mehr als 6.000 Fälle umfasst, dauert das Durchsuchen bei dieser Anzahl von Fällen nicht besonders lang.
2. Verschieben Sie den Eintrag *Haushaltseinkommen in Tausend (einkomm)* durch Ziehen und Ablegen aus der Liste "Variablen" in die Liste "Variablen für Klassierung" und klicken Sie anschließend auf **Weiter**.

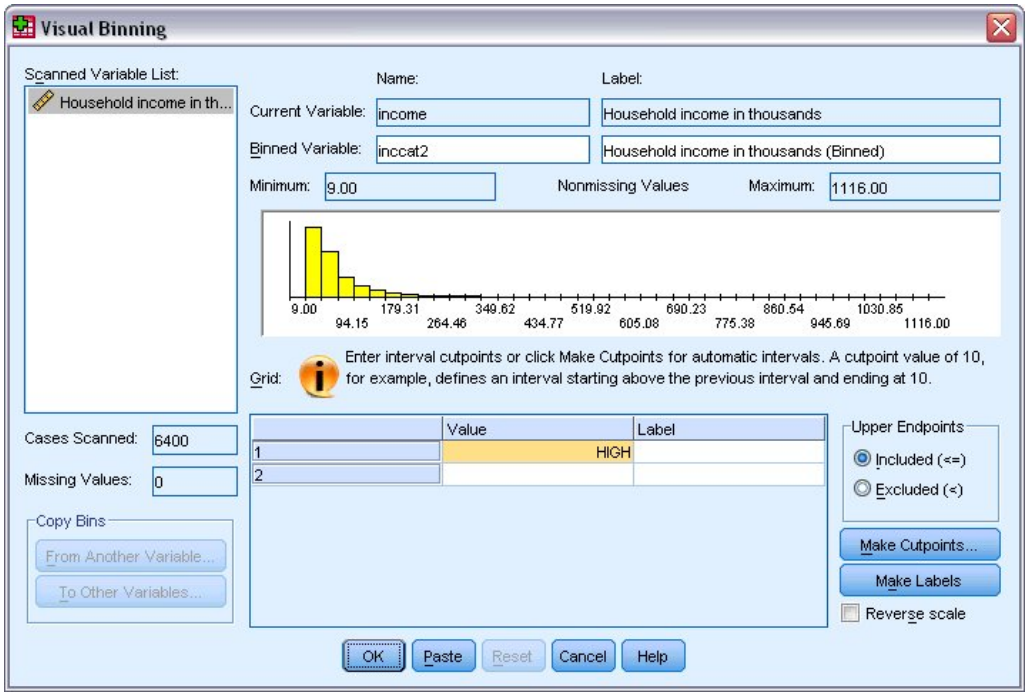


Abbildung 73. Visuelle Klassierung – Hauptdialogfeld

3. Wählen Sie im Hauptdialogfeld von "Visuelle Klassierung" in der Liste der durchsuchten Variablen den Eintrag *Haushaltseinkommen in Tausend [einkomm]* aus.

   Ein Histogramm zeigt die Verteilung der ausgewählten Variablen an, die in diesem Fall sehr schief ist.

4. Geben Sie `inccat2` für den neuen Namen der gebinnten Variablen und `Income category [in thousands]` für die Variablenbezeichnung ein.
5. Klicken Sie auf **Trennwerte erstellen**.
6. Wählen Sie **Intervalle mit gleicher Breite** aus.
7. Geben Sie 25 als ersten Trennwert, 3 als Anzahl der Trennwerte und 25 als Breite ein.

   Die Anzahl der klassierten Kategorien ist um den Wert 1 größer als die Anzahl der Trennwerte. In diesem Beispiel weist die neue, klassierte Variable also vier Kategorien auf, wobei die ersten drei Kategorien Bereiche von 25 (tausend) umfassen und die letzte Kategorie alle Werte über dem höchsten Trennwert 75 (tausend) enthält.

8. Klicken Sie auf **Anwenden**.

   Die nun im Raster angezeigten Werte stellen die festgelegten Trennwerte dar, die oberen Endpunkte der einzelnen Kategorien. Die Position der Trennwerte wird auch durch vertikale Linien im Histogramm angezeigt.

   Standardmäßig sind diese Trennwerte in den entsprechenden Kategorien enthalten. Der erste Wert (25) bedeutet beispielsweise, dass alle Werte kleiner oder gleich 25 eingeschlossen werden. In diesem Beispiel sollen jedoch die Klassen "unter 25", "25-49", "50-74" und "75 oder höher" verwenden werden.

9. Wählen Sie in der Gruppe "Obere Endpunkte" die Option **Ausgeschlossen (<)** aus.
10. Klicken Sie anschließend auf **Beschriftungen erstellen**.

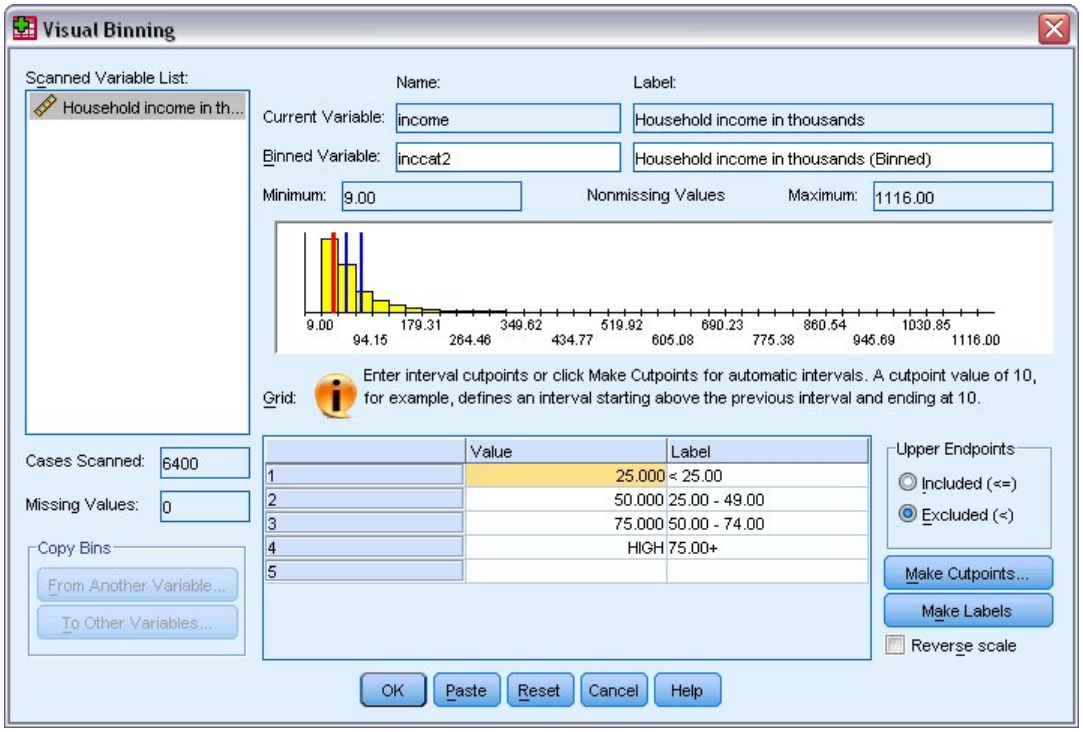


Abbildung 74. Automatisch erstellte Wertbeschriftungen

Dadurch werden automatisch beschreibende Wertbeschriftungen für die einzelnen Kategorien erstellt. Da es sich bei den Werten, die der neuen klassierten Variablen zugewiesen werden, einfach um aufeinander folgende Ganzzahlen, beginnend bei 1, handelt, können die Wertbeschriftungen sehr hilfreich sein.

Außerdem können Sie manuell Trennwerte und Beschriftungen im Raster eingeben oder ändern, die Position von Trennwerten durch Ziehen und Ablegen der Trennwertlinien im Histogramm ändern und Trennwerte löschen, indem Sie die Trennwertlinien vom Histogramm weg ziehen.

11. Klicken Sie auf **OK**, um die neue, klassierte Variable zu erstellen.

Die neue Variable wird im Dateneditor angezeigt. Da die Variable am Ende der Liste hinzugefügt wird, wird sie in der Ansicht "Daten" in der äußersten rechten Spalte und in der Ansicht "Variable" in der letzten Zeile angezeigt.

Berechnen von neuen Variablen

Sie können eine Vielfalt von mathematischen Funktionen einsetzen, um neue Variablen auf der Grundlage von komplizierten Gleichungen zu berechnen. In diesem Beispiel wird jedoch lediglich eine neue Variable berechnet, die die Differenz zwischen den Werten von zwei vorhandenen Variablen darstellt.

Die Datendatei *demo.sav* enthält eine Variable für das aktuelle Alter der Befragten und eine Variable für die Anzahl der Jahre beim aktuellen Arbeitgeber. Sie enthält dagegen keine Variable dafür, welches Alter die Befragten hatten, als sie diese Arbeitsstelle antraten. Sie können eine neue Variable erstellen, welche die berechnete Differenz zwischen dem aktuellen Alter und der Anzahl der Jahre beim aktuellen Arbeitgeber darstellt. Diese Differenz ist ungefähr das Alter, in dem die Befragten diese Arbeitsstelle antraten.

1. Wählen Sie in den Menüs im Fenster "Dateneditor" Folgendes aus:

   Transformation > **Variable berechnen...**

2. Geben Sie als Zielvariable `jobstart` ein.
3. Wählen Sie in der Quellenvariablenliste *Alter in Jahren [Alter]* aus und klicken Sie auf die Pfeilschaltfläche, um die Variable in das Textfeld "Numerischer Ausdruck" zu kopieren.

4. Klicken Sie im Rechnerbereich des Dialogfelds auf die Schaltfläche mit dem Minuszeichen (–) oder drücken Sie die Minustaste auf der Tastatur.
5. Wählen Sie *Jahre beim aktuellen Arbeitgeber [arbeit]* aus und klicken Sie auf die Pfeilschaltfläche, um die Variable in die Liste für den Ausdruck zu kopieren.

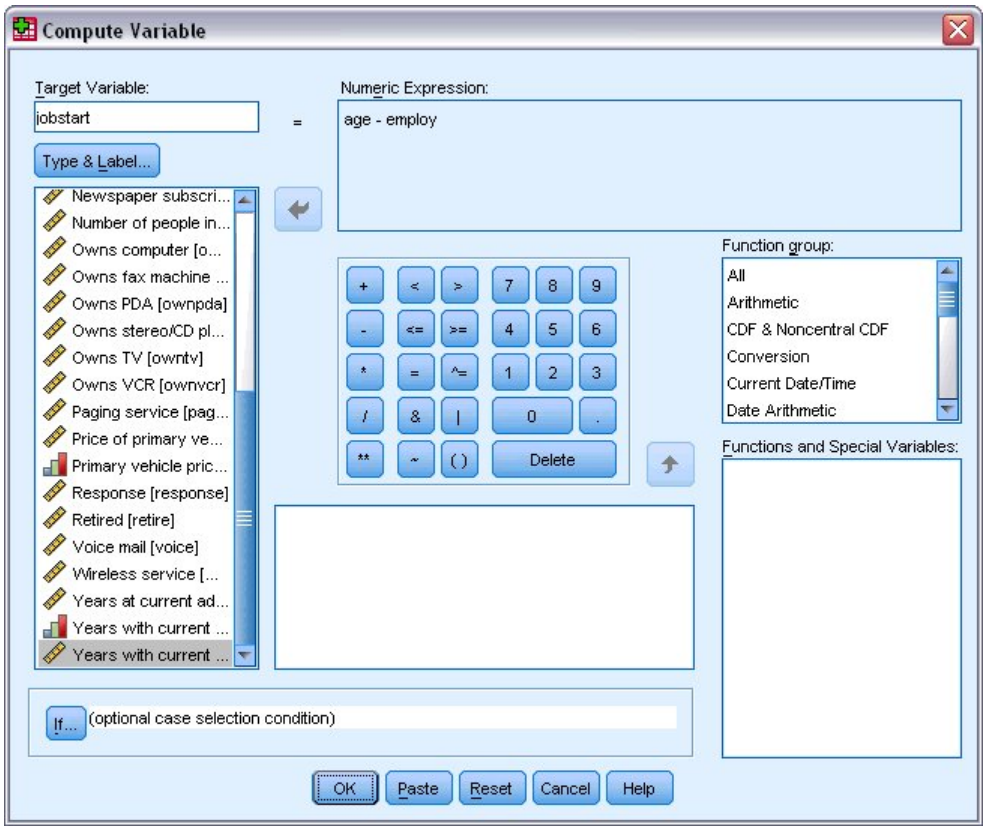


Abbildung 75. Dialogfeld "Variable berechnen"

Hinweis: Achten Sie darauf, die richtige Variable für den Arbeitsplatz auszuwählen. Es gibt auch eine umcodierte kategoriale Version der Variablen. Dies ist *nicht* die Variable, die Sie benötigen. Der numerische Ausdruck sollte *alter–arbeit* sein, nicht *alter–arbeit_kl*.

