

# Datenanalyse mit SPSS für Windows

(SPSS-Kurs April 2000)

Dr. Reinhard Wittenberg

Lehrstuhl für Soziologie

Sozialwissenschaftliches Institut

Universität Erlangen-Nürnberg

Findelgasse 7-9

90402 Nürnberg

<mailto:reinhard.wittenberg@wiso.uni-erlangen.de>

Dipl. Ing. Hans Cramer

Regionales Rechenzentrum Erlangen

Universität Erlangen-Nürnberg

Martensstraße 1

91058 Erlangen

<mailto:hans.cramer@rrze.uni-erlangen.de>

# Inhaltsverzeichnis

<b>2 SPSS-Grundlagen</b>	<b>5</b>
<b>2.1 SPSS für Windows im Überblick</b>	<b>5</b>
2.1.1 SPSS starten	5
2.1.2 Fenster und Hauptmenüs	6
2.1.2.1 Daten-Editor	6
2.1.2.2 Viewer und Text-Viewer	8
2.1.2.3 Pivot-Tabellen-Editor	10
2.1.2.4 Diagramm-Editor	11
2.1.2.5 Textausgabe-Editor	13
2.1.2.6 Syntax-Editor	14
2.1.2.7 Skript-Editor	15
2.1.3 Fensterverwaltung	16
2.1.4 SPSS beenden	16
2.1.5 Dateien	17
2.1.6 Programmooptionen	17
2.1.7 Drucken	19
2.1.7.1 Viewer	19
2.1.7.2 Text-Viewer	21
2.1.7.3 Daten-Editor	21
2.1.7.4 Syntax-Editor	21
<b>2.2 Eine SPSS-Beispielsitzung</b>	<b>21</b>
2.2.1 Lesen einer SPSS-Datendatei	21
2.2.2 Aufruf einer Statistikprozedur	22
2.2.3 Die Ergebnisse im Ausgabefenster	24
2.2.4 Die Ergebnisse drucken und weiterverarbeiten	26
<b>2.3 Grundbegriffe</b>	<b>27</b>
2.3.1 Datenmatrix	27
2.3.2 Rohdaten	27
2.3.3 Fälle, Variablen und Werte	28
<b>2.4 Datendateien</b>	<b>28</b>
2.4.1 Erstmalige Datenerfassung	28
2.4.2 Rohdatendateien lesen	29
2.4.2.1 Rohdatendateien in festem Format	29
2.4.2.2 Rohdatendateien in freiem Format	29
2.4.2.3 Rohdatendateien mit Tabulator-Trennzeichen	29
2.4.3 Tabellenkalkulations- und Datenbank-Dateien lesen	30
2.4.4 SPSS-Datendateien	30
2.4.4.1 SPSS-Datendateien schreiben	30
2.4.4.2 SPSS-Datendateien lesen	31
2.4.5 Rohdatendateien schreiben	31
2.4.6 Tabellenkalkulations- und Datenbank-Dateien schreiben	31
<b>2.5 Die SPSS-Befehlssprache</b>	<b>31</b>
2.5.1 SPSS-Befehlsdateien	32
2.5.1.1 Ein neues Syntaxfenster öffnen	32

---

2.5.1.2	Eine Befehlsdatei speichern und öffnen	32
2.5.1.3	SPSS-Befehle ausführen	33
2.5.2	Befehle	33
2.5.2.1	Unterbefehle	34
2.5.2.2	Variablennamen	34
2.5.2.3	Werte	34
2.5.2.4	Schlüsselwörter	35
2.5.2.5	Zeichenketten	35
2.5.2.6	Arithmetische Operatoren	35
2.5.2.7	Trennzeichen	36
2.5.3	Variablen	36
2.5.3.1	Variablentypen	36
2.5.3.2	Variablengruppen	37
2.5.3.3	Variablenlisten	37
2.5.4	Grundlegende SPSS-Befehle	39
2.5.5	Reihenfolge der SPSS-Befehle	40
2.5.6	Hilfsbefehle	40
2.5.6.1	COMMENT	40
2.5.6.2	EXECUTE	41
<b>2.6</b>	<b>Datenbeschreibung und Datenerfassung</b>	<b>41</b>
2.6.1	Befehle	42
2.6.1.1	DATA LIST	42
2.6.1.2	MISSING VALUES	43
2.6.1.3	VARIABLE LEVEL	45
2.6.1.4	VARIABLE LABELS	45
2.6.1.5	VALUE LABELS	46
2.6.2	ALLBUS-Datensatz	47
2.6.2.1	DATA LIST	47
2.6.2.2	Fehlende Werte, Meßniveau und Etiketten	50
2.6.3	Daten-Editor	52
2.6.3.1	Datenbeschreibung in SPSS 9	52
2.6.3.2	Datenbeschreibung in SPSS 10	55
2.6.3.3	Dateneingabe	59
<b>2.7</b>	<b>Datentransformation</b>	<b>60</b>
2.7.1	RECODE	61
2.7.2	COMPUTE	64
2.7.3	IF	67
2.7.4	WEIGHT	69
2.7.5	COUNT	69
<b>2.8</b>	<b>Datenauswahl</b>	<b>70</b>
2.8.1	SELECT IF	70
2.8.2	SAMPLE	71
<b>2.9</b>	<b>SPSS-Datendateien</b>	<b>71</b>
2.9.1	SPSS-Datendateien schreiben (SAVE)	72
2.9.2	SPSS-Datendateien lesen (GET)	73

2.9.3 Variableninformationen (SYSFILE INFO, DISPLAY) .....	<a href="#">73</a>
2.9.4 Daten auflisten (LIST) .....	<a href="#">74</a>
2.9.5 SPSS-Datendateien manipulieren .....	<a href="#">75</a>
2.9.5.1 SPSS-Datendateien sortieren .....	<a href="#">75</a>
2.9.5.2 SPSS-Datendateien aufteilen .....	<a href="#">75</a>
2.9.6 Portable SPSS-Datendateien .....	<a href="#">75</a>
2.9.6.1 Portable SPSS-Datendateien schreiben (EXPORT) .....	<a href="#">75</a>
2.9.6.2 Portable SPSS-Datendateien lesen (IMPORT) .....	<a href="#">76</a>

## 2 SPSS-Grundlagen

SPSS für Windows wird über eine grafische Benutzeroberfläche bedient. Um möglichst schnell mit dieser vertraut zu werden, beschreiben wir als erstes den Start des Programmsystems, seine verschiedenen Fenster mit ihren Menüs und das Beenden einer SPSS-Sitzung. Die unter *Programmoptionen* vorgeschlagenen Programmeinstellungen sollten möglichst vor der ersten Datenanalyse durchgeführt werden.

Danach zeigen wir mehrere Möglichkeiten auf, bestehende Datendateien in SPSS zu übernehmen und geben Ratschläge zur erstmaligen Datenerfassung. Es folgt die ausführliche Beschreibung der SPSS-Befehlssprache sowie die Vorstellung der grundlegenden SPSS-Befehle.

### 2.1 SPSS für Windows im Überblick

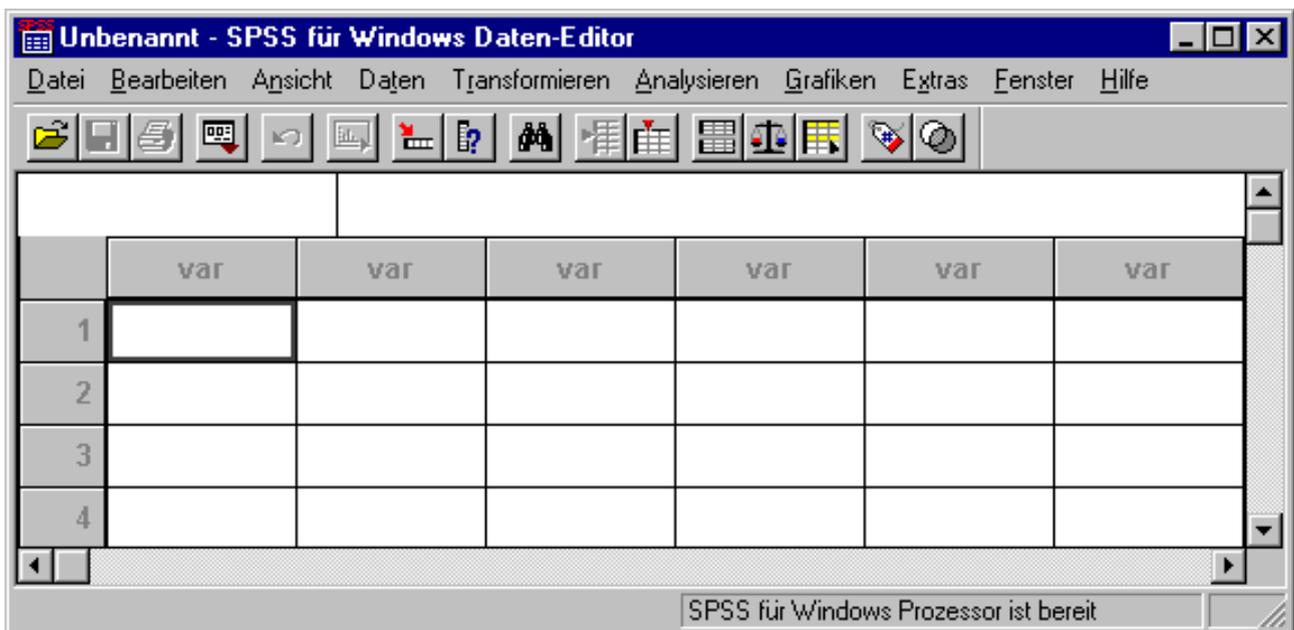
Die grafische Benutzeroberfläche von SPSS für Windows erleichtert die Datenanalyse durch selbst-erklärende Menüs und Symbole sowie übersichtlich gestaltete Dialogfelder. Ein grundlegendes Lernprogramm, der Statistik-Assistent und ausführliche Online-Hilfen unterstützen den Anfänger beim Einarbeiten und bieten dem Fortgeschrittenen ein Nachschlagewerk. Nach einer sorgfältigen Datenbeschreibung und Erstellung einer SPSS-Datendatei können viele Aufgaben einfach mit der Maus erledigt werden.

#### 2.1.1 SPSS starten

SPSS - hier Version 9 - wird über das Startmenü von Windows 95, Windows 98 und Windows NT wie folgt gestartet:

►Start ►Programme ►SPSS 9.0 für Windows

Nach dem Start öffnet SPSS standardmäßig ein Fenster mit dem Daten-Editor. In der Fensterleiste wird der Fenstertyp mitgeteilt (Unbenannt - SPSS für Windows Daten Editor); es folgt das Daten-Editor-Menü, bestehend aus einer Menü- und Symbolleiste, und am unteren Fensterrand die SPSS-Statusleiste (SPSS für Windows Prozessor ist bereit). Im Fenster selbst wird der Inhalt der - anfangs leeren und unbenannten - SPSS-Datendatei angezeigt:



### 2.1.2 Fenster und Hauptmenüs

In SPSS werden verschiedene Typen von Fenstern eingesetzt. Die Bedienung der Fenster bezüglich Windows - also das Verkleinern, Vergrößern, Verschieben und Schließen der Fenster über die Windows-Schalt symbole im rechten, oberen Fensterrahmen - ist einheitlich, Menü- und Symbolleiste sind entsprechend den fensterspezifischen Funktionen unterschiedlich. Einige Menüs sind in allen Fenstern verfügbar, z.B.:

►Datei, ►Bearbeiten und ►Hilfe .

Einige Funktionen können statt über die Menüs direkt durch Anklicken des entsprechenden Symbols in der Symbolleiste aufgerufen werden. Bei Berührung des Symbols mit dem Mauszeiger wird eine Beschreibung der damit verbundenen Funktion ausgegeben. Nur die voll ausgeleuchteten Symbole sind aktiviert. Die Funktionen der Hauptmenüs sind über die darunterliegenden Untermenüs erreichbar.

**Hilfe** bieten alle Hauptmenüs und viele Untermenüs. Zusätzlich verfügen die meisten Dialogfelder über eine Schaltfläche ►Hilfe. Auch durch Anklicken einer Schaltfläche mit der rechten Maustaste kann über das Pop-up-Kontextmenü Hilfe (►Direkthilfe) angefordert werden.

Jedes Fenster enthält, sofern nicht über ►Ansicht ►Statusleiste ausgeblendet, die **SPSS-Statusleiste** mit folgenden Informationen:

1. Befehlsstatus: Anzeige des Befehls, der gerade von SPSS ausgeführt wird bzw. die Anzahl der bereits abgearbeiteten Fälle. Nach Beendigung eines SPSS-Programms erscheint hier „SPSS für Windows Prozessor ist bereit“. Erst jetzt ist die Ausgabe komplett und erst jetzt können neue Befehle zur Bearbeitung an SPSS übergeben werden.
2. Filterstatus: Die Meldung „Filter“ weist daraufhin, daß nicht alle Fälle in der aktuellen Analyse verarbeitet werden. Es wurde durch eine Datenauswahl eine Teilmenge herausgefiltert.
3. Gewichtungstatus: Die Meldung „Gewichtung“ besagt, daß die Fälle der Analyse unterschiedlich gewichtet wurden.
4. Aufspaltungsstatus: Bei der Meldung „Datei aufteilen an“ ist die Datendatei in verschiedene Gruppen (z.B. Frauen und Männer) aufgeteilt.

Die Fensterinhalte können, bis auf die des Pivot-Tabellen- und Textausgabe-Editors, getrennt in einzelnen Dateien gespeichert und zur erneuten Benutzung geöffnet werden. SPSS verwendet die folgenden 8 Fenstertypen, die zugehörigen SPSS-Dateitypen sind in eckigen Klammern angegeben:

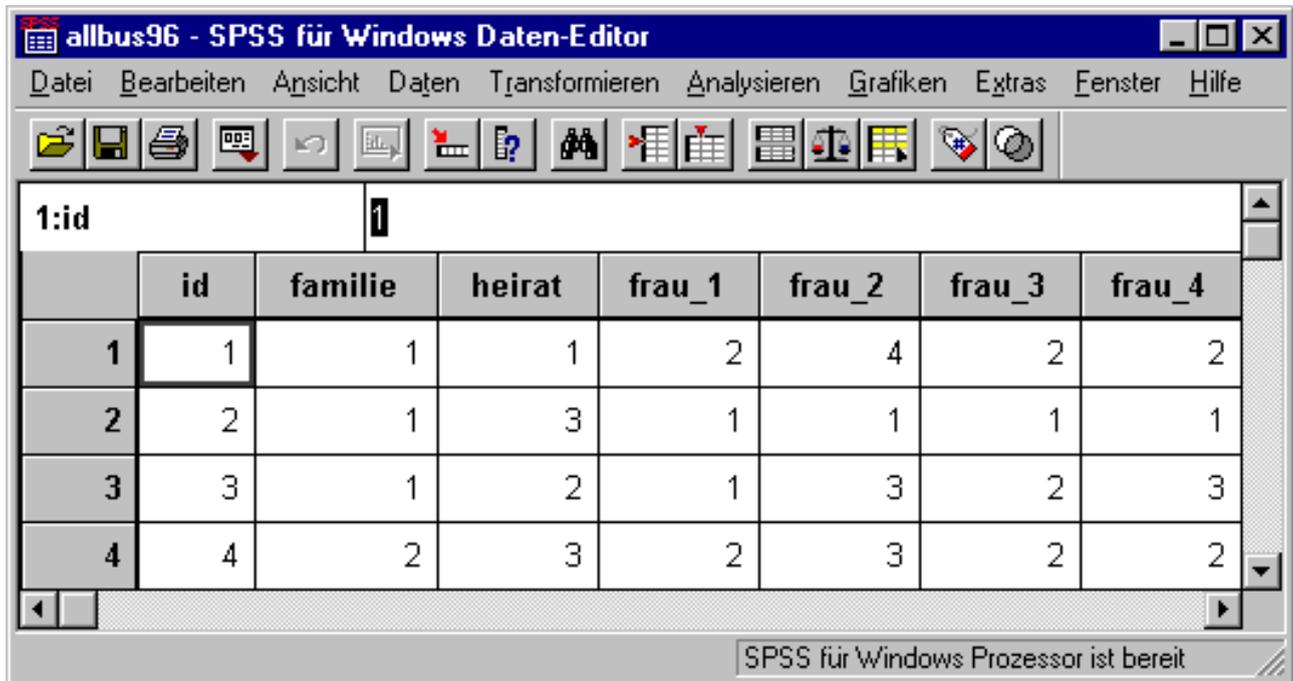
5. Daten-Editor-Fenster [`*.sav`]
6. Ausgabe-Fenster (alternativ)
  - a. Viewer-Fenster [`*.spo`]
  - b. Text-Viewer-Fenster [`*.rtf`]
7. Pivot-Tabellen-Editor-Fenster (im Viewer-Fenster enthalten)
8. Diagramm-Editor-Fenster [`*.sct`]
9. Textausgabe-Editor-Fenster (im Viewer-Fenster enthalten)
10. Syntax-Editor-Fenster [`*.sps`]
11. Skript-Editor-Fenster [`*.sbs`]

Statt der langen Fensteramen kann auch deren Kurzform, wie z.B. Datenfenster statt Daten-Editor-Fenster, benutzt werden.

#### 2.1.2.1 Daten-Editor

Im Daten-Editor-Fenster wird der Inhalt der Datendatei angezeigt. Mit Hilfe des Daten-Editors kön-

neue neue Datendateien erstellt und bestehende bearbeitet werden. Es kann nur eine Datendatei und damit auch nur ein Daten-Editor-Fenster geöffnet sein.



## Hauptmenüs

- ▶ Datei      Dateien öffnen, lesen, speichern, drucken; Rohdaten-, Tabellenkalkulations- (z.B. Excel-Dateien) und Datenbank-Dateien (z.B. dBASE- und Access-Dateien) lesen und speichern; SPSS beenden
- ▶ Bearbeiten      Datenwerte markieren, ausschneiden, kopieren, einfügen, löschen, suchen, ersetzen; Programmoptionen
- ▶ Ansicht      Symbol- und Statusleiste gestalten; Schriftarten und Anzeige der Datenwerte, Wertetiketten und Gitterlinien ändern
- ▶ Daten      Variablendefinition, Datenauswahl/-gewichtung, Datendateifunktionen
- ▶ Transformieren  
                  Daten ändern und neu berechnen, Zeitreihen erstellen, fehlende Werte ersetzen
- ▶ Analysieren      Statistikprozedur-Auswahl<sup>1</sup>
- ▶ Grafiken      Grafikdiagramm-Auswahl
- ▶ Extras      Variableninformationen
- ▶ Fenster      Fensterwechsel und SPSS-Minimierung
- ▶ Hilfe      Online-Dokumentation, Lernprogramm, Statistikassistent, Informationen zur SPSS-Programmversion

SPSS 10 bietet zusätzlich zu der in den vorhergehenden Versionen üblichen, tabellarischen **Daten-**

<sup>1</sup> So werden in SPSS die Befehle zur Durchführung statistischer Berechnungen genannt.

**ansicht** eine **Variablenansicht** zur übersichtlichen Darstellung der Variablenbeschreibung:

	Name	Typ	Spaltenformat	Dezimalstellen	Variablenlabel	Wertelabels	Fehlende Wert
1	id	Numerisch	4	0	Identifikationsn	Kein	Kein
2	familie	Numerisch	1	0	Braucht man F	{1, man brauch	9
3	heirat	Numerisch	1	0	Heirat bei daue	{1, ja}...	8, 9
4	frau_1	Numerisch	1	0	Berufst. Frau:	{1, stimme voll	8, 9
5	frau_2	Numerisch	1	0	Frau: Lieber M	{1, stimme voll	8, 9
6	frau_3	Numerisch	1	0	Frau: Nicht arb	{1, stimme voll	8, 9
7	frau_4	Numerisch	1	0	Frau: Zuhause	{1, stimme voll	8, 9
8	frau_5	Numerisch	1	0	Frau: Bessere	{1, stimme voll	8, 9
9	frau_6	Numerisch	1	0	Frau: Nach Hei	{1, stimme voll	8, 9
10	anomia_1	Numerisch	1	0	Lageverschlec	{1, bin derselb	8, 9
11	anomia_2	Numerisch	1	0	Bei dieser Zuk	{1, bin derselb	8, 9
12	anomia_3	Numerisch	1	0	Politiker uninte	{1, bin derselb	8, 9
13	anomia_4	Numerisch	1	0	Mehrheit unint	{1, bin derselb	8, 9

Die Variablenansicht zeigt die Namen und Attribute aller Variablen. Die Nutzung der Variablenansicht und die Variablenattribute sind in [Kapitel 2.6.3](#) beschrieben.

### 2.1.2.2 Viewer und Text-Viewer

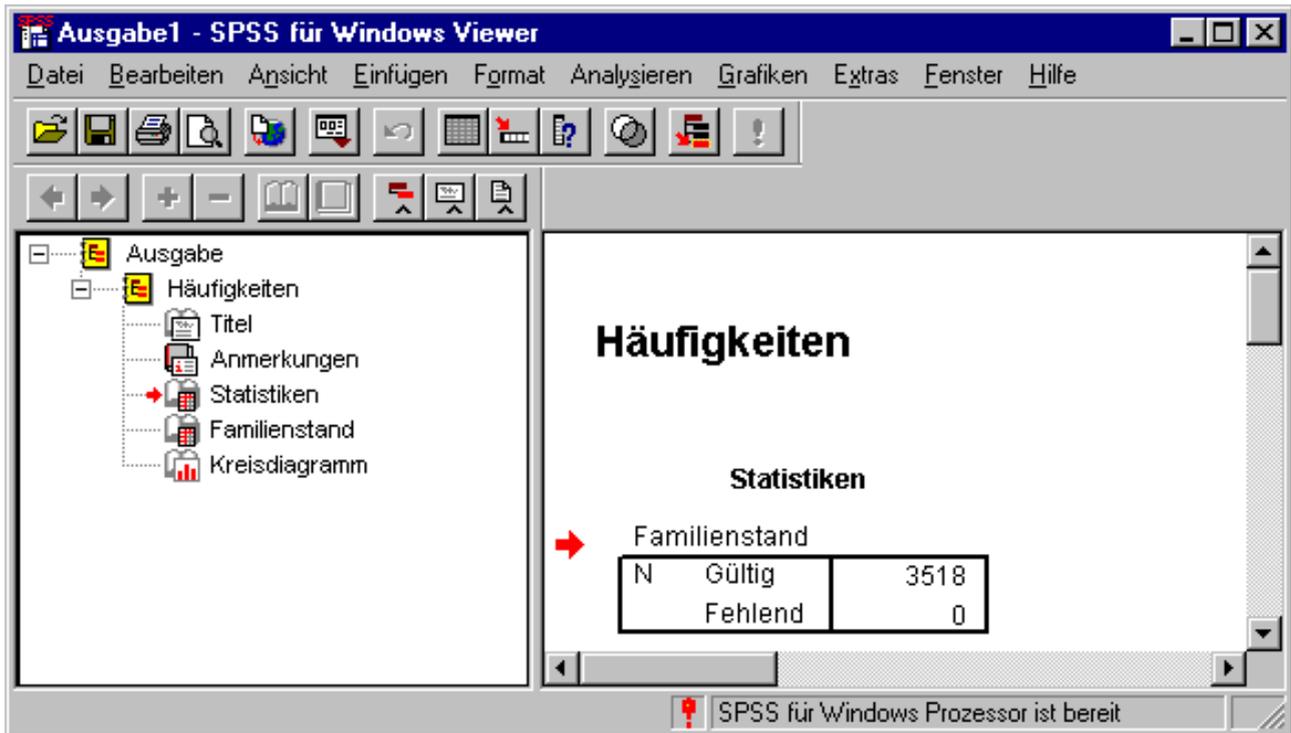
Die statistischen Ergebnisse, Tabellen und Diagramme werden im Viewer-Fenster oder Text-Viewer-Fenster angezeigt und können bequem mit dem Viewer bzw. Text-Viewer gelesen, gespeichert und gedruckt werden. Es können mehrere Ausgabefenster (jedoch nur eines Viewer-Typs) gleichzeitig geöffnet sein. Der Viewer-Typ wird ausgewählt über:

► Bearbeiten ► Optionen ► Allgemein ► Ausgabetyt beim Starten

#### Viewer

Im Viewer-Fenster werden Tabellen als interaktive Pivot-Tabellen und Grafiken als interaktive Diagramme dargestellt, die nachträglich bearbeitet und als interaktive Objekte in andere Windows-Anwendungen übernommen werden können.

Der linke Flügel des Viewer-Fensters enthält das gegliederte Inhaltsverzeichnis der Ausgabe. Die Ausgabe selbst wird im rechten Fensterflügel angezeigt. Das Navigieren durch die Ausgabe erfolgt durch Anklicken von Ausgabeelementen im Gliederungsfenster.



## Hauptmenüs

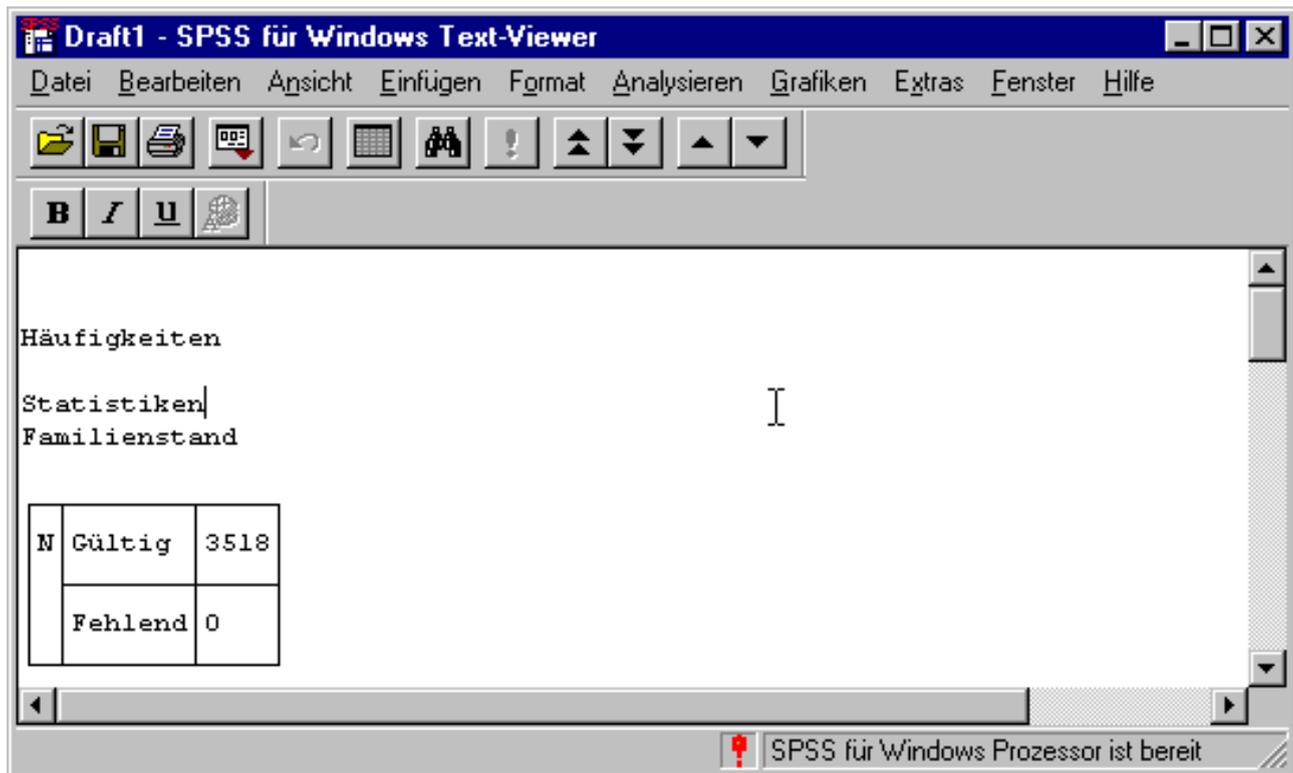
- ▶Datei        Dateien öffnen, lesen, speichern, drucken und in Fremdformaten, z.B. als Text- oder HTML-Datei<sup>1</sup>, exportieren
- ▶Bearbeiten    Ausgabegliederung im Gliederungsfenster (linker Fensterflügel) ändern; Ausgaben markieren, ausschneiden, kopieren, einfügen, löschen, verschieben; Programmoptionen
- ▶Ansicht        Symbol- und Statusleiste gestalten, Gliederungsansicht im Gliederungsfenster ändern
- ▶Einfügen      Ausgabegestaltung: Seitenumbrüche, Titel usw.
- ▶Format        Ausgabegestaltung: Tabelleninhalte ausrichten
- ▶Analysieren    Statistikprozedur-Auswahl
- ▶Grafiken        Grafikdiagramm-Auswahl
- ▶Extras         Variableninformationen, Hauptfensterauswahl
- ▶Fenster        Fensterwechsel und SPSS-Minimierung
- ▶Hilfe         Online-Dokumentation, Lernprogramm, Statistikassistent, Informationen zur SPSS-Programmversion

## Text-Viewer

Der Text-Viewer stellt Ergebnisse als einfachen Text dar. Die Textausgabe kann bearbeitet und in andere Anwendungen eingefügt werden. Diagramme können nachträglich nicht bearbeitet werden, lediglich ihre Größe ist veränderbar. Eine Übernahme in andere Anwendungen ist ebenfalls möglich.

Das Ausgabefenster ist nicht über ein Inhaltsverzeichnis gegliedert, das Navigieren erfolgt durch Blättern und Suchen im Text-Viewer-Fenster.

<sup>1</sup> HyperText Markup Language ist die Beschreibungssprache für World-Wide-Web-Dokumente (WWW).



### Hauptmenüs

- ▶Datei        Dateien öffnen, lesen, speichern, drucken und in Fremdformaten, z.B. als Textdatei exportieren
- ▶Bearbeiten    Ausgaben markieren, ausschneiden, kopieren, einfügen, löschen, verschieben; Programmooptionen
- ▶Ansicht        Symbol- und Statusleiste gestalten
- ▶Einfügen      Ausgabegestaltung: Seitenumbrüche einfügen
- ▶Format        Ausgabegestaltung: Schriftart ändern
- ▶Analysieren    Statistikprozedur-Auswahl
- ▶Grafiken        Grafikdiagramm-Auswahl
- ▶Extras        Variableninformationen, Hauptfensterauswahl
- ▶Fenster        Fensterwechsel und SPSS-Minimierung
- ▶Hilfe         Online-Dokumentation, Lernprogramm, Statistikassistent, Informationen zur SPSS-Programmversion

#### 2.1.2.3 Pivot-Tabellen-Editor<sup>1</sup>

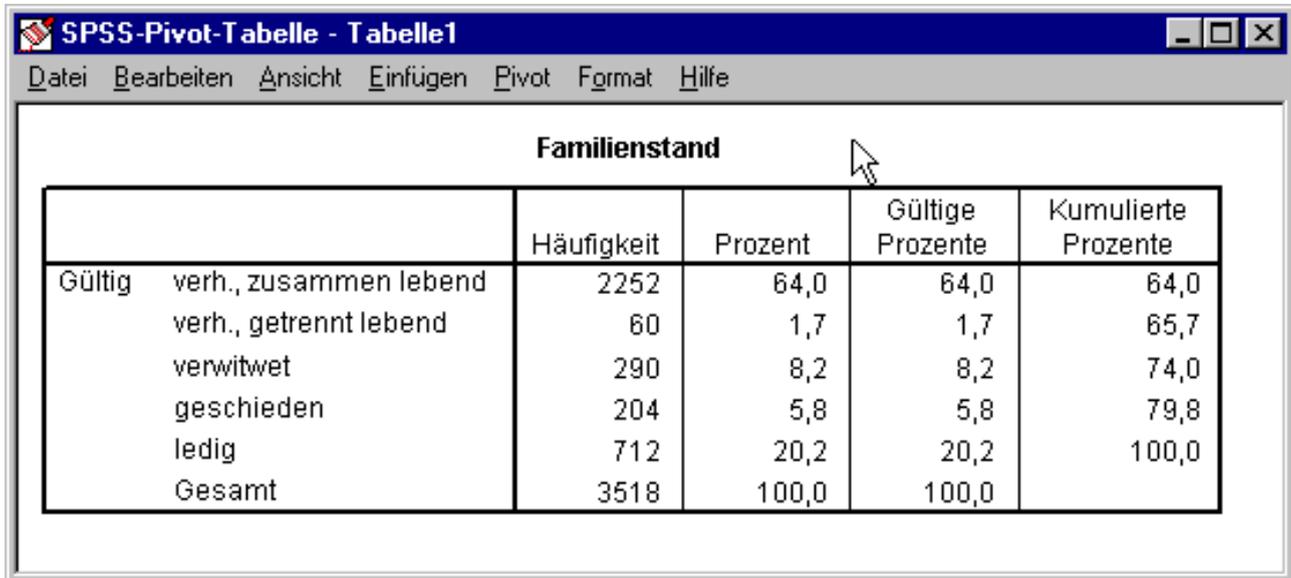
Ausgaben, die in Form von Pivot-Tabellen im Viewer-Fenster angezeigt werden, können mit dem Pivot-Tabellen-Editor bearbeitet werden, d.h. Zeilen, Spalten und Schichten dieser Tabellen können interaktiv umsortiert werden.

Es können mehrere Pivot-Tabellen gleichzeitig bearbeitet werden. Die Bearbeitung erfolgt über das

<sup>1</sup> Franz. „pivot“: Dreh-, Angelpunkt

kontextsensitive Menü der rechten Maustaste:

- ▶Klick mit rechter Maustaste auf die gewünschte Pivot-Tabelle
- ▶Objekt: SPSS Pivot-Tabelle
- ▶Öffnen



The screenshot shows a window titled "SPSS-Pivot-Tabelle - Tabelle1" with a menu bar containing "Datei", "Bearbeiten", "Ansicht", "Einfügen", "Pivot", "Format", and "Hilfe". The main content area displays a table titled "Familienstand". The table has five columns: "Gültig", "Häufigkeit", "Prozent", "Gültige Prozente", and "Kumulierte Prozente". The rows represent different marital statuses.