6. Klicken Sie auf **OK**, um die neue Variable zu berechnen.

Die neue Variable wird im Dateneditor angezeigt. Da die Variable am Ende der Liste hinzugefügt wird, wird sie in der Ansicht "Daten" in der äußersten rechten Spalte und in der Ansicht "Variable" in der letzten Zeile angezeigt.

Verwenden von Funktionen in Ausdrücken

Sie können außerdem vordefinierte Funktionen in Ausdrücken verwenden. Es stehen mehr als 70 integrierte Funktionen zur Verfügung, darunter folgende:

- Arithmetische Funktionen
- Statistische Funktionen
- Verteilungsfunktionen
- Logische Funktionen
- Funktionen zur Aggregation und Extraktion von Datum und Uhrzeit
- Funktionen für fehlende Werte
- Fallübergreifende Funktionen
- Zeichenfolgefunktionen

Funktionen werden in logisch getrennte Gruppen eingeteilt, beispielsweise eine Gruppe für arithmetische Operationen und eine Gruppe zur Berechnung statistischer Metriken. Zur Erhöhung der Benutzerfreundlichkeit ist auch eine Anzahl von häufig verwendeten Systemvariablen, beispielsweise *$TIME* (aktuelles Datum und aktuelle Uhrzeit) in den entsprechenden Funktionsgruppen enthalten.

Einfügen einer Funktion in einen Ausdruck

So fügen Sie eine Funktion in einen Ausdruck ein:

1. Setzen Sie den Mauszeiger an die Stelle im Ausdruck, an der die Funktion eingefügt werden soll.
2. Wählen Sie aus der Liste "Funktionsgruppe" die geeignete Gruppe aus. Die Gruppe mit der Beschriftung **Alle** bietet eine Auflistung aller verfügbaren Funktionen und Systemvariablen.
3. Doppelklicken Sie in der Liste "Funktionen und Sondervariablen" auf die Funktion. (Sie können auch die Funktion auswählen und auf den Pfeil neben der Liste "Funktionsgruppe" klicken.)

Die Funktion wird in den Ausdruck eingefügt. Wenn Sie einen Teil des Ausdrucks markieren und anschließend die Funktion einfügen, wird der hervorgehobene Teil des Ausdrucks als erstes Argument der Funktion verwendet.

Bearbeiten einer Funktion in einem Ausdruck

Die Funktion ist erst vollständig, nachdem Sie die Argumente eingegeben haben, die in der eingefügten Funktion durch Fragezeichen dargestellt werden. Die Anzahl der Fragezeichen gibt die Mindestanzahl der Argumente an, die zur Vervollständigung der Funktion erforderlich sind.

1. Markieren Sie das bzw. die Fragezeichen in der eingefügten Funktion.
2. Geben Sie die Argumente ein. Wenn es sich bei den Argumenten um Variablennamen handelt, können Sie sie aus der Variablenliste einfügen.

Verwenden von bedingten Ausdrücken

Mit bedingten Ausdrücken, auch logische Ausdrücke genannt, können Sie Transformationen auf ausgewählte Subsets von Fällen anwenden. Ein bedingter Ausdruck gibt für jeden Fall den Wert "Wahr", "Falsch" oder "Fehlend" zurück. Wenn das Ergebnis eines bedingten Ausdrucks "Wahr" lautet, wird die Transformation für den Fall durchgeführt. Wenn als Ergebnis der Wert "Falsch" oder "Fehlend" vorliegt, wird die Transformation nicht auf den Fall angewendet.

So geben Sie einen bedingten Ausdruck ein:

1. Klicken Sie im Dialogfeld "Variable berechnen" auf **Falls**. Dadurch wird das Dialogfeld "Falls Bedingung erfüllt ist" aufgerufen.
2. Wählen Sie **Fall einschließen, wenn Bedingung erfüllt ist** aus.
3. Geben Sie den bedingten Ausdruck ein.

Die meisten bedingten Ausdrücke enthalten mindestens einen relationalen Operator. Beispiel:

```
age>=21
```

oder

```
income*3<100
```

Im ersten Beispiel werden nur Fälle ausgewählt, in denen der Wert für *Alter [alter]* größer-gleich 21 ist. Im zweiten Beispiel muss der Wert für *Haushaltseinkommen in Tausend [einkomm]* multipliziert mit 3 unter 100 liegen, damit ein Fall ausgewählt wird.

Es besteht außerdem die Möglichkeit, mehr als zwei bedingte Ausdrücke über logische Operatoren zu verbinden. Beispiel:

```
age>=21 | ed>=4
```

oder

```
income*3<100 & ed=5
```

Im ersten Beispiel werden Fälle ausgewählt, die die Bedingung für *Alter [alter]* oder für *Schulabschluss [schulab]* erfüllen. Im zweiten Beispiel müssen sowohl die Bedingungen für *Haushaltseinkommen in Tausend [einkomm]* als auch die Bedingungen für *Schulabschluss [schulab]* erfüllt sein, damit ein Fall ausgewählt wird.

Arbeiten mit Datumsangaben und Uhrzeiten

Mit dem Assistenten für Datum und Uhrzeit lässt sich eine Reihe von Aufgaben, die häufig mit Datumsangaben und Uhrzeiten durchgeführt werden, problemlos bewältigen. Mit diesem Assistenten können Sie folgende Aufgaben ausführen:

- Erstellen einer Datums-/Zeitvariablen aus einer Zeichenfolgevariablen, die ein Datum oder eine Uhrzeit enthält.
- Erstellen einer Datums-/Zeitvariablen durch Zusammenführen von Variablen, die verschiedene Teile des Datums bzw. der Uhrzeit enthalten.
- Addieren oder Subtrahieren von Werten zu bzw. von einer Datums-/Zeitvariablen (einschließlich Addition bzw. Subtraktion von zwei Datums-/Zeitvariablen).
- Extrahieren eines Teils einer Datums- oder Zeitvariablen, beispielsweise des Tags im Monat aus einer Datums-/Zeitvariablen mit dem Format mm/tt/jjjj.

Für die Beispiele in diesem Abschnitt wird die Datendatei *upgrade.sav* verwendet. Weitere Informationen finden Sie im Thema Kapitel 10, „Stichprobendateien”, auf Seite 83.

So verwenden Sie den Assistenten für Datum und Uhrzeit:

1. Wählen Sie in den Menüs Folgendes aus:

Transformation > **Assistent für Datum und Uhrzeit...**

Abbildung 76. Assistent für Datum und Uhrzeit: Einführungsbildschirm

Im Einführungsbildschirm des Assistenten für Datum und Uhrzeit wird eine Reihe allgemeiner Aufgaben angezeigt. Aufgaben, die nicht auf die aktuellen Daten zutreffen, sind inaktiviert. So enthält die Datendatei *upgrade.sav* beispielsweise keine Zeichenfolgevariablen, sodass die Aufgabe zur Erstellung einer Datenvariablen aus einer Zeichenfolge inaktiviert ist.

Wenn Sie sich mit Datum und Uhrzeit in IBM SPSS Statistik nicht auskennen, können Sie die Option **Erfahren, wie Datum und Uhrzeit dargestellt werden** aktivieren und auf **Weiter** klicken. Dadurch gelangen Sie

auf einen Bildschirm, der einen kurzen Überblick über Datums-/Zeitvariablen und (über die Schaltfläche "Hilfe") einen Link zu detaillierteren Informationen bietet.

Berechnen des Zeitabstands zwischen zwei Datumsangaben

Zu den häufigsten Aufgaben, die mit Datumsangaben zu tun haben, gehört die Berechnung des Zeitabstands zwischen zwei Datumsangaben. Betrachten wir folgendes Beispiel: Ein Softwareunternehmen möchte die Verkäufe von Upgradelizenzen analysieren, indem die Anzahl der Jahre ermittelt wird, die vergangen sind, seitdem die einzelnen Kunden zuletzt ein Upgrade erworben haben. Die Datendatei *upgrade.sav* enthält eine Variable für das Datum, an dem die einzelnen Kunden zuletzt ein Upgrade erwarben, und nicht die Anzahl der Jahre seit diesem Kauf. Eine neue Variable, die die Zeitspanne in Jahren zwischen dem Datum der letzten Aktualisierung und dem Datum der nächsten Produktveröffentlichung angibt, bietet ein Maß für diesen Wert.

So berechnen Sie den Zeitabstand zwischen zwei Datumsangaben:

1. Wählen Sie im Einführungsbildschirm des Assistenten für Datum und Uhrzeit die Option **Berechnungen mit Datums- und Zeitwerten durchführen** aus und klicken Sie anschließend auf **Weiter**.
2. Wählen Sie die Option **Berechnen der Anzahl der Zeiteinheiten zwischen zwei Datumswerten** und klicken Sie auf **Weiter**.

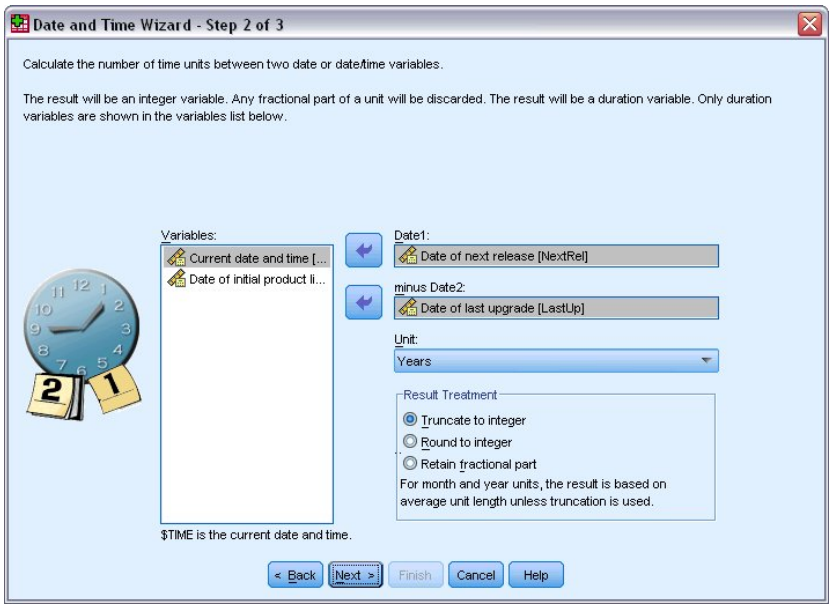


Abbildung 77. Berechnen des Zeitabstands zwischen zwei Datumsangaben: Schritt 2

3. Wählen Sie in Schritt 2 *Date of next release* (Datum der nächsten Veröffentlichung) als "Datum1" aus.
4. Wählen Sie *Date of last upgrade* (Datum des letzten Upgrades) als "Datum2" aus.
5. Wählen Sie **Jahre** als Einheit und **Auf ganze Zahl kürzen** als Ergebnisbehandlung aus. (Dies sind die Standardeinstellungen.)
6. Klicken Sie auf **Weiter**.
7. Geben Sie in Schritt 3 *YearsLastUp* als Name der Ergebnisvariablen ein. Ergebnisvariablen können nicht denselben Namen haben wie bestehende Variablen.
8. Geben Sie *Years since last upgrade* (Jahre seit dem letzten Upgrade) als Beschriftung für die Ergebnisvariable ein. Die Variablenbeschriftungen für Ergebnisvariablen sind optional.
9. Behalten Sie die Standardauswahl **Variable jetzt erstellen** bei und klicken Sie auf **Fertigstellen**, um die neue Variable zu erstellen.

Bei der neuen Variablen *YearsLastUp*, die im Dateneditor angezeigt wird, handelt es sich um die ganzzahlige Angabe der Jahre zwischen den beiden Datumswerten. Bruchteile von Jahren wurden gekürzt.

Hinzufügen einer Dauer zu einem Datum

Sie können Werte für die Zeitdauer zu einem Datum addieren bzw. subtrahieren, beispielsweise 10 Tage oder 12 Monate. Setzen wir das Beispiel des Softwareunternehmens aus dem vorangegangen Abschnitt fort: Wir bestimmen das Datum, an dem der ursprüngliche Vertrag über den Technischen Support für die einzelnen Kunden ausläuft. Die Datendatei *upgrade.sav* enthält eine Variable für die Anzahl der Jahre für den vertraglich zugesicherten Support und eine Variable für das ursprüngliche Kaufdatum. Nun kann das Enddatum des ursprünglichen Supports bestimmt werden, indem die Jahre des Supportzeitraums zum Kaufdatum hinzugefügt werden.

So fügen Sie eine Dauer zu einem Datum hinzu:

1. Wählen Sie im Einführungsbildschirm des Assistenten für Datum und Uhrzeit die Option **Berechnungen mit Datums- und Zeitwerten durchführen** aus und klicken Sie anschließend auf **Weiter**.
2. Wählen Sie in Schritt 1 die Option **Addieren bzw. Subtrahieren einer Dauer zu bzw. von einem Datum** und klicken Sie auf **Weiter**.

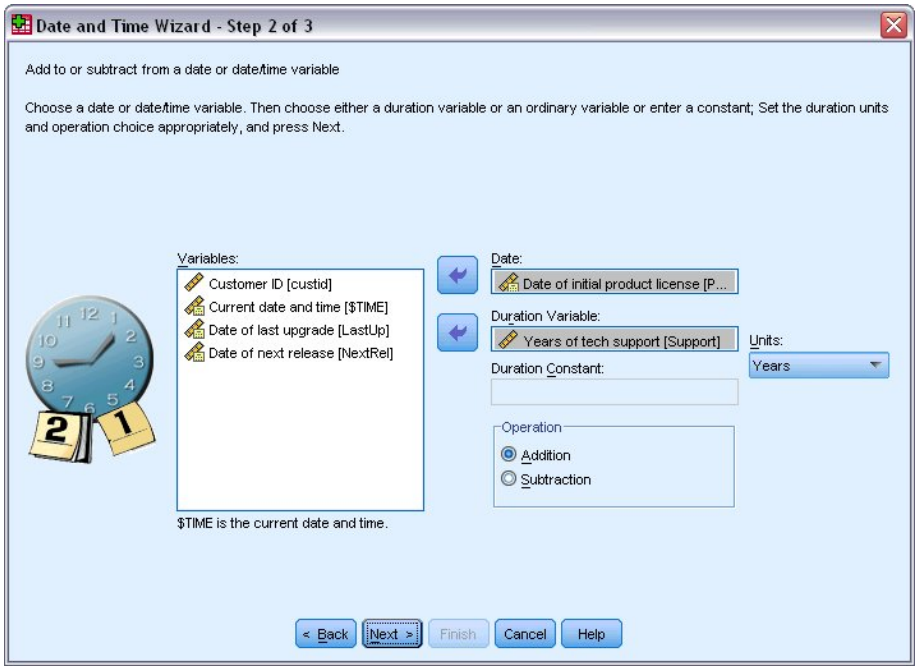


Abbildung 78. Hinzufügen einer Dauer zu einem Datum: Schritt 2

3. Wählen Sie als Datum die Option *Date of initial product license* (Datum der ersten Produktlizenz) aus.
4. Wählen Sie in Schritt 2 *Years of tech support* (Jahre an technischem Support) als Variable für die Dauer aus.

   Da es sich bei *Years of tech support* (Jahre an technischem Support) einfach um eine numerische Variable handelt, müssen Sie angeben, welche Einheiten verwendet werden sollen, wenn diese Variable als Dauer hinzugefügt wird.
5. Wählen Sie in der Dropdown-Liste für die Einheiten die Option **Jahre** aus.
6. Klicken Sie auf **Weiter**.
7. Geben Sie in Schritt 3 *SupEndDate* als Name der Ergebnisvariablen ein. Ergebnisvariablen können nicht denselben Namen haben wie bestehende Variablen.
8. Geben Sie *End date for support* (Enddatum für den Support) als Beschriftung für die Ergebnisvariable ein. Die Variablenbeschriftungen für Ergebnisvariablen sind optional.
9. Klicken Sie auf **Fertigstellen**, um die neue Variable zu erstellen.

Die neue Variable wird im Dateneditor angezeigt.

Kapitel 9. Sortieren und Auswählen von Daten

Datendateien liegen nicht immer genau in der Form vor, die Sie gerade benötigen. Um Daten für die Analyse vorzubereiten, stehen Ihnen mehrere Möglichkeiten zur Dateitransformation zur Verfügung, darunter folgende Funktionen:

- **Sortieren von Daten.** Sie können Fälle nach dem Wert einer oder mehrerer Variablen sortieren lassen.
- **Auswählen von Subsets von Fällen.** Sie können die Analyse auf ein Subset von Fällen beschränken oder Analysen für verschiedene Subsets gleichzeitig vornehmen.

Für die Beispiele dieses Kapitels wird die Datendatei *demo.sav* verwendet. Weitere Informationen finden Sie im Thema Kapitel 10, „Stichprobendateien", auf Seite 83.

Sortieren von Daten

Das Sortieren von Fällen, d. h. der Zeilen der Datendatei, ist für bestimmte Analysetypen hilfreich und mitunter auch erforderlich.

Um die Reihenfolge der Fälle in der Datendatei auf der Grundlage des Werts mindestens einer Sortiervariablen zu ändern, gehen Sie folgendermaßen vor:

1. Wählen Sie in den Menüs Folgendes aus:

   Daten > **Sortieren von Fällen...**

   Das Dialogfeld "Fälle sortieren" wird angezeigt.

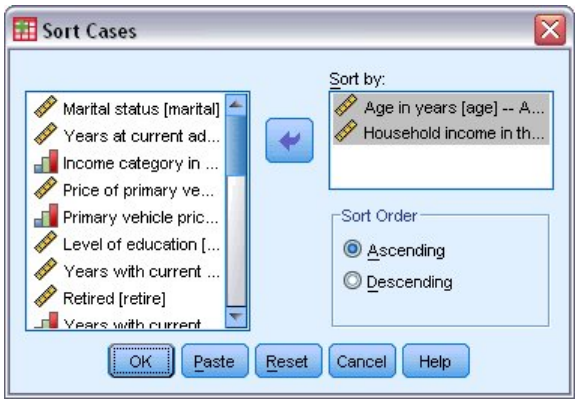


Abbildung 79. Dialogfeld "Fälle sortieren"

2. Fügen Sie der Liste "Sortieren nach" die Variablen *Alter in Jahren [alter]* und *Haushaltseinkommen in Tausend [einkomm]* hinzu.

Bei Auswahl mehrerer Sortiervariablen bestimmt die Reihenfolge, in der diese in der Liste "Sortieren nach" angezeigt werden, die Sortierreihenfolge der Fälle. In diesem Beispiel werden die Fälle auf der Grundlage der Liste "Sortieren nach" nach dem Wert von *Haushaltseinkommen in Tausend [einkomm]* innerhalb von Kategorien von *Alter in Jahren [alter]* sortiert. Bei Zeichenfolgevariablen stehen Großbuchstaben in der Sortierreihenfolge vor den entsprechenden Kleinbuchstaben. Der Zeichenfolgewert *Ja* steht in der Sortierreihenfolge beispielsweise vor *ja*.

Verarbeitung von aufgeteilten Dateien

So teilen Sie die Datendatei zu Analysezwecken in einzelne Gruppen auf:

1. Wählen Sie in den Menüs Folgendes aus:

   Daten > **Datei aufteilen ...**

Das Dialogfeld "Datei aufteilen" wird geöffnet.

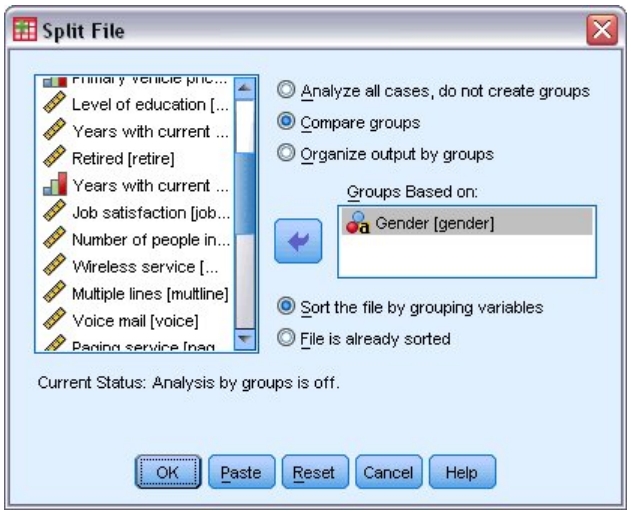


Abbildung 80. Dialogfeld "Datei aufteilen"

2. Aktivieren Sie das Optionsfeld **Gruppen vergleichen** oder **Ausgabe nach Gruppen aufteilen**. (Die nach diesen Schritten aufgeführten Beispiele zeigen die Unterschiede zwischen den beiden Optionen.)
3. Wählen Sie *Geschlecht [geschl]* aus, um die Datei für diese Variablen in getrennte Gruppen aufzuteilen.

Sie können numerische Variablen sowie kurze und lange Zeichenfolgevariablen als Gruppierungsvariablen verwenden. Für jede der durch die Gruppierungsvariablen definierten Untergruppen wird eine gesonderte Analyse durchgeführt. Bei der Auswahl mehrerer Gruppierungsvariablen bestimmt die Reihenfolge, in der diese in der Liste "Gruppen basierend auf" angezeigt werden, die Gruppierung der Fälle.

Wenn Sie **Gruppen vergleichen** auswählen, werden die Ergebnisse aller Gruppen der aufgeteilten Datei in derselben Tabelle bzw. denselben Tabellen aufgeführt, wie beispielsweise in der folgenden Tabelle mit Auswertungsstatistiken gezeigt, die von der Prozedur "Häufigkeiten" erstellt wurde.

Statistics

Household income in thousands

Female	N	Valid	3179
		Missing	0
	Mean		68.7798
	Median		44.0000
	Std. Deviation		75.73510
Male	N	Valid	3221
		Missing	0
	Mean		70.1608
	Median		45.0000
	Std. Deviation		81.56216

Abbildung 81. Ausgabe für die aufgeteilte Datei mit einer einzelnen Pivot-Tabelle

Wenn Sie **Ausgabe nach Gruppen aufteilen** auswählen und die Prozedur "Häufigkeiten" ausführen, werden zwei Pivot-Tabellen erstellt: eine Tabelle für Frauen und eine Tabelle für Männer.

Statistics[a]

Household income in thousands

N	Valid	3179
	Missing	0
Mean		68.7798
Median		44.0000
Std. Deviation		75.73510

a. Gender = Female

Abbildung 82. Ausgabe für die aufgeteilte Datei mit der Pivot-Tabelle für Frauen

Statistics[a]

Household income in thousands

N	Valid	3221
	Missing	0
Mean		70.1608
Median		45.0000
Std. Deviation		81.56216

a. Gender = Male

Abbildung 83. Ausgabe für die aufgeteilte Datei mit der Pivot-Tabelle für Männer

Sortieren von Fällen für die Verarbeitung von aufgeteilten Dateien

In der Prozedur "Datei aufteilen" wird eine neue Untergruppe erstellt, sobald ein anderer Wert für eine der Gruppierungsvariablen ermittelt wird. Es ist daher wichtig, die Fälle anhand der Werte der Gruppierungsvariablen zu sortieren, ehe Sie die aufgeteilte Datei verarbeiten.

In der Standardeinstellung wird die Datendatei in der Prozedur "Datei aufteilen" auf der Grundlage der Werte der Gruppierungsvariablen sortiert. Wenn die Datei bereits in der richtigen Reihenfolge sortiert ist, können Sie Verarbeitungszeit sparen, indem Sie **Datei ist sortiert** auswählen.

Aktivieren und Inaktivieren der Verarbeitung von aufgeteilten Dateien

Nachdem Sie die Verarbeitung von aufgeteilten Dateien aktiviert haben, bleibt sie für den Rest der Sitzung aktiv. Sie muss ausdrücklich inaktiviert werden.

- **Alle Fälle analysieren.** Diese Option inaktiviert die Verarbeitung von aufgeteilten Dateien.
- **Gruppen vergleichen** und **Ausgabe nach Gruppen aufteilen.** Diese Option aktiviert die Verarbeitung von aufgeteilten Dateien.

Wenn die Verarbeitung von aufgeteilten Dateien aktiviert ist, wird in der Statusleiste unten im Anwendungsfenster die Nachricht **Datei aufteilen an** angezeigt.