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	verh., zusammen lebend	2252	64,0	64,0	64,0
	verh., getrennt lebend	60	1,7	1,7	65,7
	verwitwet	290	8,2	8,2	74,0
	geschieden	204	5,8	5,8	79,8
	ledig	712	20,2	20,2	100,0
	Gesamt	3518	100,0	100,0	

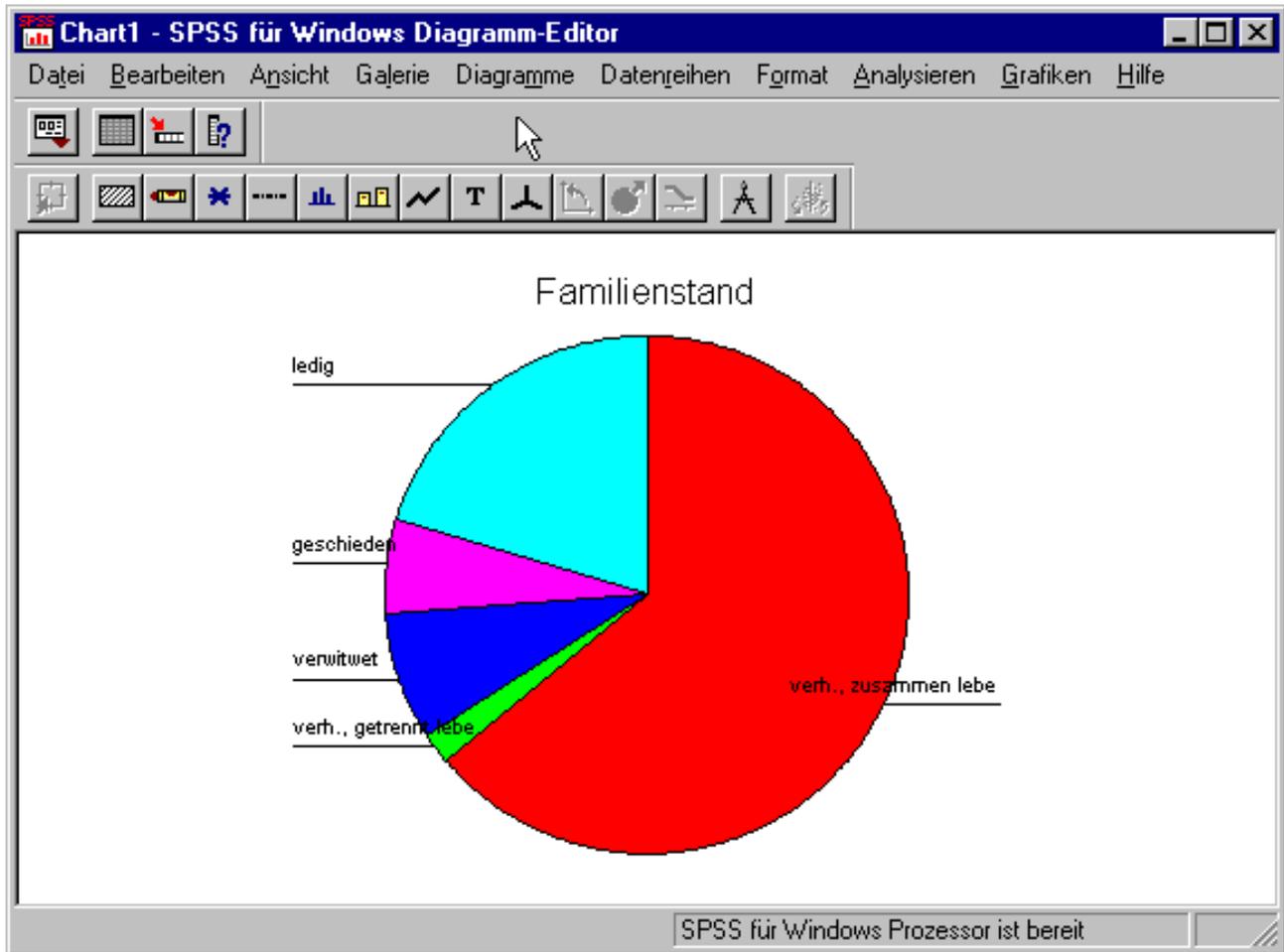
### Hauptmenüs

- ▶Datei        Dateien öffnen, lesen, speichern, drucken
- ▶Bearbeiten    Datenwerte in Pivot-Tabellen markieren, ausschneiden, kopieren, einfügen, löschen; Pivot-Aktionen wiederholen und zurücknehmen
- ▶Ansicht        Symbol- und Statusleiste gestalten, Elemente und Gitterlinien der Pivot-Tabellen ein-/ausblenden
- ▶Einfügen      Pivot-Tabellengestaltung: Titel, Fußnoten usw.
- ▶Pivot          Pivot-Aufgaben ausführen, Pivot-Leisten ein-/ausblenden, mehrdimensionale Pivot-Tabellen bearbeiten
- ▶Format        Pivot-Tabellengestaltung: Tabellen- und Zellenformat, Schriftmerkmale usw.
- ▶Hilfe          Online-Dokumentation, Informationen zum Pivot-Tabellen-Editor

#### 2.1.2.4 Diagramm-Editor

Die im Viewer-Fenster (nicht Text-Viewer) ausgegebenen Diagramme und Grafiken können mit dem Diagramm-Editor nachbearbeitet werden. Ein Diagramm-Editor-Fenster wird wie folgt geöffnet:

- ▶Klick mit rechter Maustaste auf das gewünschte Diagramm
- ▶Objekt: SPSS Diagramm
- ▶Öffnen



### Hauptmenüs

- ▶Datei Diagrammvorlagen speichern; Diagramme in Fremdfomaten wie z.B. BMP<sup>1</sup>, JPEG<sup>2</sup>, TIFF<sup>3</sup> und WMF<sup>4</sup> exportieren
- ▶Bearbeiten Diagramme kopieren; Programmooptionen
- ▶Ansicht Symbol- und Statusleiste gestalten
- ▶Galerie Diagrammtypänderung: z.B. Balkendiagramm in Kreisdiagramm umwandeln
- ▶Diagramme Diagrammbearbeitung: Layout, Beschriftung, Rahmen
- ▶Datenreihen Datenauswahl für Diagramme, Daten transponieren
- ▶Format Ausgabegestaltung: Füllmuster, Farben, Linienstil, Balkenarten, Beschriftungen, Schriften, Achsenvertauschung, 3D-Rotation, Darstellung der fehlenden Werte
- ▶Analysieren Statistikprozedur-Auswahl
- ▶Grafiken Grafikdiagramm-Auswahl

<sup>1</sup> BMP: Windows Bitmap

<sup>2</sup> JPEG: Grafik-Kompressionsverfahren der Joint Photographic Experts Group

<sup>3</sup> TIFF: Tagged Image File Format

<sup>4</sup> WMF: Windows Metafile

- ▶ Hilfe      Online-Dokumentation, Lernprogramm, Statistikassistent, Informationen zur SPSS-Programmversion

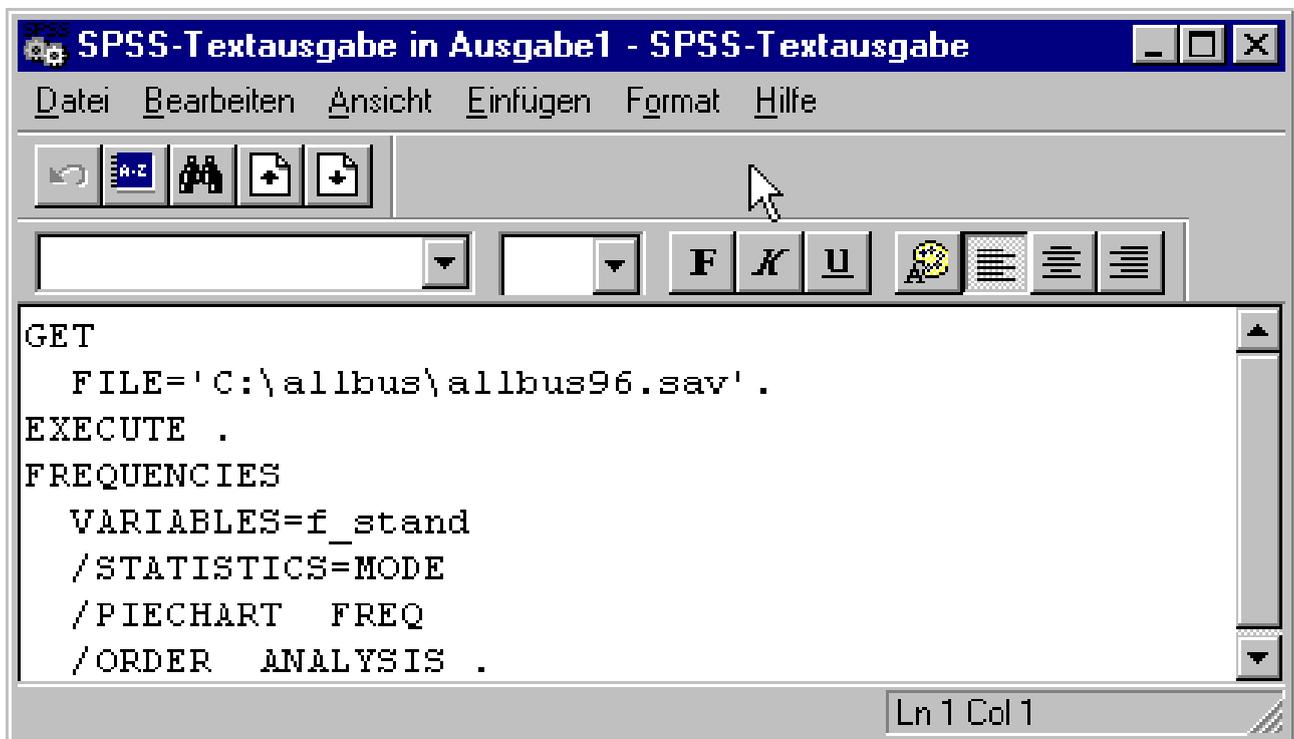
Für die Übernahme in Textverarbeitungssysteme stehen die gängigen Grafikdateiformate zur Verfügung.

### 2.1.2.5 Textausgabe-Editor

Im Viewer-Fenster können Textausgaben außerhalb von Pivot-Tabellen mit dem Textausgabe-Editor nachbearbeitet und gespeichert werden.

Die Bearbeitung einer Textausgabe erfolgt über das kontextsensitive Menü der rechten Maustaste:

- ▶ Klick mit rechter Maustaste auf das gewünschte Textobjekt
- ▶ Objekt: SPSS Rtf Document Object
- ▶ Open



### Hauptmenüs

- ▶ Datei      Dateien öffnen, lesen, speichern, drucken
- ▶ Bearbeiten      Ausgabetexte markieren, ausschneiden, kopieren, einfügen, löschen, suchen und ersetzen; Farbeinstellungen ändern
- ▶ Ansicht      Symbol- und Statusleiste gestalten
- ▶ Einfügen      Ausgabegestaltung: Seitenumbruch
- ▶ Format      Ausgabegestaltung: Schriften, Ausrichtung usw.
- ▶ Hilfe      Online-Dokumentation, Lernprogramm, Informationen zum Text-Editor

### 2.1.2.6 Syntax-Editor

Mit dem Syntax-Editor werden SPSS-Befehle oder komplette SPSS-Programme im Syntaxfenster erstellt und ausgeführt. Die SPSS-Programme können gespeichert, erneut geladen, bearbeitet und ausgeführt werden.