Auswählen von Subsets von Fällen

Sie können die Analyse anhand von Kriterien, zu denen Variablen und komplexe Ausdrücke gehören, auf eine bestimmte Untergruppe beschränken. Sie können auch eine Zufallsstichprobe aus den Fällen auswählen. Die Kriterien zum Festlegen der Untergruppen können folgende Elemente enthalten:

- Variablenwerte und -bereiche
- Datums- und Zeitbereiche
- Fallnummern (Zeilennummern)
- Arithmetische Ausdrücke
- Logische Ausdrücke
- Funktionen

So wählen Sie ein Subset der Fälle für die Analyse aus:

1. Wählen Sie in den Menüs Folgendes aus:

   Daten > **Fälle auswählen...**

Dadurch wird das Dialogfeld "Fälle auswählen" aufgerufen.

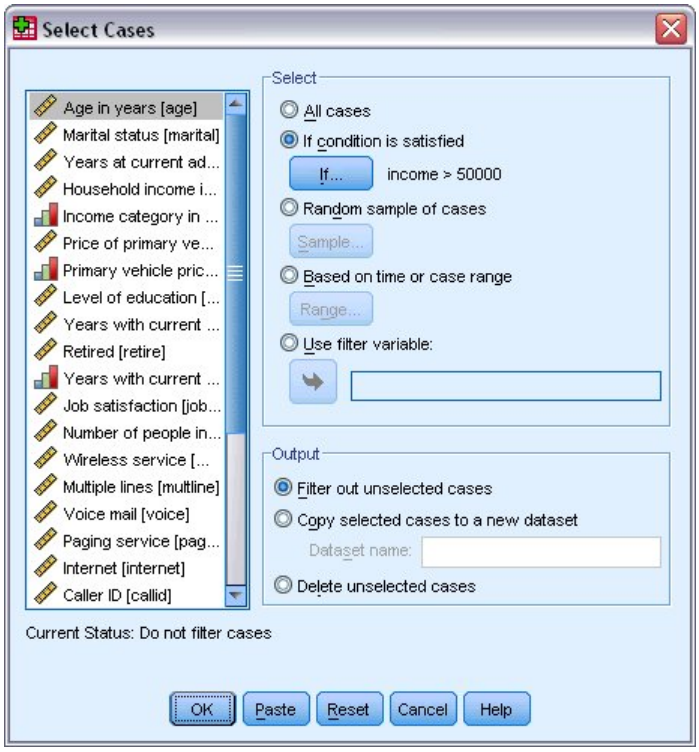


Abbildung 84. Dialogfeld "Fälle auswählen"

Auswählen von Fällen anhand eines bedingten Ausdrucks

So wählen Sie Fälle auf der Grundlage eines bedingten Ausdrucks aus:

1. Wählen Sie **Falls Bedingung zutrifft** aus und klicken Sie im Dialogfeld "Fälle auswählen" auf **Falls**.

Dadurch wird das Dialogfeld "Fälle auswählen: Falls" aufgerufen.

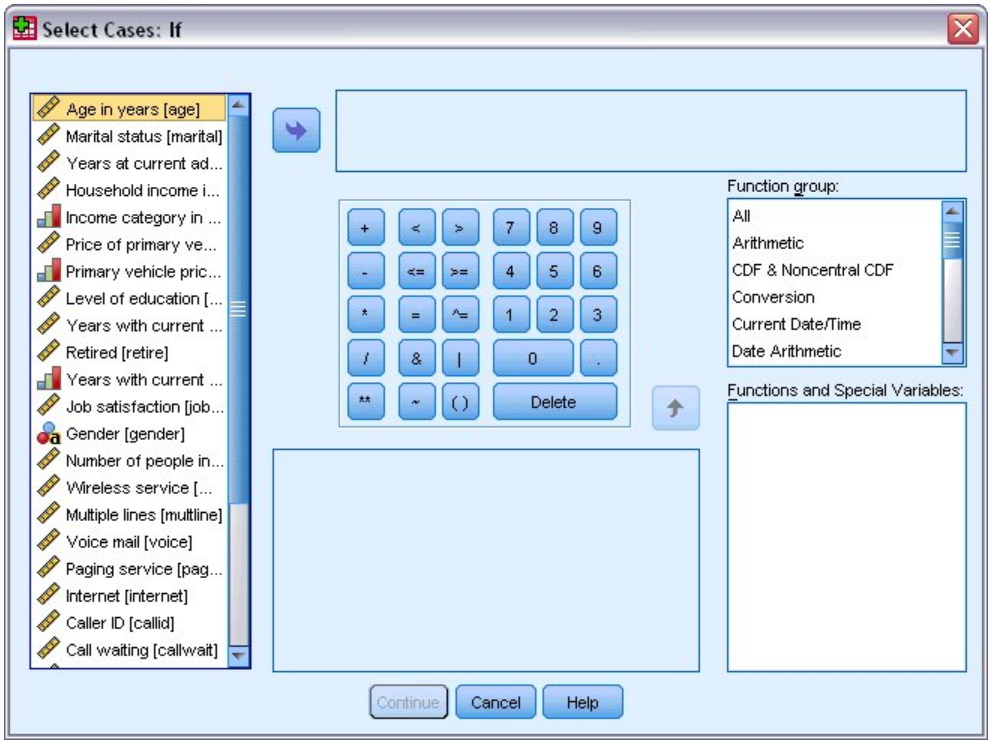


Abbildung 85. Dialogfeld "Fälle auswählen: Falls"

Der bedingte Ausdruck kann vorhandene Variablennamen, Konstanten, arithmetische Operatoren, logische Operatoren, relationale Operatoren und Funktionen enthalten. Sie können den Ausdruck im Textfeld genau wie Text in einem Ausgabefenster eingeben und bearbeiten. Sie können auch die Tastatur im Rechnerbereich, die Variablenliste und die Funktionsliste verwenden, um Elemente in den Ausdruck einzufügen. Weitere Informationen finden Sie in „Verwenden von bedingten Ausdrücken” auf Seite 71.

Auswählen einer Zufallsstichprobe aus den Fällen

So erhalten Sie eine Zufallsstichprobe:

1. Wählen Sie im Dialogfeld "Fälle auswählen" die Option **Zufallsstichprobe** aus.
2. Klicken Sie auf **Stichprobe**.

Dadurch wird das Dialogfeld "Fälle auswählen: Zufallsstichprobe" aufgerufen.

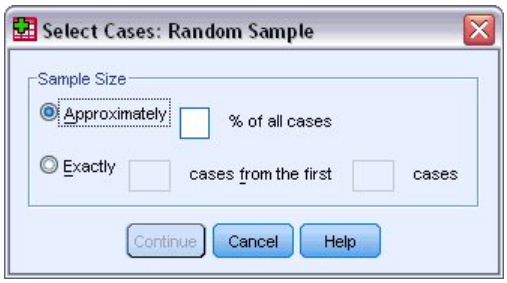


Abbildung 86. Dialogfeld "Fälle auswählen: Zufallsstichprobe"

Für die Stichprobengröße stehen die folgenden Optionen zur Auswahl:

- **Ungefähr.** Geben Sie einen Prozentsatz ein. SPSS erstellt eine Zufallsstichprobe, die ungefähr den angegebenen Prozentsatz aller Fälle enthält.
- **Ja, genau.** Geben Sie die gewünschte Anzahl der Fälle ein. Sie müssen außerdem die Anzahl der Fälle angeben, aus denen die Stichprobe gezogen werden soll. Diese zweite Zahl muss kleiner-gleich der Gesamtanzahl der Fälle in der Datendatei sein. Wenn die angegebene Anzahl die Gesamtanzahl der

Fälle in der Datendatei übersteigt, enthält die Stichprobe entsprechend weniger Fälle als die geforderte Anzahl.

Auswählen eines Zeit- oder Fallbereichs

So wählen Sie einen Fallbereich anhand von Datum, Uhrzeit oder Beobachtungs- bzw. Zeilennummern aus.

1. Wählen Sie die Option **Nach Zeit- oder Fallbereich** aus und klicken Sie im Dialogfeld "Fälle auswählen" auf **Bereich**.

   Dadurch wird das Dialogfeld "Fälle auswählen: Bereich" geöffnet, in dem Sie einen Bereich von Beobachtungs- bzw. Zeilennummern auswählen können.

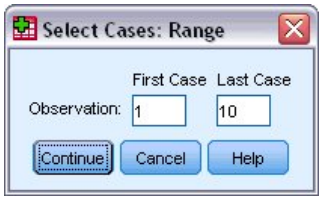


Abbildung 87. Dialogfeld "Fälle auswählen: Bereich"

- **Erster Fall.** Geben Sie das Startdatum und/oder die Startzeit für den Bereich ein. Wenn keine Datumsvariablen definiert sind, geben Sie die erste Beobachtungsnummer (die Zeilennummer im Dateneditor, wenn die Option "Datei aufteilen" inaktiviert ist) ein. Wenn Sie im Feld "Letzter Fall" keinen Wert angeben, werden alle Fälle vom Startdatum/von der Startzeit bis zum Ende der Zeitreihe ausgewählt.
- **Letzter Fall.** Geben Sie das Enddatum und/oder die Abschlusszeit für den Bereich ein. Wenn keine Datumsvariablen definiert sind, geben Sie die letzte Beobachtungsnummer (die Zeilennummer im Dateneditor, wenn die Option "Datei aufteilen" inaktiviert ist) ein. Wenn Sie im Feld "Erster Fall" keinen Wert angeben, werden alle Fälle vom Beginn der Zeitreihe bis zum Enddatum/zur Abschlusszeit ausgewählt.

Für Zeitreihendaten mit definierten Datumsvariablen können Sie einen Datumsbereich und/oder Zeitbereich auf der Grundlage von definierten Datumsvariablen auswählen. Jeder Fall steht für Beobachtungen zu einem anderen Zeitpunkt und die Datei ist in chronologischer Reihenfolge sortiert.

Abbildung 88. Dialogfeld "Fälle auswählen: Bereich" (Zeitreihen)

So generieren Sie Datumsvariablen für Zeitreihendaten:

2. Wählen Sie in den Menüs Folgendes aus:

   Daten > **Datum definieren...**

Behandlung von nicht ausgewählten Fällen

Die folgenden Optionen stehen für die Behandlung nicht ausgewählter Fälle zur Auswahl:

- **Nicht ausgewählte Fälle filtern.** Nicht ausgewählte Fälle werden nicht in die Analyse aufgenommen, verbleiben jedoch im Dataset. Sie können die nicht ausgewählten Fälle später in der Sitzung verwenden, wenn Sie die Filterfunktion inaktivieren. Wenn Sie eine Zufallsstichprobe oder Fälle anhand eines be-

dingten Ausdrucks auswählen, wird die Variable *filter_$* mit dem Wert 1 für ausgewählte Fälle und dem Wert 0 für nicht ausgewählte Fälle erzeugt.

- **Ausgewählte Fälle in neues Dataset kopieren.** Die ausgewählten Fälle werden in ein neues Dataset kopiert, das ursprüngliche Dataset bleibt unverändert. Nicht ausgewählte Fälle werden nicht in das neue Dataset übernommen. Sie werden im ursprünglichen Zustand im ursprünglichen Dataset belassen.
- **Nicht ausgewählte Fälle löschen.** Nicht ausgewählte Fälle werden aus dem Dataset gelöscht. Gelöschte Fälle können nur wiederhergestellt werden, indem Sie die Datei ohne Speichern der Änderungen schließen und sie dann erneut öffnen. Wenn Sie die Änderungen in der Datendatei speichern, werden die Fälle dauerhaft gelöscht.

Hinweis: Wenn Sie nicht ausgewählte Fälle löschen und die Datei speichern, können die Fälle nicht wiederhergestellt werden.

Status der Fallauswahl

Wenn Sie ein Subset von Fällen ausgewählt, nicht ausgewählte Fälle jedoch nicht verworfen haben, sind die nicht ausgewählten Fälle im Dateneditor mit einer diagonalen Linie durch die Zeilennummer gekennzeichnet.

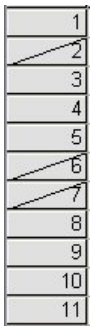


Abbildung 89. Status der Fallauswahl

Kapitel 10. Stichprobendateien

Die zusammen mit dem Produkt installierten Beispieldateien finden Sie im Unterverzeichnis *Samples* des Installationsverzeichnisses. Innerhalb des Unterverzeichnisses Samples gibt es für jede der folgenden Sprachen einen eigenen Ordner: Englisch, Französisch, Deutsch, Italienisch, Japanisch, Koreanisch, Polnisch, Vereinfachtes Chinesisch, Spanisch und Traditionelles Chinesisch.

Nicht alle Beispieldateien stehen in allen Sprachen zur Verfügung. Wenn eine Beispieldatei nicht in einer Sprache zur Verfügung steht, enthält der jeweilige Sprachordner eine englische Version der Beispieldatei.

Beschreibungen

Im Folgenden finden Sie Kurzbeschreibungen der in den verschiedenen Beispielen in der Dokumentation verwendeten Beispieldateien.

- **accidents.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um eine Versicherungsgesellschaft geht, die alters- und geschlechtsabhängige Risikofaktoren für Autounfälle in einer bestimmten Region untersucht. Jeder Fall entspricht einer Kreuzklassifikation von Alterskategorie und Geschlecht.
- **adl.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um Bemühungen geht, die Vorteile einer vorgeschlagenen Therapieform für Schlaganfallpatienten zu ermitteln. Ärzte teilten weibliche Schlaganfallpatienten nach dem Zufallsprinzip jeweils einer von zwei Gruppen zu. Die erste Gruppe erhielt die physische Standardtherapie, die zweite erhielt eine zusätzliche Emotionaltherapie. Drei Monate nach den Behandlungen wurden die Fähigkeiten der einzelnen Patienten, übliche Alltagsaktivitäten auszuführen, als ordinale Variablen bewertet.
- **advert.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen eines Einzelhändlers geht, die Beziehungen zwischen den in Werbung investierten Beträgen und den daraus resultierenden Umsätzen zu untersuchen. Zu diesem Zweck hat er die Umsätze vergangener Jahre mit den zugehörigen Werbeausgaben zusammengestellt.
- **aflatoxin.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um Tests von Maisernten auf Aflatoxin geht, ein Gift, dessen Konzentration stark zwischen und innerhalb von Ernteerträgen schwankt. Ein Getreideverarbeitungsbetrieb hat aus 8 Ernteerträgen je 16 Proben erhalten und das Aflatoxinniveau in Teilen pro Milliarde (PPB - Parts per Billion) gemessen.
- **anorectic.sav.** Während der Arbeit an einer standardisierten Symptomatik für anorektisches/bulimisches Verhalten führten die Forscher[1]eine Studie an 55 Jugendlichen mit bekannten Essstörungen durch. Jeder Patient wurde vier Mal über einen Zeitraum von vier Jahren untersucht, es fanden also insgesamt 220 Beobachtungen statt. Bei jeder Beobachtung erhielten die Patienten Scores für jedes von 16 Symptomen. Die Symptomscores fehlen für Patient 71 zum Zeitpunkt 2, Patient 76 zum Zeitpunkt 2 und Patient 47 zum Zeitpunkt 3, wodurch 217 gültige Beobachtungen verbleiben.
- **bankloan.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen einer Bank geht, den Anteil der nicht zurückgezahlten Kredite zu reduzieren. Die Datei enthält Informationen zum Finanzstatus und demografischen Hintergrund von 850 früheren und potenziellen Kunden. Bei den ersten 700 Fällen handelt es sich um Kunden, denen bereits ein Kredit gewährt wurde. Bei den letzten 150 Fällen handelt es sich um potenzielle Kunden, deren Kreditrisiko die Bank als gering oder hoch einstufen möchte.
- **bankloan_binning.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die Informationen zum Finanzstatus und demografischen Hintergrund von 5.000 früheren Kunden enthält.
- **behavior.sav.** In einem klassischen Beispiel[2]wurden 52 Studenten gebeten, die Kombinationen von 15 Situationen und 15 Verhaltensweisen auf einer 10-stufigen Skala von 0 = "äußerst angemessen" bis 9 =

[1] Van der Ham, T., J. J. Meulman, D. C. Van Strien, und H. Van Engeland. 1997. Empirically based subgrouping of eating disorders in adolescents: A longitudinal perspective. *British Journal of Psychiatry,* 170, 363-368.

[2] Price, R. H., und D. L. Bouffard. 1974. Behavioral appropriateness and situational constraints as dimensions of social behavior. *Journal of Personality and Social Psychology,* 30, 579-586.

"äußerst unangemessen" einzustufen. Die Werte werden über die einzelnen Personen gemittelt und als Unähnlichkeiten verwendet.

- **behavior_ini.sav.** Diese Datendatei enthält eine Ausgangskonfiguration für eine zweidimensionale Lösung für *behavior.sav*.
- **brakes.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Qualitätskontrolle in einer Fabrik geht, die Scheibenbremsen für Hochleistungsfahrzeuge herstellt. Die Datendatei enthält Messungen des Durchmessers von 16 Scheiben aus 8 Produktionsmaschinen. Der Zieldurchmesser für die Scheiben ist 322 Millimeter.
- **breakfast.sav.** In einer klassischen Studie[3]wurden 21 MBA-Studenten der Wharton School und ihre Ehepartner gebeten, 15 Elemente des Frühstücks in der Reihenfolge ihrer Präferenz von 1 = "am meisten bevorzugt" bis 15 = "am wenigsten bevorzugt" einzustufen. Die Präferenzen wurden in sechs unterschiedlichen Szenarios erfasst, von "Overall preference" (Allgemein bevorzugt) bis "Snack, with beverage only" (Imbiss, nur mit Getränk).
- **breakfast-overall.sav.** Diese Datei enthält die Daten zu den bevorzugten Frühstücksartikeln, allerdings nur für das erste Szenario, "Overall preference" (Allgemein bevorzugt).
- **broadband_1.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die die Anzahl der Abonnenten eines Breitbandservice, nach Region geordnet, enthält. Die Datendatei enthält die monatlichen Abonnentenzahlen für 85 Regionen über einen Zeitraum von vier Jahren.
- **broadband_2.sav.** Diese Datendatei stimmt mit *broadband_1.sav* überein, enthält jedoch Daten für weitere drei Monate.
- **car_insurance_claims.sav.** Ein an anderer Stelle vorgestellter und analysierter Datensatz,[4]betrifft Schadenersatzansprüche für Autos. Die durchschnittliche Höhe der Schadensansprüche lässt sich mit Gammaverteilung modellieren. Dazu wird eine inverse Verknüpfungsfunktion verwendet, um den Mittelwert der abhängigen Variablen mit einer linearen Kombination aus Alter des Versicherungsnehmers, Fahrzeugtyp und Fahrzeugalter in Bezug zu setzen. Die Anzahl der eingereichten Schadensansprüche kann als Skalierungsgewichtung verwendet werden.
- **car_sales.sav.** Diese Datendatei enthält hypothetische Verkaufsschätzungen, Listenpreise und physische Spezifikationen für verschiedene Fahrzeugfabrikate und -modelle. Die Listenpreise und physischen Spezifikationen wurden von *edmunds.com* und Herstellerwebsites entnommen.
- **car_sales_uprepared.sav.** Hierbei handelt es sich um eine modifizierte Version der Datei *car_sales.sav*, die keinerlei transformierte Versionen der Felder enthält.
- **carpet.sav.** In einem vielfach eingesetzten Beispiel[5]möchte ein Unternehmen, das einen neuen Teppichreiniger auf den Markt bringen will, den Einfluss von fünf Faktoren auf die Kundenpräferenz untersuchen: Verpackungsdesign, Markenname, Preis, *Good Housekeeping*-Siegel und Geld-zurück-Garantie. Die Verpackungsgestaltung setzt sich aus drei Faktorenebenen zusammen, die sich durch die Position der Auftragebürste unterscheiden. Außerdem gibt es drei Markennamen (*K2R*, *Glory* und *Bissell*), drei Preisstufen sowie je zwei Ebenen (Nein oder Ja) für die letzten beiden Faktoren. 10 Kunden stufen 22 Profile ein, die durch diese Faktoren definiert sind. Die Variable *Preference* enthält den Rang der durchschnittlichen Einstufung für die verschiedenen Profile. Ein niedriger Rang bedeutet eine starke Präferenz. Diese Variable gibt ein Gesamtmaß der Präferenz für die Profile an.
- **carpet_prefs.sav.** Diese Datendatei beruht auf denselben Beispielen, wie für *carpet.sav* beschrieben, enthält jedoch die tatsächlichen Einstufungen durch jeden der 10 Kunden. Die Kunden wurden gebeten, die 22 Produktprofile in der Reihenfolge ihrer Präferenzen einzustufen. Die Variablen *PREF1* bis *PREF22* enthalten die IDs der zugeordneten Profile, wie in *carpet_plan.sav* definiert.
- **catalog.sav.** Diese Datendatei enthält hypothetische monatliche Verkaufszahlen für drei Produkte, die von einem Versandhaus verkauft werden. Daten für fünf mögliche Prädiktorvariablen wurden ebenfalls aufgenommen.

---

[3] Green, P. E., und V. Rao. 1972. *Applied multidimensional scaling*. Hinsdale, Ill.: Dryden Press.

[4] McCullagh, P., und J. A. Nelder. 1989. *Verallgemeinerte lineare Modelle*, 2nd ed. London: Chapman & Hall.

[5] Green, P. E., und Y. Wind. 1973. *Multiattribute decisions in marketing: A measurement approach*. Hinsdale, Ill.: Dryden Press.

- **catalog_seasfac.sav.** Diese Datendatei ist mit *catalog.sav* identisch, außer, dass ein Set von saisonalen Faktoren, die mithilfe der Prozedur "Saisonale Zerlegung" berechnet wurden, sowie die zugehörigen Datumsvariablen hinzugefügt wurden.
- **cellular.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen eines Mobiltelefonunternehmens geht, die Kundenabwanderung zu verringern. Propensity-Scores für die Abwanderungsneigung (von 0 bis 100) werden auf die Kunden angewendet. Kunden mit einem Score von 50 oder höher streben vermutlich einen Anbieterwechsel an.
- **ceramics.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen eines Herstellers geht, der ermitteln möchte, ob ein neue, hochwertige Keramiklegierung eine größere Hitzebeständigkeit aufweist als eine Standardlegierung. Jeder Fall entspricht einem Test einer der Legierungen; die Temperatur, bei der das Keramikwälzlager versagte, wurde erfasst.
- **cereal.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um eine Umfrage geht, bei der 880 Personen nach ihren Frühstückgewohnheiten befragt wurden. Außerdem wurden Alter, Geschlecht, Familienstand und Vorliegen bzw. Nichtvorliegen eines aktiven Lebensstils (auf der Grundlage von mindestens zwei Trainingseinheiten pro Woche) erfasst. Jeder Fall entspricht einem Befragten.
- **clothing_defects.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Qualitätskontrolle in einer Bekleidungsfabrik geht. Aus jeder in der Fabrik produzierten Charge entnehmen die Kontrolleure eine Stichprobe an Bekleidungsartikeln und zählen die Anzahl der Bekleidungsartikel die inakzeptabel sind.
- **coffee.sav.** Diese Datendatei bezieht sich auf die wahrgenommenen Bilder von sechs Eiskaffeemarken[6]. Für jedes der 23 Bildattribute von Eiskaffee haben die Befragten alle Marken ausgewählt, die durch das Attribut beschrieben wurden. Die sechs Marken werden als "AA", "BB", "CC", "DD", "EE" und "FF" bezeichnet, um Vertraulichkeit zu gewährleisten.
- **contacts.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Kontaktlisten einer Gruppe von Vertretern geht, die Computer an Unternehmen verkaufen. Die einzelnen Kontaktpersonen werden anhand der Abteilung, in der sie in ihrem Unternehmen arbeiten und anhand ihrer Stellung in der Unternehmenshierarchie in Kategorien eingeteilt. Außerdem werden der Betrag des letzten Verkaufs, die Zeit seit dem letzten Verkauf und die Größe des Unternehmens, in dem die Kontaktperson arbeitet, aufgezeichnet.
- **creditpromo.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen eines Kaufhauses geht, die Wirksamkeit einer kürzlich durchgeführten Kreditkartenwerbeaktion einzuschätzen. Dazu wurden 500 Karteninhaber nach dem Zufallsprinzip ausgewählt. Die Hälfte erhielt eine Werbebeilage, die einen reduzierten Zinssatz für Einkäufe in den nächsten drei Monaten ankündigte. Die andere Hälfte erhielt eine Standardwerbebeilage.
- **customer_dbase.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen eines Unternehmens geht, das die Informationen in seinem Data Warehouse nutzen möchte, um spezielle Angebote für Kunden zu erstellen, die mit der größten Wahrscheinlichkeit darauf ansprechen. Nach dem Zufallsprinzip wurde ein Subset des Kundenstamms ausgewählt. Diese Kunden erhielten die speziellen Angebote und die Reaktionen wurden aufgezeichnet.
- **customer_information.sav.** Eine hypothetische Datendatei mit Kundenmailing-Daten wie Name und Adresse.
- **customer_subset.sav.** Ein Subset von 80 Fällen aus der Datei *customer_dbase.sav*.
- **debate.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die paarige Antworten auf eine Umfrage unter den Zuhörern einer politischen Debatte enthält (Antworten vor und nach der Debatte). Jeder Fall entspricht einem Befragten.
- **debate_aggregate.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, in der die Antworten aus *debate.sav* aggregiert wurden. Jeder Fall entspricht einer Kreuzklassifikation der favorisierten Politiker vor und nach der Debatte.