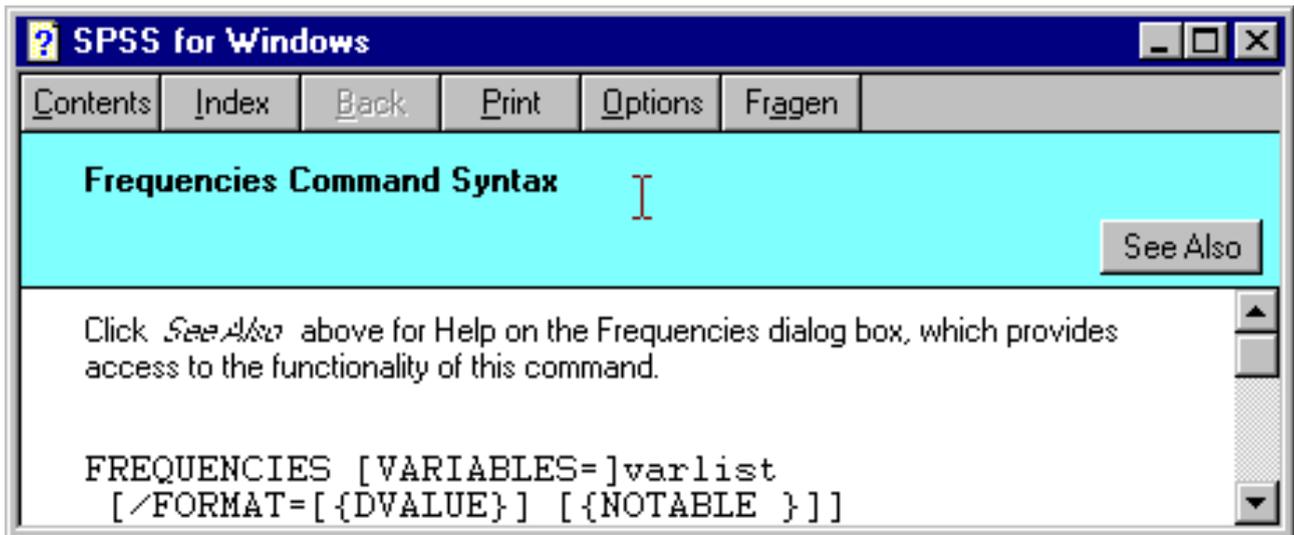


#### Hauptmenüs

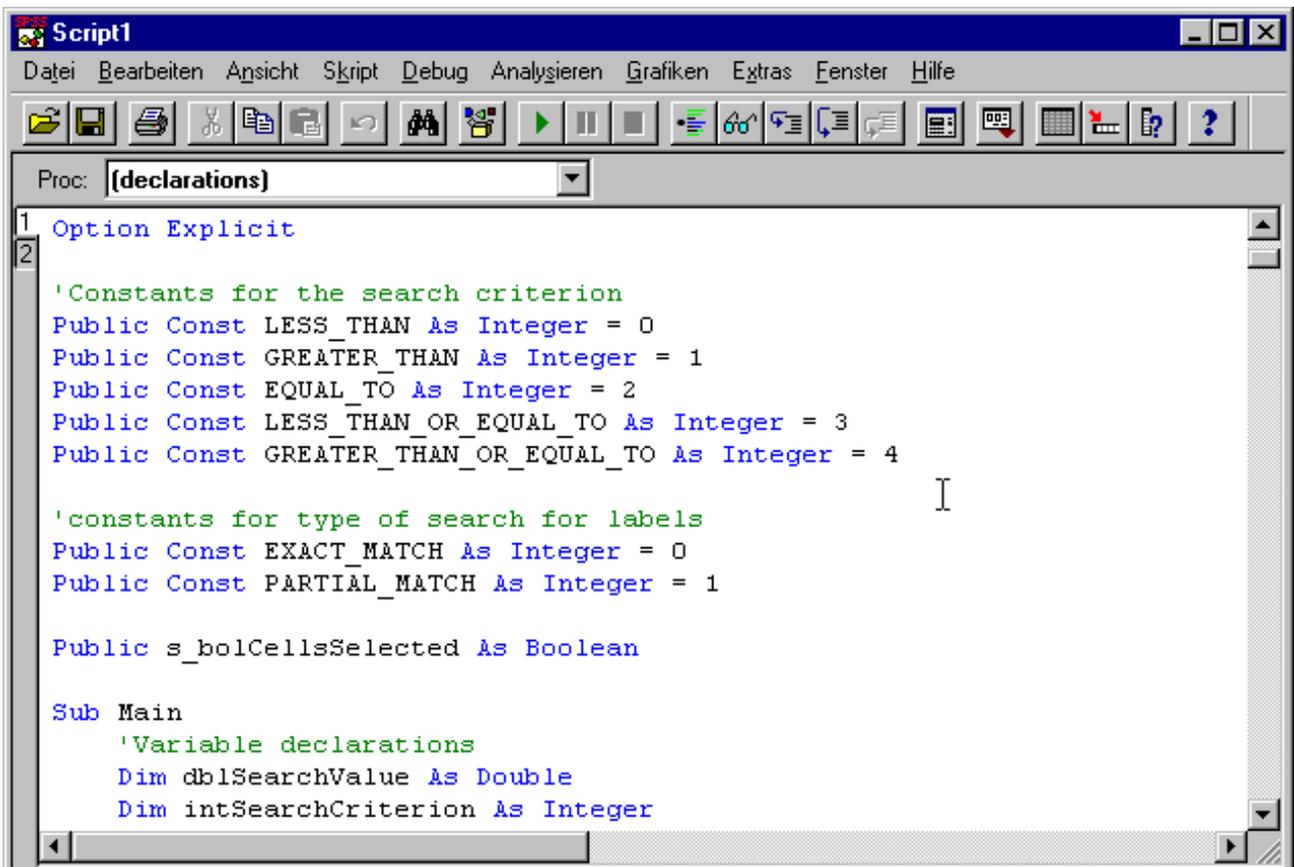
- ▶ Datei        Dateien öffnen, lesen, speichern, drucken
- ▶ Bearbeiten   Text markieren, ausschneiden, kopieren, einfügen, löschen, suchen und ersetzen; Programmooptionen
- ▶ Ansicht      Symbol- und Statusleiste gestalten, Schriftart ändern
- ▶ Analysieren   Statistikprozedur-Auswahl
- ▶ Grafiken     Grafikdiagramm-Auswahl
- ▶ Extras        Variableninformationen, Hauptfensterauswahl
- ▶ Ausführen    SPSS-Befehle ausführen
- ▶ Fenster      Fensterwechsel und SPSS-Minimierung
- ▶ Hilfe         Online-Dokumentation, Lernprogramm, Statistikassistent, Informationen zur SPSS-Programmversion

**Syntaxbeschreibung:** ▶ 

Die Syntaxdiagramme, das sind die Beschreibungen der SPSS-Befehle, können über das Hilfe-Menü oder im Syntax-Editor-Fenster über obiges Symbol nachgeschlagen werden. Sie werden kontextabhängig angezeigt, d.h. es wird das Diagramm des Befehls angezeigt, auf dem der Mauszeiger sich gerade im Syntax-Editor-Fenster befindet. Im obigen Beispiel ist das die Statistikprozedur *FREQUENCIES*:



### 2.1.2.7 Skript-Editor



Skripts dienen der Automatisierung von SPSS-Programmabläufen wie z.B. dem Öffnen und Speichern von Datendateien und der Anpassung der Ausgabe im Ausgabefenster. Skripts werden mit dem Skript-Editor in der speziellen Programmiersprache *Sax BASIC* erstellt. Da das Arbeiten mit Skripts sehr gute Kenntnisse der internen Funktionsweise von SPSS voraussetzt, wird in diesem Buch nicht näher darauf eingegangen. Erfahrenen SPSS-Benutzern, die häufiger umfangreiche Analysen mit SPSS durchführen, können Skripts jedoch zeitraubende Aufgaben abnehmen.

### Hauptmenüs

- ▶Datei           Dateien öffnen, lesen, speichern, drucken
- ▶Bearbeiten    Text markieren, ausschneiden, kopieren, einfügen, löschen, suchen und ersetzen; Programmooptionen
- ▶Ansicht        Symbol- und Statusleiste gestalten
- ▶Skript         Skripts erstellen, bearbeiten und ausführen
- ▶Debug         Skripts testen
- ▶Analysieren   Statistikprozedur-Auswahl
- ▶Grafiken       Grafikdiagramm-Auswahl
- ▶Extras         Variableninformationen
- ▶Fenster        Fensterwechsel und SPSS-Minimierung
- ▶Hilfe         Online-Dokumentation zur SPSS-Skriptsprache, Informationen zur SPSS-Programmversion

### 2.1.3 Fensterverwaltung

Beim Aufruf von SPSS wird standardmäßig das leere Daten-Editor- Fenster geöffnet. Andere Fenster wie z.B. ein Syntax-Editor-Fenster können vom Benutzer über

▶Datei ▶Neu ▶Syntax bzw. ▶Datei ▶Öffnen

geöffnet werden, ebenso erzeugt SPSS bei Bedarf neue Fenster( z.B. Ausgabe- und Diagramm-Editor-Fenster).

Der Inhalt der Fenster kann über das Menü ▶Datei in Dateien gespeichert und in folgenden SPSS-Sitzungen über ▶Datei ▶Öffnen wieder in das entsprechende Fenster geladen werden ( *Datendateien*).

Dem Anfänger raten wir, damit er den „Durchblick“ behält, nicht zu viele Fenster gleichzeitig zu öffnen und nicht mehr benötigte zu schließen. Bei gleichzeitigem Arbeiten mit mehreren Windows-Programmen ist die Funktion „SPSS minimieren“ in den Hauptmenüs „Fenster“ sehr hilfreich.. Mit einem einzigen Mausklick verschwinden alle SPSS-Fenster von der „Bildfläche“, um dann genauso einfach wieder geöffnet zu werden.

#### Aktives Fenster: ▶Fenster

Bei mehreren geöffneten Fenstern ist das Fenster, das sich auf dem Bildschirm im Vordergrund befindet, das **aktive Fenster**. Ein Fenster wird durch Anklicken im Menü ▶Fenster zum aktiven Fenster. Die über die Menü- und Symbolleiste angebotenen Funktionen werden auf den Inhalt des aktiven Fensters angewendet.

**Hauptfenster:** ▶ 

Von einigen Fenstertypen, z.B. Syntax-Editor und Viewer, können gleichzeitig mehrere Fenster geöffnet sein. Von diesen kann jedoch nur eins das **Hauptfenster** sein. SPSS verwendet das Hauptfenster als Ein- bzw. Ausgabefenster. Ein Hauptfenster wird über das Menü mit ▶Extras ▶Hauptfenster oder durch Anklicken des oben abgebildeten Ausrufezeichens in der Symbolleiste bestimmt, das daraufhin in der Statuszeile des Fensters erscheint.

### 2.1.4 SPSS beenden



SPSS wird mit dem Windows-Schaltersymbol  oder über die Menüfolge ▶Datei ▶Beenden beendet. Falls nicht alle Dateien (=Fensterinhalte), die während der SPSS-Sitzung erzeugt wurden, bereits gespeichert sind, fordert SPSS nun dazu auf.