[6] Kennedy, R., C. Riquier und B. Scharf. 1996. Practical applications of correspondence analysis to categorical data in market research. *Journal of Targeting, Measurement, and Analysis for Marketing,* 5, 56-70.

- **demo.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um eine Kundendatenbank geht, die zum Zwecke der Zusendung monatlicher Angebote erworben wurde. Neben verschiedenen demografischen Informationen ist erfasst, ob der Kunde auf das Angebot geantwortet hat.
- **demo_cs_1.sav.** Hierbei handelt es sich um eine hypothetische Datendatei für den ersten Schritt eines Unternehmens, das eine Datenbank mit Umfrageinformationen zusammenstellen möchte. Jeder Fall entspricht einer anderen Stadt. Außerdem sind IDs für Region, Provinz, Landkreis und Stadt erfasst.
- **demo_cs_2.sav.** Hierbei handelt es sich um eine hypothetische Datendatei für den zweiten Schritt eines Unternehmens, das eine Datenbank mit Umfrageinformationen zusammenstellen möchte. Jeder Fall entspricht einem anderen Stadtteil aus den im ersten Schritt ausgewählten Städten. Außerdem sind IDs für Region, Provinz, Landkreis, Stadt, Stadtteil und Wohneinheit erfasst. Die Informationen zur Stichprobenziehung aus den ersten beiden Stufen des Stichprobenplans sind ebenfalls enthalten.
- **demo_cs.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die Umfrageinformationen enthält, die mit einem komplexen Stichprobendesign erfasst wurden. Jeder Fall entspricht einer anderen Wohneinheit. Es sind verschiedene Informationen zum demografischen Hintergrund und zur Stichprobenziehung erfasst.
- **diabetes_costs.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die Informationen enthält, die von einer Versicherungsgesellschaft für an Diabetes erkrankte Versicherungsnehmer verwaltet werden. Jeder Fall entspricht einem anderen Versicherungsnehmer.
- **dietstudy.sav.** Diese hypothetische Datendatei enthält die Ergebnisse einer Studie über die "Stillman-Diät"[7]. Jeder Fall entspricht einem separaten Proband und zeichnet sein oder ihr Gewicht vor und nach der Diät in Pfund und den Triglyzeridspiegel in mg/100 ml auf.
- **dmdata.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die Informationen über Demografie und Einkäufe für ein Direktmarketing-Unternehmen enthält. *dmdata2.sav* enthält Informationen für ein Subset von Kontakten, die eine Testsendung erhalten, und *dmdata3.sav* enthält Informationen zu den verbleibenden Kontakten, die keine Testsendung erhalten.
- **dvdplayer.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Entwicklung eines neuen DVD-Players geht. Mithilfe eines Prototyps hat das Marketingteam Zielgruppendaten erfasst. Jeder Fall entspricht einem befragten Benutzer und enthält demografische Daten zu dem Benutzer sowie dessen Antworten auf Fragen zum Prototyp.
- **german_credit.sav.** Diese Datendatei wurde aus dem Datensatz "German credit" im Repository der Datenbanken für maschinelles Lernen entnommen[8]an der University of California, Irvine.
- **grocery_1month.sav.** Bei dieser hypothetischen Datendatei handelt es sich um die Datendatei *grocery_coupons.sav*, wobei die wöchentlichen Einkäufe zusammengefasst sind, sodass jeder Fall einem anderen Kunden entspricht. Dadurch entfallen einige der Variablen, die wöchentlichen Änderungen unterworfen waren, und der verzeichnete ausgegebene Betrag ist nun die Summe der Beträge, die in den vier Wochen der Studie ausgegeben wurden.
- **grocery_coupons.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die Umfragedaten enthält, die von einer Lebensmittelkette erfasst wurden, die sich für die Kaufgewohnheiten ihrer Kunden interessiert. Jeder Kunde wird über vier Wochen beobachtet und jeder Fall entspricht einer Kundenwoche und enthält Informationen zu den Geschäften, in denen der Kunde einkauft sowie zu anderen Merkmalen, beispielsweise welcher Betrag in der betreffenden Woche für Lebensmittel ausgegeben wurde.
- **guttman.sav.** Glocke[9]präsentierte eine Tabelle zur Veranschaulichung möglicher sozialer Gruppen. Guttman[10]verwendete einen Teil dieser Tabelle, in der fünf Variablen, die Dinge wie soziale Interaktion, Zugehörigkeitsgefühl zu einer Gruppe, physische Nähe der Mitglieder und Formalität der Beziehung be-

[7] Rickman, R., :NONE. Mitchell, J. Dingman, und J. E. Dalen. 1974. Changes in serum cholesterol during the Stillman Diet. *Journal of the American Medical Association,* 228:, 54-58.

[8] Blake, C. L., und C. J. Merz. 1998. "UCI Repository of machine learning databases." Verfügbar unter http://www.ics.uci.edu/~mlearn/MLRepository.html.

[9] Bell, E. H. 1961. *Social foundations of human behavior: Introduction to the study of sociology*. New York: Harper & Row.

[10] Guttman, L. 1968. A general nonmetric technique for finding the smallest coordinate space for configurations of points. *Psychometrika,* 33, 469-506.

schreiben, mit sieben theoretischen sozialen Gruppen gekreuzt wurden, dazu einschließlich Menschenmengen (zum Beispiel Menschen bei einem Fußballspiel), Publikum (zum Beispiel Menschen in einem Theater oder einer Vorlesung), Öffentlichkeit (zum Beispiel Zeitungs- oder Fernsehpublikum), Mobs (wie eine Menschenmenge, aber mit viel intensiverer Interaktion), Primärgruppen (intim), Sekundärgruppen (freiwillig) und die moderne Gemeinschaft (loser Zusammenschluss mit enger räumlicher Nähe und einem Bedarf an spezialisierten Services).

- **health_funding.sav.** Hierbei handelt es sich um eine hypothetische Datei, die Daten zur Finanzierung des Gesundheitswesens (Betrag pro 100 Personen), Krankheitsraten (Rate pro 10.000 Personen der Bevölkerung) und Besuche bei medizinischen Einrichtungen/Ärzten (Rate pro 10.000 Personen der Bevölkerung) enthält. Jeder Fall entspricht einer anderen Stadt.
- **hivassay.sav.** Hierbei handelt es sich um eine hypothetische Datendatei zu den Bemühungen eines pharmazeutischen Labors, einen Schnelltest zur Erkennung von HIV-Infektionen zu entwickeln. Die Ergebnisse des Tests sind acht kräftiger werdende Rotschattierungen, wobei kräftigeren Schattierungen auf eine höhere Infektionswahrscheinlichkeit hindeuten. Bei 2.000 Blutproben, von denen die Hälfte mit HIV infiziert war, wurde ein Labortest durchgeführt.
- **hourlywagedata.sav.** Hierbei handelt es sich um eine hypothetische Datendatei zum Stundenlohn von Pflegepersonal in Praxen und Krankenhäusern mit unterschiedlich langer Berufserfahrung.
- **insurance_claims.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um eine Versicherungsgesellschaft geht, die ein Modell zur Kennzeichnung verdächtiger, potenzieller Betrugsversuche erstellen möchte. Jeder Fall entspricht einem Anspruch.
- **insure.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um eine Versicherungsgesellschaft geht, die die Risikofaktoren untersucht, die darauf hinweisen, ob ein Kunde die Leistungen einer mit einer Laufzeit von 10 Jahren abgeschlossenen Lebensversicherung in Anspruch nehmen wird. Jeder Fall in der Datendatei entspricht einem Paar von Verträgen, je einer mit Leistungsforderung und der andere ohne, wobei die beiden Versicherungsnehmer in Alter und Geschlecht übereinstimmen.
- **judges.sav.** Hierbei handelt es sich um eine hypothetische Datendatei mit den Wertungen von ausgebildeten Kampfrichtern (sowie eines Sportliebhabers) zu 300 Kunstturnleistungen. Jede Zeile stellt eine Leistung dar; die Kampfrichter bewerteten jeweils dieselben Leistungen.
- **kinship_dat.sav.** Rosenberg und Kim[11]untersuchten 15 Verwandtschaftsbegriffe (Tante, Bruder, Cousin, Tochter, Vater, Enkelin, Großvater, Großmutter, Enkel, Mutter, Neffe, Nichte, Schwester, Sohn, Onkel). Die beiden Analytiker baten vier Gruppen von College-Studenten (zwei weibliche und zwei männliche Gruppen), diese Bezeichnungen auf der Grundlage der Ähnlichkeiten zu sortieren. Zwei Gruppen (eine weibliche und eine männliche Gruppe) wurden gebeten, die Bezeichnungen zweimal zu sortieren; die zweite Sortierung sollte dabei nach einem anderen Kriterium erfolgen als die erste. So wurden insgesamt sechs "Quellen" erzielt. Jede Quelle entspricht einer Ähnlichkeitsmatrix mit 15 x 15 Elementen. Die Anzahl der Zellen ist dabei gleich der Anzahl der Personen in einer Quelle minus der Anzahl der gemeinsamen Platzierungen der Objekte in dieser Quelle.
- **kinship_ini.sav.** Diese Datendatei enthält eine Ausgangskonfiguration für eine dreidimensionale Lösung für *kinship_dat.sav*.
- **kinship_var.sav.** Diese Datendatei enthält die unabhängigen Variablen *gender* (Geschlecht), *gener* (Generation) und *degree* (Verwandtschaftsgrad), die zur Interpretation der Dimensionen einer Lösung für *kinship_dat.sav* verwendet werden können. Insbesondere können sie verwendet werden, um den Lösungsraum auf eine lineare Kombination dieser Variablen zu beschränken.
- **marketvalues.sav.** Diese Datei betrifft den Verkauf von Wohnräumen in einer neuen Wohnbauentwicklung in Algonquin, Illinois, in den Jahren 1999-2000. Diese Verkäufe sind in Grundbucheinträgen dokumentiert.
- **nhis2000_subset.sav.** Die "National Health Interview Survey (NHIS)" ist eine große, bevölkerungsbezogene Umfrage unter der US-amerikanischen Zivilbevölkerung. Es werden persönliche Interviews in einer landesweit repräsentativen Stichprobe von Haushalten durchgeführt. Für die Mitglieder jedes Haushalts werden demografische Informationen und Beobachtungen zum Gesundheitsverhalten und

[11] Rosenberg, S., und M. P. Kim. 1975. The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research,* 10, 489-502.

Gesundheitsstatus eingeholt. Diese Datendatei enthält ein Subset der Informationen aus der Umfrage des Jahres 2000. National Center for Health Statistics. National Health Interview Survey, 2000. Allgemein zugängliche Datendatei und Dokumentation. *ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHIS/2000/*. Zugriff erfolgte 2003.

- **ozone.sav.** Die Daten enthalten 330 Beobachtungen zu sechs meteorologischen Variablen zur Vorhersage der Ozonkonzentration aus den übrigen Variablen. Frühere Forscher[12],[13]fanden unter anderem Nichtlinearitäten zwischen diesen Variablen, die Standardregressionsansätze behindern.
- **pain_medication.sav.** Diese hypothetische Datendatei enthält die Ergebnisse eines klinischen Tests für ein entzündungshemmendes Medikament zur Schmerzbehandlung bei chronischer Arthritis. Von besonderem Interesse ist die Zeitdauer, bis die Wirkung des Medikaments einsetzt und wie es im Vergleich mit bestehenden Medikamenten abschneidet.
- **patient_los.sav.** Diese hypothetische Datendatei enthält die Behandlungsaufzeichnungen zu Patienten, die wegen des Verdachts auf Herzinfarkt in das Krankenhaus eingeliefert wurden. Jeder Fall entspricht einem Patienten und enthält diverse Variablen in Bezug auf den Krankenhausaufenthalt.
- **patlos_sample.sav.** Diese hypothetische Datendatei enthält die Behandlungsaufzeichnungen für eine Stichprobe von Patienten, denen während der Behandlung eines Herzinfarkts Thrombolytika verabreicht wurden. Jeder Fall entspricht einem Patienten und enthält diverse Variablen in Bezug auf den Krankenhausaufenthalt.
- **poll_cs.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um Bemühungen geht, die öffentliche Unterstützung für einen Gesetzentwurf zu ermitteln, bevor er im Parlament eingebracht wird. Die Fälle entsprechen registrierten Wählern. Für jeden Fall sind County, Gemeinde und Wohnviertel des Wählers erfasst.
- **poll_cs_sample.sav.** Diese hypothetische Datendatei enthält eine Stichprobe der in *poll_cs.sav* aufgeführten Wähler. Die Stichprobe wurde gemäß dem in der Plandatei *poll.csplan* angegebenen Stichprobenplan gezogen und in dieser Datendatei sind die Einschlusswahrscheinlichkeiten und Stichprobengewichtungen erfasst. Beachten Sie jedoch Folgendes: Da im Stichprobenplan die PPS-Methode (Probability Proportional To Size - Wahrscheinlichkeit proportional zur Größe) verwendet wird, gibt es außerdem eine Datei mit den gemeinsamen Auswahlwahrscheinlichkeiten (*poll_jointprob.sav*). Die zusätzlichen Variablen zum demografischen Hintergrund der Wähler und ihrer Meinung zum vorgeschlagenen Gesetzentwurf wurden nach der Ziehung der Stichprobe erfasst und zur Datendatei hinzugefügt.
- **property_assess.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, in der es um die Bemühungen eines für einen Bezirk (County) zuständigen Immobilienbewerters geht, trotz eingeschränkter Ressourcen die Einschätzungen des Werts von Immobilien auf dem aktuellsten Stand zu halten. Die Fälle entsprechen den Immobilien, die im vergangenen Jahr in dem betreffenden County verkauft wurden. Jeder Fall in der Datendatei enthält die Gemeinde, in der sich die Immobilie befindet, den Bewerter, der die Immobilie besichtigt hat, die seit dieser Bewertung verstrichene Zeit, den zu diesem Zeitpunkt ermittelten Wert sowie den Verkaufswert der Immobilie.
- **property_assess_cs.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, in der es um die Bemühungen eines für einen US-Bundesstaat zuständigen Immobilienbewerters geht, trotz eingeschränkter Ressourcen die Einschätzungen des Werts von Immobilien auf dem aktuellsten Stand zu halten. Die Fälle entsprechen den Immobilien in dem betreffenden Bundesstaat. Jeder Fall in der Datendatei enthält das County, die Gemeinde und das Wohnviertel, in dem sich die Immobilie befindet, die seit der letzten Bewertung verstrichene Zeit sowie den zu diesem Zeitpunkt ermittelten Wert.
- **property_assess_cs_sample.sav.** Diese hypothetische Datendatei enthält eine Stichprobe der in *property_assess_cs.sav* aufgeführten Immobilien. Die Stichprobe wurde gemäß dem in der Plandatei *property_assess.csplan* angegebenen Stichprobenplan gezogen und in dieser Datendatei sind die Einschlusswahrscheinlichkeiten und Stichprobengewichtungen erfasst. Die zusätzliche Variable *Current value* (Aktueller Wert) wurde nach der Ziehung der Stichprobe erfasst und zur Datendatei hinzugefügt.
- **recidivism.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen einer Strafverfolgungsbehörde geht, einen Einblick in die Rückfallraten in ihrem Zuständigkeits-

[12] Breiman, L., und J. H. Friedman. 1985. Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Association,* 80, 580-598.

[13] Hastie, T., und R. Tibshirani. 1990. *Generalized additive models*. London: Chapman and Hall.

bereich zu gewinnen. Jeder Fall entspricht einem früheren Straftäter und erfasst Daten zu dessen demografischen Hintergrund, einige Details zu seinem ersten Verbrechen sowie die Zeit bis zu seiner zweiten Festnahme, sofern diese innerhalb von zwei Jahren nach der ersten Festnahme erfolgte.

- **recidivism_cs_sample.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen einer Strafverfolgungsbehörde geht, einen Einblick in die Rückfallraten in ihrem Zuständigkeitsbereich zu gewinnen. Jeder Fall entspricht einem früheren Straftäter, der im Juni 2003 erstmals aus der Haft entlassen wurde, und erfasst Daten zu dessen demografischen Hintergrund, einige Details zu seinem ersten Verbrechen sowie die Daten zu seiner zweiten Festnahme, sofern diese bis Ende Juni 2006 erfolgte. Die Straftäter wurden aus per Stichprobenziehung ermittelten Polizeidirektionen ausgewählt (gemäß dem in *recidivism_cs.csplan* angegebenen Stichprobenplan). Da hierbei eine PPS-Methode (Probability Proportional To Size - Wahrscheinlichkeit proportional zur Größe) verwendet wird, gibt es außerdem eine Datei mit den gemeinsamen Auswahlwahrscheinlichkeiten (*recidivism_cs_jointprob.sav*).
- **rfm_transactions.sav.** Eine hypothetische Datendatei mit Kauftransaktionsdaten wie Kaufdatum, gekauften Artikeln und Geldbetrag für jede Transaktion.
- **salesperformance.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um Evaluierung von zwei neuen Verkaufsschulungen geht. 60 Mitarbeiter, die in drei Gruppen unterteilt sind, erhalten jeweils eine Standardschulung. Zusätzlich erhält Gruppe 2 eine technische Schulung und Gruppe 3 eine Praxisschulung. Die einzelnen Mitarbeiter wurden am Ende der Schulung einem Test unterzogen und der erzielte Score wurde erfasst. Jeder Fall in der Datendatei stellt einen Lehrgangsteilnehmer dar und enthält die Gruppe, der der Lehrgangsteilnehmer zugeteilt wurde sowie der von ihm in der Prüfung erreichte Score.
- **satisf.sav.** Hierbei handelt es sich um eine hypothetische Datendatei zu einer Zufriedenheitsumfrage, die von einem Einzelhandelsunternehmen in 4 Filialen durchgeführt wurde. Insgesamt wurden 582 Kunden befragt. Jeder Fall gibt die Antworten eines einzelnen Kunden wieder.
- **screws.sav.** Diese Datendatei enthält Informationen über die Eigenschaften von Schrauben, Bolzen, Muttern und Nägeln[14].
- **shampoo_ph.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Qualitätskontrolle in einer Fabrik für Haarpflegeprodukte geht. In regelmäßigen Zeitintervallen werden Messwerte von sechs separaten Ausgangschargen erhoben und ihr pH-Wert erfasst. Der Zielbereich ist 4,5-5,5.
- **ships.sav.** Ein an anderer Stelle vorgestellter und analysierter Datensatz,[15]der durch Wellen verursachte Schäden an Frachtschiffen betrifft. Die Vorfallshäufigkeiten können unter Angabe von Schiffstyp, Konstruktionszeitraum und Betriebszeitraum gemäß einer Poisson-Rate modelliert werden. Das Aggregat der Betriebsmonate für jede Zelle der durch die Kreuzklassifizierung der Faktoren gebildeten Tabelle gibt die Werte für die Risikoanfälligkeit an.
- **site.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen eines Unternehmens geht, neue Standorte für die betriebliche Expansion auszuwählen. Das Unternehmen beauftragte zwei Berater unabhängig voneinander mit der Bewertung der Standorte. Neben einem umfassenden Bericht gaben die Berater auch eine zusammenfassende Wertung für jeden Standort als "good" (gut) "fair" (mittelmäßig) oder "poor" (schlecht) ab.
- **smokers.sav.** Diese Datendatei ist ein Auszug aus der National Household Survey of Drug Abuse von 1998 und stellt eine Wahrscheinlichkeitsstichprobe amerikanischer Haushalte dar. (http://dx.doi.org/10.3886/ICPSR02934) Der erste Schritt bei der Analyse dieser Datendatei sollte daher darin bestehen, die Daten zu gewichten, um die Bevölkerungsentwicklung widerzuspiegeln.
- **stocks.sav.** Diese hypothetische Datendatei umfasst Börsenkurse und -volumina für ein Jahr.
- **stroke_clean.sav.** Diese hypothetische Datendatei enthält den Zustand einer medizinischen Datenbank, nachdem diese mithilfe der Prozeduren in der Option "Data Preparation" bereinigt wurde.
- **stroke_invalid.sav.** Diese hypothetische Datendatei enthält den ursprünglichen Zustand einer medizinischen Datenbank, der mehrere Dateneingabefehler aufweist.

[14] Hartigan, J. A. 1975. *Clustering algorithms*. New York: John Wiley and Sons.
[15] McCullagh, P., und J. A. Nelder. 1989. *Verallgemeinerte lineare Modelle*, 2nd ed. London: Chapman & Hall.

- **stroke_survival.** In dieser hypothetischen Datendatei geht es um die Überlebenszeiten von Patienten, die nach einem Rehabilitationsprogramm wegen eines ischämischen Schlaganfalls mit einer Reihe von Problemen zu kämpfen haben. Nach dem Schlaganfall werden das Auftreten von Herzinfarkt, ischämischem Schlaganfall und hämorrhagischem Schlaganfall sowie der Zeitpunkt des Ereignisses aufgezeichnet. Die Stichprobe ist auf der linken Seite abgeschnitten, da sie nur Patienten enthält, die bis zum Ende des Rehabilitationsprogramms, das nach dem Schlaganfall durchgeführt wurde, überlebten.
- **stroke_valid.sav.** Diese hypothetische Datendatei enthält den Zustand einer medizinischen Datenbank, nachdem diese mithilfe der Prozedur "Daten validieren" überprüft wurde. Sie enthält immer noch potenziell anomale Fälle.
- **survey_sample.sav.** Diese Datendatei enthält Umfragedaten einschließlich demografischer Daten und verschiedener Meinungskennzahlen. Sie beruht auf einem Subset der Variablen aus der NORC General Social Survey aus dem Jahr 1998. Allerdings wurden zu Demonstrationszwecken einige Daten abgeändert und weitere fiktive Variablen hinzugefügt.
- **tcm_kpi.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die die Werte der wöchentlichen wesentlichen Leistungsindikatoren für ein Unternehmen enthält. Sie enthält darüber hinaus wöchentliche Daten für zahlreiche steuerbare Metriken über denselben Zeitraum.
- **tcm_kpi_upd.sav.** Diese Datendatei stimmt mit *tcm_kpi.sav* überein, enthält jedoch Daten für weitere vier Wochen.
- **telco.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen eines Telekommunikationsunternehmens geht, die Kundenabwanderung zu verringern. Jeder Fall entspricht einem Kunden und enthält verschiedene Informationen zum demografischen Hintergrund und zur Servicenutzung.
- **telco_extra.sav.** Diese Datendatei ähnelt der Datendatei *telco.sav*, aber die Variablen "Beschäftigungsdauer" und protokolltransformierte Kundenausgaben wurden entfernt und durch standardisierte, protokolltransformierte Variablen für Kundenausgaben ersetzt.
- **telco_missing.sav.** Diese Datendatei ist ein Subset der Datendatei *telco.sav*, allerdings wurde ein Teil der demografischen Datenwerte durch fehlende Werte ersetzt.
- **testmarket.sav.** Diese hypothetische Datendatei bezieht sich auf die Pläne einer Fast-Food-Kette, einen neuen Artikel in ihr Menü aufzunehmen. Es gibt drei mögliche Kampagnen zur Verkaufsförderung für das neue Produkt. Daher wird der neue Artikel in Filialen in mehreren zufällig ausgewählten Märkten eingeführt. An jedem Standort wird eine andere Form der Verkaufsförderung verwendet und die wöchentlichen Verkaufszahlen für das neue Produkt werden für die ersten vier Wochen aufgezeichnet. Jeder Fall entspricht einer Standortwoche.
- **testmarket_1month.sav.** Bei dieser hypothetischen Datendatei handelt es sich um die Datendatei *testmarket.sav*, wobei die wöchentlichen Verkaufszahlen zusammengefasst sind, sodass jeder Fall einem Standort entspricht. Dadurch entfallen einige der Variablen, die wöchentlichen Änderungen unterworfen waren und die verzeichneten Verkaufszahlen sind nun die Summe der Verkaufszahlen während der vier Wochen der Studie.
- **tree_car.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die demografische Daten sowie Daten zum Kaufpreis von Fahrzeugen enthält.
- **tree_credit.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die demografische Daten sowie Daten zu früheren Bankkrediten enthält.
- **tree_missing_data.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die demografische Daten sowie Daten zu früheren Bankkrediten enthält und eine große Anzahl fehlender Werte aufweist.
- **tree_score_car.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, die demografische Daten sowie Daten zum Kaufpreis von Fahrzeugen enthält.
- **tree_textdata.sav.** Eine einfache Datendatei mit nur zwei Variablen, die vor allem den Standardzustand von Variablen vor der Zuweisung von Messniveau und Wertbeschriftungen zeigen soll.
- **tv-survey.sav.** Hierbei handelt es sich um eine hypothetische Datendatei zu einer Studie, die von einem Fernsehstudio durchgeführt wurde, das überlegt, ob die Laufzeit eines erfolgreichen Programms verlängert werden soll. 906 Personen wurden gefragt, ob sie das Programm unter verschiedenen Bedingungen ansehen würden. Jede Zeile entspricht einem Befragten; jede Spalte entspricht einer Bedingung.