### 2.1.5 Dateien

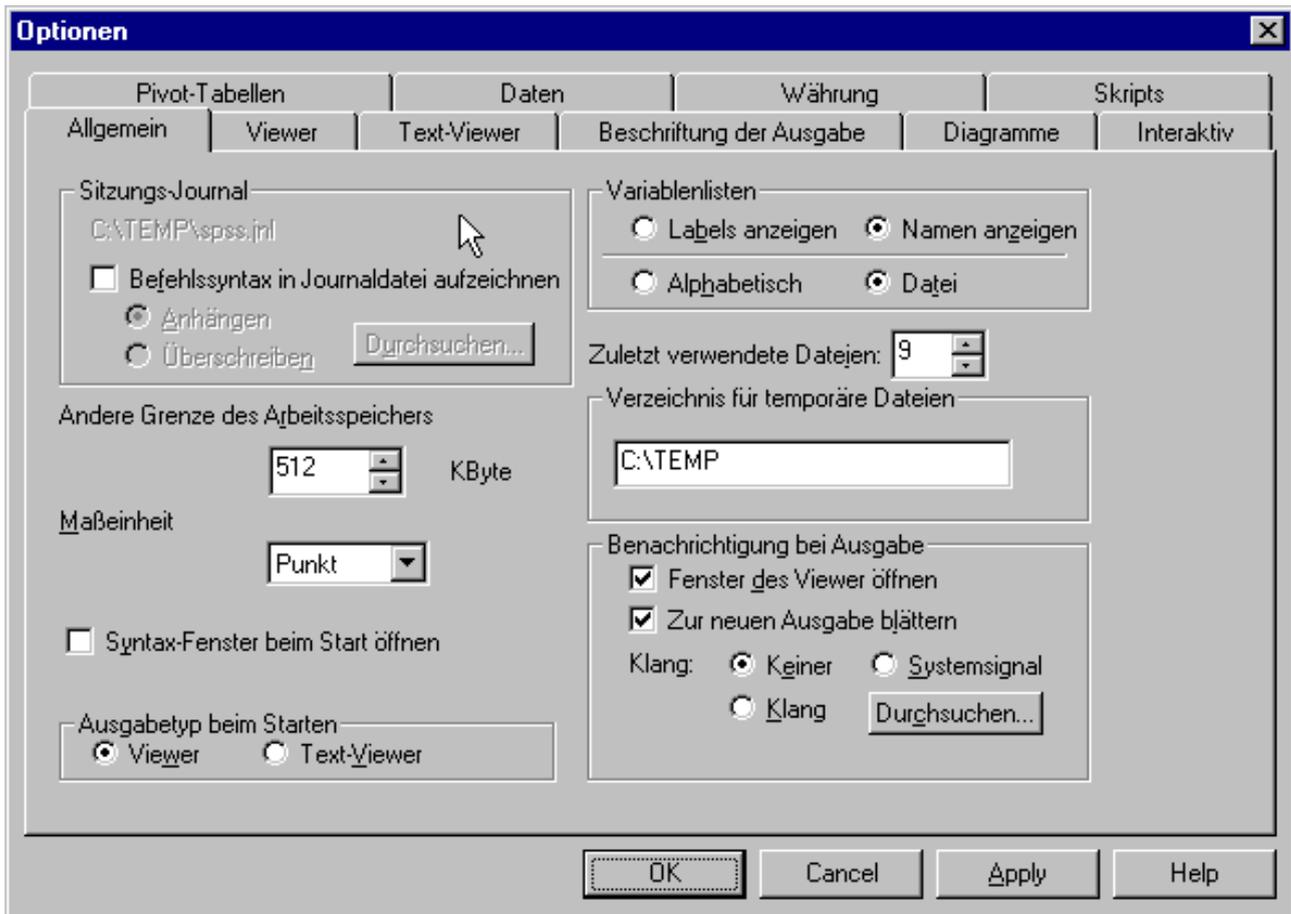
SPSS liest und schreibt Dateien unterschiedlicher Art, von denen wir schon einige kennengelernt haben. Die wichtigsten sind:

Bezeichnung	Inhalt bzw. Fenster	Typ	Format
Rohdatendatei	Daten	*.dat	Text
Befehlsdatei	Syntax-Editor	*.sps	Text
Arbeitsdatei	Daten, Variablen	intern	binär
SPSS-Datendatei	Daten-Editor	*.sav	binär
Ausgabedatei	Viewer	*.spo	Text
Ausgabedatei	Textausgabe-Editor	*.txt *.htm	Text
Ausgabedatei	Pivot-Tabellen-Editor	*.txt *.htm	Text
Ausgabedatei	Text-Viewer	*.rtf	Text
Grafikdatei	Diagramm-Editor	*.jpg *.tif *.wmf	binär
Portable Datei	Daten, Variablen	*.por	Text

Die Dateinamen (\*) sind frei wählbar, nach dem Punkt muß der Dateityp angegeben werden. Textdateien sind plattformunabhängig, Binärdateien haben ein rechner- und betriebssystemspezifisches Format.

### 2.1.6 Programmooptionen

Das Programmsystem SPSS sollte an die individuelle Arbeitsumgebung angepaßt werden. Eine Vielzahl dieser permanenten, d.h. auch für folgende Sitzungen noch wirksamen Programmeinstellungen, kann in fast allen Fenstern über das Dialogfeld ▶Bearbeiten ▶Optionen geändert werden. Folgende Einstellungen, die erst nach einem Neustart von SPSS wirksam sind, sollten in der ersten SPSS-Sitzung vorgenommen werden:

**Registerkarte „Allgemein“**

## 1. Sitzungs-Journal:

- Befehlssyntax in Journaldatei aufzeichnen

Durch Entfernen des Häkchens per Mausklick wird die Aufzeichnung abgeschaltet. Erfahrene Benutzer können sie wieder aktivieren und die in der Journaldatei gesammelten SPSS-Befehle im Syntaxfenster verwenden.

## 2. Variablenlisten:

- a.  Namen anzeigen
- b.  Datei (d.h. Reihenfolge der Variablen wie im Fragebogen)

3. Ausgabebetyp beim Starten:  Viewer

(wird in diesem Buch für alle Beispiele verwendet)

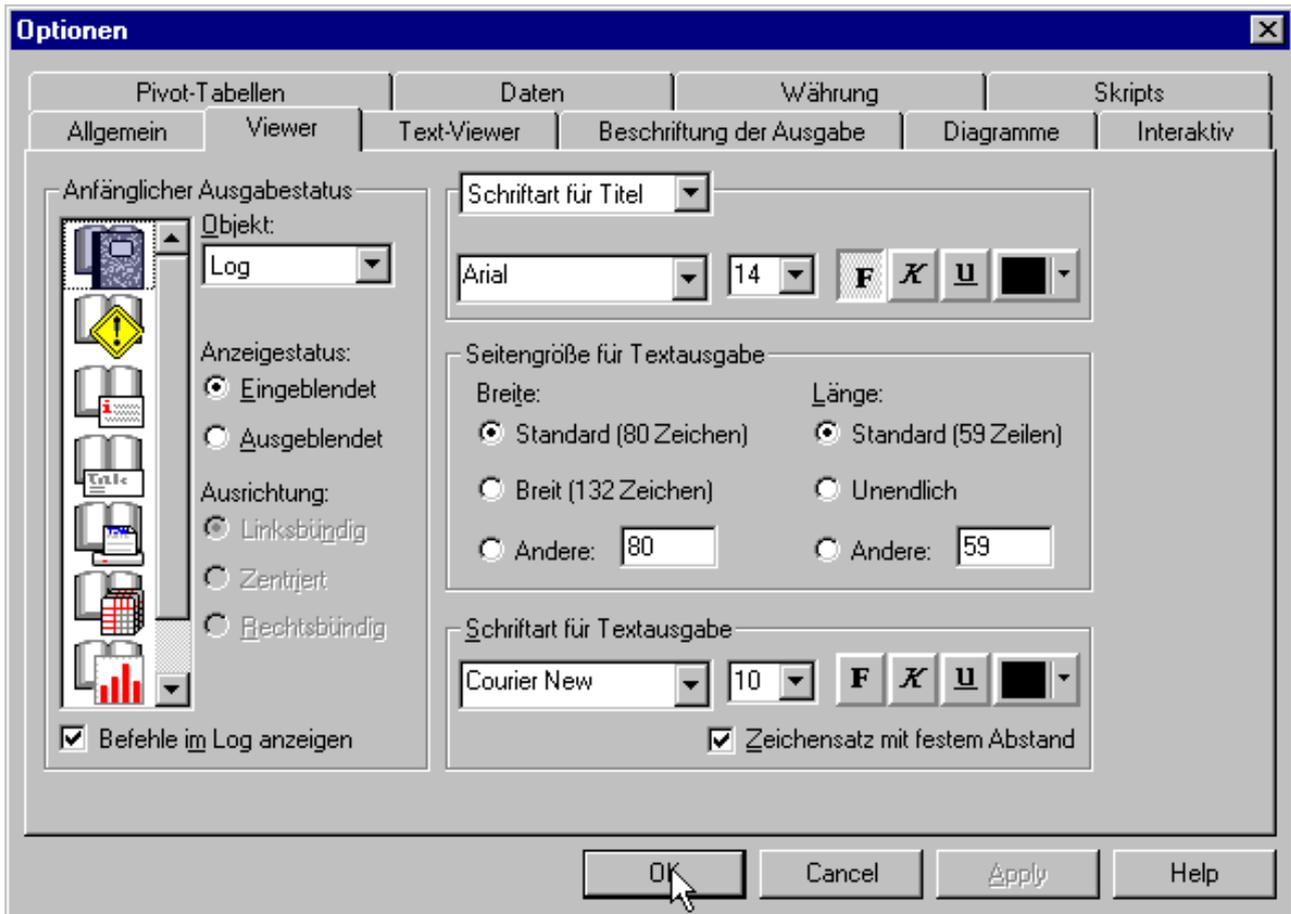
**Registerkarte „Viewer“**

## 4. Anfänglicher Ausgabestatus:

- a. Objekt: Log (SPSS-Befehle werden in der Ausgabe protokolliert.)

5. Anzeige status:  Eingebledet

6.  Befehle im Log anzeigen



Auch bei den anderen Ausgabeobjekten des Viewers sollte der Anzeigestatus „Eingebledet“ sein: Warnungen, Anmerkungen, Seitentitel, Pivot-Tabelle, Diagramm.

Alle anderen SPSS-Programm-Einstellungen werden übernommen und sollten erst bei Bedarf geändert werden.

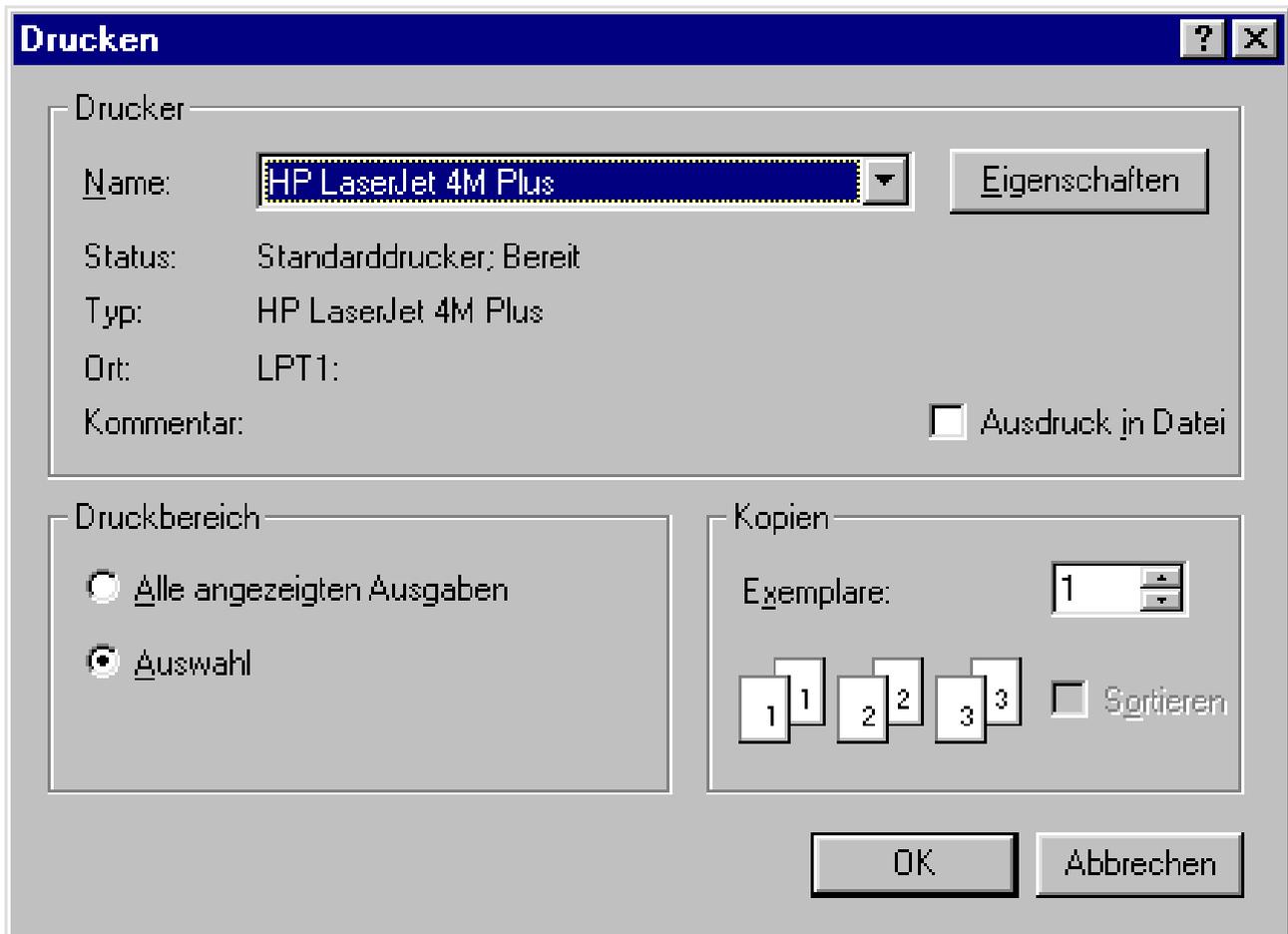
### 2.1.7 Drucken

Die SPSS-Druckfunktionen sind fensterspezifisch, d.h. auf die Fensterinhalte abgestimmt. Sie werden über die Datei-Menüs der Fenster angesprochen. Drucken kann man aus den Fenstern der Viewer und denen des Daten- und Syntax-Editors.

#### 2.1.7.1 Viewer

Pivot-Tabellen, Diagramme und Textausgaben können insgesamt oder einzeln gedruckt werden:

►Datei ►Drucken



Wie bei anderen Windows-Programmen kann man vor dem Drucken noch die Druckereigenschaften verändern und die Anzahl der Kopien wählen. Es ist auch möglich, die Druckausgabe in eine Datei umzulenken.

Vor dem Drucken sollte man sich das Layout der zu druckenden Seiten zeigen lassen. Die im Ausgabefenster nicht sichtbaren Seiten- und Tabellenumbrüche, die verborgenen Pivot-Tabellen-Schichten sowie die Kopf- und Fußzeilen werden in der Vorschau dargestellt:

►Datei ►Seitenansicht

Bei Bedarf können vor dem Drucken noch das Papierformat und die Seitenränder verändert und Kopf- und Fußzeilen mit Seitennumerierung, Datum und Uhrzeit definiert werden:

**Pivot-Tabellen** sind häufig zu groß für eine Seite. Der Pivot-Tabellen-Editor bietet Möglichkeiten der Seitengestaltung für Tabellen. So können die Zeilen und Spalten festgelegt werden, an denen die Tabelle aufgetrennt werden darf

►Doppelklick auf Pivot-Tabelle  
 ►Klick auf Umbruchstelle  
 ►Format ►Umbruch hier

bzw. nicht geteilt werden sollte:

►Doppelklick auf Pivot-Tabelle  
 ►Zusammenzuhaltende Zeilen und Spalten markieren: Klick und Ziehen  
 ►Format ►Zusammenhalten

Eine Tabelle kann an die Seitengröße angepaßt werden:

- ▶ Doppelklick auf Pivot-Tabelle
- ▶ Format ▶ Tabelleneigenschaften

Alle über „Einstellungen“ und mit dem Tabellen-Editor vorgenommenen Änderungen wirken sich nur auf die Druckausgabe und nicht auf das Ausgabefenster aus. Das neue Layout sollte vor dem Drucken mit „Seitenansicht“ überprüft werden.

### 2.1.7.2 Text-Viewer

Ähnlich wie beim Viewer kann die Ausgabe insgesamt gedruckt werden oder es kann vor dem Druck eine Auswahl getroffen werden. Die Textauswahl erfolgt durch Markieren der zu druckenden Stellen, die Diagrammauswahl durch Anklicken. Eine Seitenvorschau ist ebenfalls möglich.

### 2.1.7.3 Daten-Editor

Daten werden so gedruckt, wie sie im Daten-Editor-Fenster dargestellt sind. Über das Menü „Ansicht“ können Gitterlinien ein- oder ausgeblendet werden, statt der Datenwerte kann man sich deren Etiketten anzeigen lassen. Nach „Datei“ und „Drucken“ kann „Alles“ oder eine „Auswahl“ der Daten zum Drucker geschickt werden.

Für das Drucken größerer Datendateien, wie z.B. des ALLBUS-Datensatzes, empfehlen wir den SPSS-Befehl  *LIST*, mit dem sich Variablen und Fälle für die Datenauflistung auswählen und geeignete Ausgabeformate vergeben lassen.

### 2.1.7.4 Syntax-Editor

Syntaxdateien oder Teile davon werden über das Datei-Menü des Syntax-Editors gedruckt. Nach „Datei“ und „Drucken“ kann man „Alles“ oder bestimmte Seiten auswählen.

## 2.2 Eine SPSS-Beispielsitzung

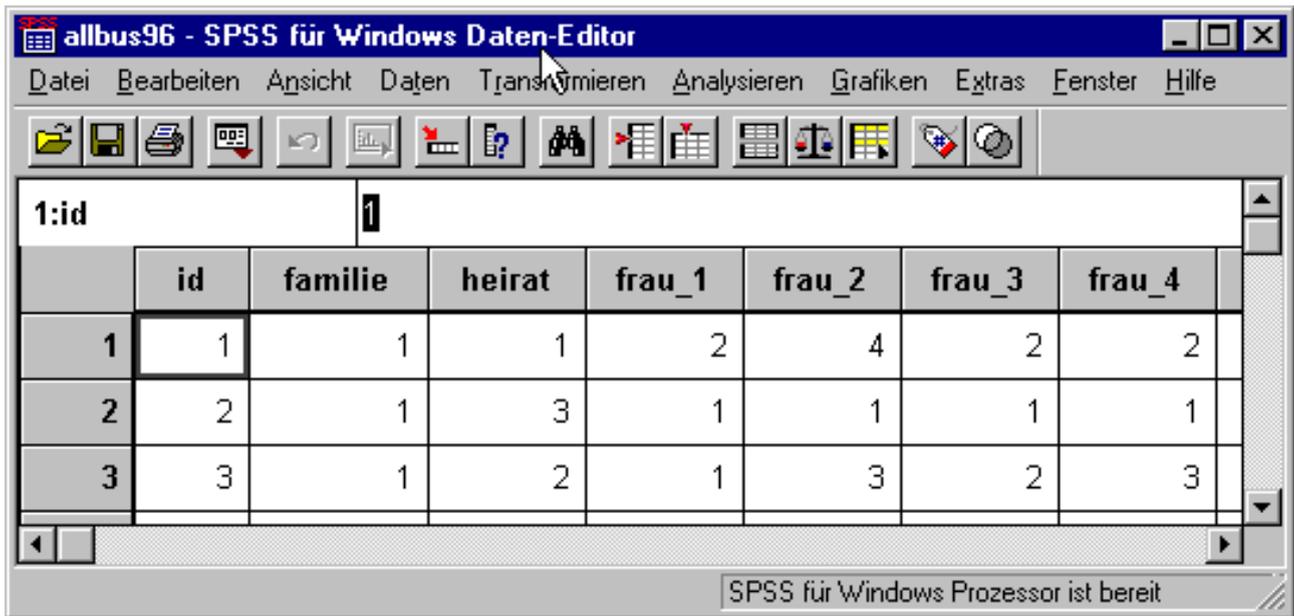
In dieser Beispielsitzung gehen wir von der SPSS-Datendatei aus, die aus dem ALLBUS96-Datensatz erzeugt wurde (*allbus96.sav*). Im Anhang ist beschrieben, wo diese Datei erhältlich ist und wie sie auf der Festplatte angelegt wird. Außerdem müssen die SPSS-Programmeinstellungen entsprechend den Angaben im Kapitel „Programmoptionen“ vorgenommen werden.

### 2.2.1 Lesen einer SPSS-Datendatei

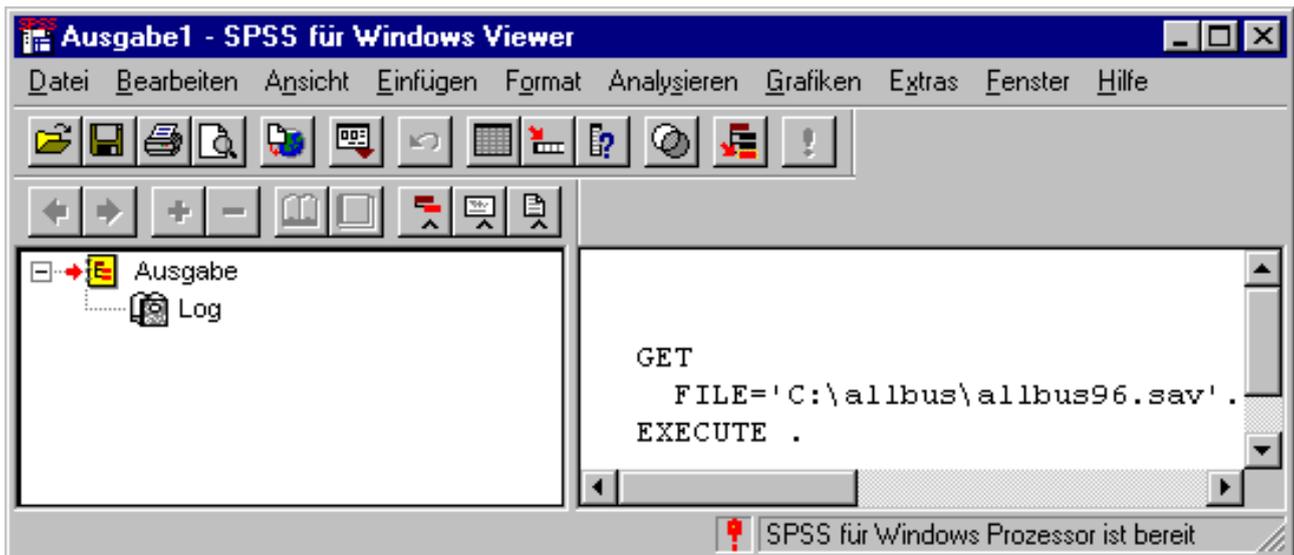
Wir rufen SPSS auf und öffnen die SPSS-Datendatei im Daten-Editor-Fenster:

- ▶ Datei ▶ Öffnen
- ▶ Verzeichnis wählen: c:\allbus
- ▶ Dateityp: SPSS (\*.sav)
- ▶ Dateiname: allbus96.sav
- ▶ Öffnen

Im Datenfenster werden nun die Namen und Werte der Variablen angezeigt. Vom Gesamtdatensatz kann nur ein Ausschnitt betrachtet werden. Mit Maus und Pfeiltasten, mit Symbolen zum Rauf- und Runterschieben sowie zum Schieben nach rechts und links kann man sich durch die gesamte Datendatei bewegen. Mit der Suchfunktion kann man gezielt nach einzelnen Werten suchen, mit der Sprungfunktion zu einem bestimmten Fall positionieren.