- **ulcer_recurrence.sav.** Diese Datei enthält Teilinformationen aus einer Studie zum Vergleich der Wirksamkeit zweier Therapien zur Vermeidung des Wiederauftretens von Geschwüren. Sie bietet ein gutes Beispiel für intervallzensierte Daten und wurde an anderer Stelle vorgestellt und analysiert[16].
- **ulcer_recurrence_recoded.sav.** In dieser Datei sind die Daten aus *ulcer_recurrence.sav* so umstrukturiert, dass das Modell der Ereigniswahrscheinlichkeit für jedes Intervall der Studie berechnet werden kann und nicht nur die Ereigniswahrscheinlichkeit am Ende der Studie. Es wurde an anderer Stelle vorgestellt und analysiert[17].
- **verd1985.sav.** Diese Datendatei betrifft eine Umfrage[18]. Die Antworten von 15 Probanden auf 8 Variablen wurden aufgezeichnet. Die relevanten Variablen sind in drei Sets unterteilt. Set 1 umfasst *age* und *marital*, Set 2 besteht aus *pet* und *news* und in Set 3 finden sich *music* und *live*. Die Variable *pet* wird mehrfach nominal skaliert und die Variable *age* ordinal. Alle anderen Variablen werden einzeln nominal skaliert.
- **virus.sav.** Hierbei handelt es sich um eine hypothetische Datendatei, bei der es um die Bemühungen eines Internet-Service-Providers geht, der die Auswirkungen eines Virus auf seine Netze ermitteln möchte. Dabei wurde vom Moment der Virusentdeckung bis zu dem Zeitpunkt, zu dem die Virusinfektion unter Kontrolle war, der (ungefähre) prozentuale Anteil infizierter E-Mail in den Netzen erfasst.
- **wheeze_steubenville.sav.** Dies ist ein Teil einer Längsschnittstudie über die gesundheitlichen Auswirkungen der Luftverschmutzung auf Kinder[19]. Die Daten enthalten wiederholte binäre Messungen des Keuchstatus für Kinder aus Steubenville, Ohio, im Alter von 7, 8, 9 und 10 Jahren, zusammen mit einer festen Aufzeichnung darüber, ob die Mutter im ersten Jahr der Studie Raucherin war oder nicht.
- **workprog.sav.** Hierbei handelt es sich um eine hypothetische Datendatei zu einem Arbeitsprogramm der Regierung, das versucht, benachteiligten Personen bessere Arbeitsplätze zu verschaffen. Eine Stichprobe potenzieller Programmteilnehmer wurde beobachtet. Von diesen Personen wurden nach dem Zufallsprinzip einige für die Teilnahme an dem Programm ausgewählt. Jeder Fall entspricht einem Programmteilnehmer.
- **worldsales.sav.** Diese hypothetische Datendatei enthält Verkaufserlöse nach Kontinent und Produkt.

16 Collett, D. 2003. *Modellierung von Überlebensdaten in der medizinischen Forschung*, 2 ed. Boca Raton: Chapman & Hall/CRC.

17 Collett, D. 2003. *Modellierung von Überlebensdaten in der medizinischen Forschung*, 2 ed. Boca Raton: Chapman & Hall/CRC.

18 Verdegaal, R. 1985. *Meer sets analyse voor kwalitatieve gegevens (in Dutch)*. Leiden: Department of Data Theory, University of Leiden.

19 Ware, J. H., D. W. Dockery, A. Spiro III, F. E. Speizer, und B. G. Ferris Jr. 1984. Passive smoking, gas cooking, and respiratory health of children living in six cities. *American Review of Respiratory Diseases,* 129, 366-374.

Bemerkungen

Die vorliegenden Informationen wurden für Produkte und Services entwickelt, die auf dem deutschen Markt angeboten werden. IBM stellt dieses Material möglicherweise auch in anderen Sprachen zur Verfügung. Für den Zugriff auf das Material in einer anderen Sprache kann eine Kopie des Produkts oder der Produktversion in der jeweiligen Sprache erforderlich sein.

Möglicherweise bietet IBM die in diesem Dokument beschriebenen Produkte, Services oder Funktionen in anderen Ländern nicht an. Informationen über die gegenwärtig im jeweiligen Land verfügbaren Produkte und Services sind beim zuständigen IBM Ansprechpartner erhältlich. Hinweise auf Produkte, Programme oder Services von IBM bedeuten nicht, dass nur Produkte, Programme oder Services von IBM verwendet werden können. Anstelle der IBM Produkte, Programme oder Services können auch andere, ihnen äquivalente Produkte, Programme oder Services verwendet werden, solange diese keine gewerblichen oder anderen Schutzrechte von IBM verletzen. Die Verantwortung für den Betrieb von Produkten, Programmen und Services anderer Anbieter liegt beim Kunden.

Für in diesem Handbuch beschriebene Erzeugnisse und Verfahren kann es IBM Patente oder Patentanmeldungen geben. Mit der Auslieferung dieses Handbuchs ist keine Lizenzierung dieser Patente verbunden. Lizenzanforderungen sind schriftlich an folgende Adresse zu richten (Anfragen an diese Adresse müssen auf Englisch formuliert werden):

IBM Director of Licensing
IBM Europe, Middle East & Africa
Tour Descartes
2, avenue Gambetta
92066 Paris La Defense
France

Bei Lizenzanforderungen zu Double-Byte-Information (DBCS) wenden Sie sich bitte an die IBM Abteilung für geistiges Eigentum in Ihrem Land oder senden Sie Anfragen schriftlich an folgende Adresse:

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan

<!-- INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you. -->

Trotz sorgfältiger Bearbeitung können technische Ungenauigkeiten oder Druckfehler in dieser Veröffentlichung nicht ausgeschlossen werden. Die hier enthaltenen Informationen werden in regelmäßigen Zeitabständen aktualisiert und als Neuausgabe veröffentlicht. IBM kann ohne weitere Mitteilung jederzeit Verbesserungen und/oder Änderungen an den in dieser Veröffentlichung beschriebenen Produkten und/oder Programmen vornehmen.

Verweise in diesen Informationen auf Websites anderer Anbieter als IBM werden lediglich als Service für den Kunden bereitgestellt und stellen keinerlei Billigung des Inhalts dieser Websites dar. Das über diese Websites verfügbare Material ist nicht Bestandteil des Materials für dieses IBM Produkt. Die Verwendung dieser Websites geschieht auf eigene Verantwortung.

Werden an IBM Informationen eingesandt, können diese beliebig verwendet werden, ohne dass eine Verpflichtung gegenüber dem Einsender entsteht.

Lizenznehmer des Programms, die Informationen zu diesem Produkt wünschen mit der Zielsetzung: (i) den Austausch von Informationen zwischen unabhängig voneinander erstellten Programmen und anderen

Programmen (einschließlich des vorliegenden Programms) sowie (ii) die gemeinsame Nutzung der ausgetauschten Informationen zu ermöglichen, wenden sich an folgende Adresse:

IBM Director of Licensing
IBM Europe, Middle East & Africa
Tour Descartes
2, avenue Gambetta
92066 Paris La Defense
France

Die Bereitstellung dieser Informationen kann unter Umständen von bestimmten Bedingungen - in einigen Fällen auch von der Zahlung einer Gebühr - abhängig sein.

Die Lieferung des in diesem Dokument beschriebenen Lizenzprogramms sowie des zugehörigen Lizenzmaterials erfolgt auf der Basis der IBM Rahmenvereinbarung bzw der Allgemeinen Geschäftsbedingungen von IBM der, IBM Internationalen Nutzungsbedingungen für Programmpakete oder einer äquivalenten Vereinbarung.

Die angeführten Leistungsdaten und Kundenbeispiele dienen nur zur Illustration. Die tatsächlichen Ergebnisse beim Leistungsverhalten sind abhängig von der jeweiligen Konfiguration und den Betriebsbedingungen.

Alle Informationen zu Produkten anderer Anbieter stammen vonIBM den Anbietern der aufgeführten Produkte, deren veröffentlichten Ankündigungen oder anderen allgemein verfügbaren Quellen. IBM hat diese Produkte nicht getestet und kann daher keine Aussagen zu Leistung, Kompatibilität oder anderen Merkmalen dieser Produkte anderer Anbieter als IBM machen. Fragen zu den Leistungsmerkmalen von Produkten anderer Anbieter als IBM sind an den jeweiligen Anbieter zu richten.

Aussagen über Pläne und Absichten von IBM unterliegen Änderungen oder können zurückgenommen werden und repräsentieren nur die Ziele von IBM.

Diese Veröffentlichung enthält Beispiele für Daten und Berichte des alltäglichen Geschäftsablaufs. Sie sollen nur die Funktionen des Lizenzprogramms illustrieren und können Namen von Personen, Firmen, Marken oder Produkten enthalten. Alle diese Namen sind frei erfunden; Ähnlichkeiten mit tatsächlichen Namen und Adressen sind rein zufällig.

COPYRIGHTLIZENZ:

Diese Veröffentlichung enthält Beispielanwendungsprogramme, die in Quellensprache geschrieben sind und Programmiertechniken in verschiedenen Betriebsumgebungen veranschaulichen. Sie dürfen diese Beispielprogramme kostenlos ohne Zahlung an IBM in jeder Form kopieren, ändern und verteilen, wenn dies zu dem Zweck geschieht, Anwendungsprogramme zu entwickeln, zu verwenden, zu vermarkten oder zu verteilen, die mit der Anwendungsprogrammierschnittstelle für die Betriebsumgebung konform sind, für die diese Beispielprogramme geschrieben sind. Diese Beispiele wurden nicht unter allen denkbaren Bedingungen getestet. Daher kann IBM die Zuverlässigkeit, Wartungsfreundlichkeit oder Funktion dieser Programme weder zusagen noch gewährleisten. Die Beispielprogramme werden ohne Wartung (auf "as-is"-Basis) und ohne jegliche Gewährleistung zur Verfügung gestellt. IBM übernimmt keine Haftung für Schäden, die durch die Verwendung der Beispielprogramme entstehen.

Kopien oder Teile der Beispielprogramme bzw. daraus abgeleiteter Code müssen folgenden Copyrightvermerk beinhalten:

Marken

IBM, das IBM Logo und ibm.com sind Marken oder eingetragene Marken der IBM Corporation in den USA und/oder anderen Ländern. Weitere Produkt- und Servicenamen können Marken von IBM oder anderen Herstellern sein. Eine aktuelle Liste der IBM Marken finden Sie auf der Webseite "Copyright and trademark information" unter www.ibm.com/legal/copytrade.shtml.

Adobe, das Adobe-Logo, PostScript und das PostScript-Logo sind Marken oder eingetragene Marken der Adobe Systems Incorporated in den USA und/oder anderen Ländern.

Intel, das Intel-Logo, Intel Inside, das Intel Inside-Logo, Intel Centrino, das Intel Centrino-Logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium und Pentium sind Marken oder eingetragene Marken der Intel Corporation oder ihrer Tochtergesellschaften in den USA oder anderen Ländern.

Linux ist eine eingetragene Marke von Linus Torvalds in den USA und/oder anderen Ländern.

Microsoft, Windows, Windows NT und das Windows-Logo sind Marken der Microsoft Corporation in den USA und/oder anderen Ländern.

UNIX ist eine eingetragene Marke von The Open Group in den USA und anderen Ländern.

Java und alle auf Java basierenden Marken und Logos sind Marken oder eingetragene Marken der Oracle Corporation und/oder ihrer verbundenen Unternehmen.

Index

P

Q

S

T

U

V

W

Z