Gleichzeitig öffnet SPSS ein Ausgabefenster, das die für das Öffnen der Datendatei abgearbeiteten SPSS-Befehle (*GET*, *EXECUTE*) enthält:



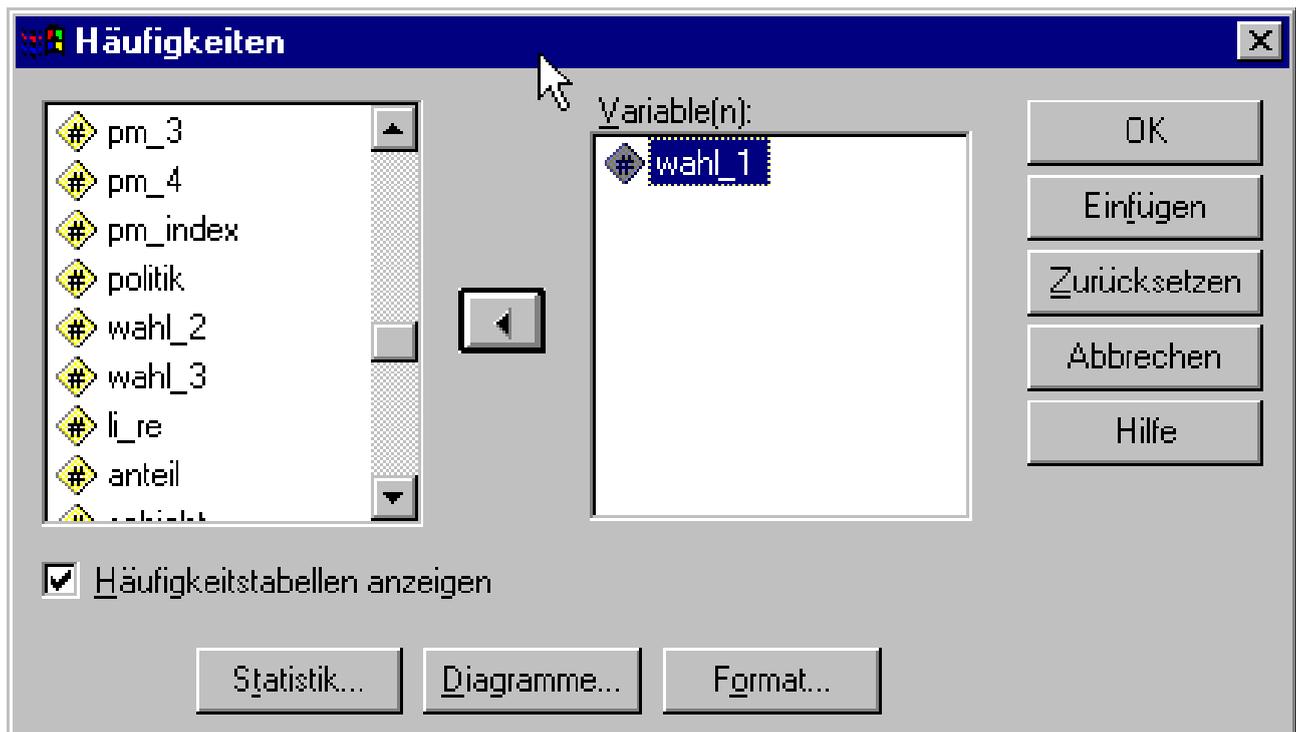
### 2.2.2 Aufruf einer Statistikprozedur

Im ALLBUS-Fragebogen wurde die sogenannte Sonntagsfrage gestellt: „Wenn am nächsten Sonntag Bundestagswahl wäre, welche Partei würden Sie dann mit Ihrer Zweitstimme wählen?“ Wir untersuchen nun die Antworten der befragten Personen durch eine Häufigkeitsauszählung und lassen uns ein Balkendiagramm ausgeben.

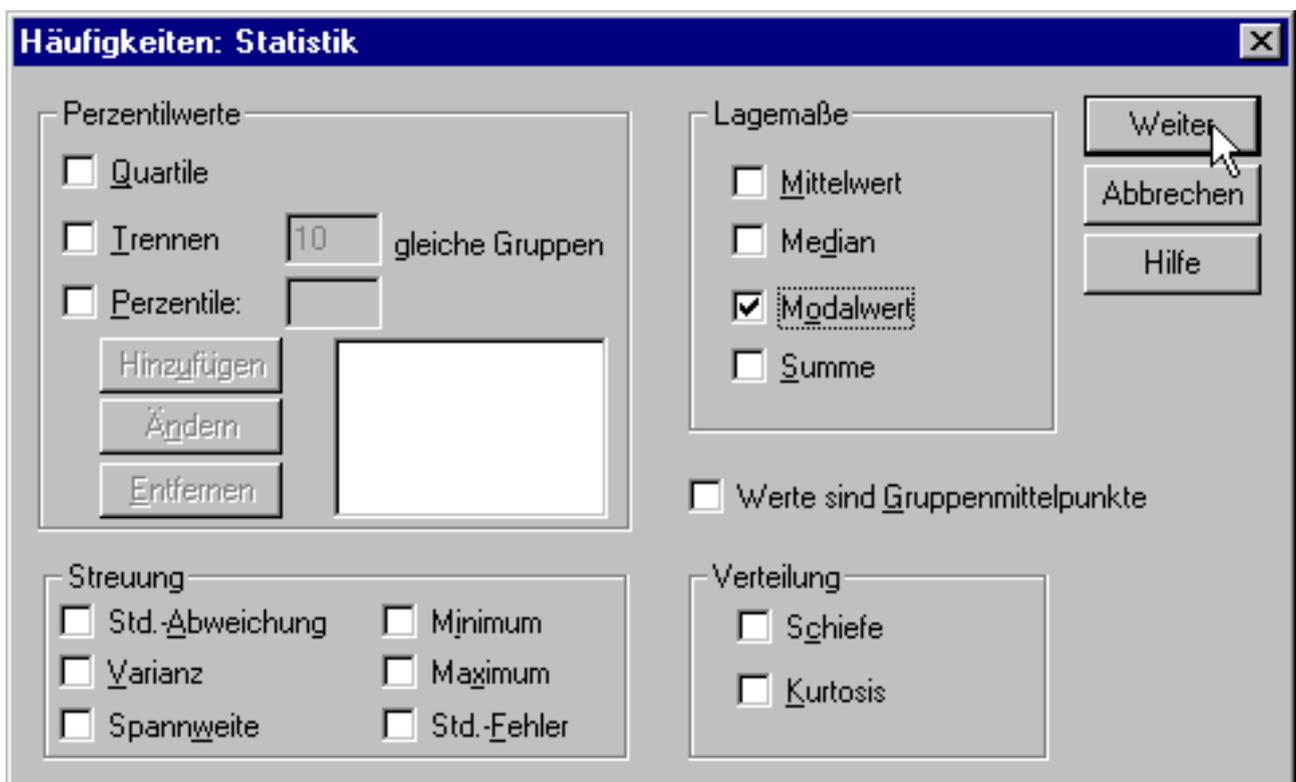
Im Daten-Editor-Menü wählen wir:

►Analysieren ►Deskriptive Statistiken ►Häufigkeiten

Im linken Fenster der Dialogbox „Häufigkeiten“ wählen wir durch Anklicken die gewünschte Variable *wahl\_1* aus und „schieben“ sie mit dem Symbol „►“ in das rechte Fenster.



Über die Schaltfläche „Statistik“ werden die zu berechnenden statistischen Kennwerte festgelegt.



Da *wahl\_1* eine nominalskalierte Variable ist, kann sinnvoller Weise nur der Modalwert (der am häufigsten auftretende Wert) berechnet werden. Mit „Weiter“ wird das „angekreuzte“ Lagemaß übernommen.

Zurück in der Dialogbox „Häufigkeiten“ wählen wir nun noch über die Schaltfläche „Diagramme“ das gewünschte Balkendiagramm aus:



Nach „Weiter“ sind wir wieder in der Dialogbox „Häufigkeiten“, in der nun mit „OK“ die Berechnung angestoßen wird.

### 2.2.3 Die Ergebnisse im Ausgabefenster

Im Ausgabefenster werden die durch die Menü-Eingaben erzeugten SPSS-Befehle, der angeforderte Statistikwert, die Häufigkeitstabelle und das Balkendiagramm angezeigt:



Im linken Fensterflügel ist die Gliederung der Ausgabe angeordnet, im rechten wird der Inhalt des durch Anklicken ausgewählten Ausgabeobjekts angezeigt. Die vorangestellten Kästchen (□) ermöglichen das Ein- (-) und Ausblenden (+) der einzelnen Objekte.

Die Ausgabe enthält unterschiedliche Typen von Objekten. Die Befehle (Log) werden als Textobjekt zusammengefaßt, Statistik- und Häufigkeitstabelle sind Pivot-Tabellen, und das Balkendiagramm ist vom Typ „SPSS-Diagramm“. Nach Anklicken der Objekte mit der rechten Maustaste öffnen sich objektabhängige Untermenüs zur weiteren Bearbeitung.

Unter „Log“ werden alle bisher von SPSS verarbeiteten Befehle aufgeführt:

```

GET
  FILE='C:\allbus\allbus96.sav'.
EXECUTE .
FREQUENCIES
  VARIABLES=wahl_1
  /STATISTICS=MODE
  /BARCHART  FREQ
  /ORDER  ANALYSIS .

```

Die Befehle werden nur dann im Ausgabefenster gelistet, wenn die von uns empfohlenen Einstellungen bei den Programmoptionen vorgenommen wurden. Die Auflistung ist hilfreich bei der Fehlersuche, da sich eventuelle Fehlermeldungen auf den ausgeführten Befehl beziehen.

Die Befehle können auch zunächst in ein Syntaxfenster geschrieben, bei Bedarf modifiziert und erst dann zur Ausführung an SPSS übergeben werden. Diese Vorgehensweise wird ausführlich bei der Vorstellung der einzelnen SPSS-Befehle beschrieben.

Nach dem Titel „Häufigkeiten“ folgen in unserem Beispiel direkt die angeforderten Statistiken, „Anmerkungen“ werden nur bei Bedarf von SPSS ausgegeben:

#### Statistiken

##### Wahlabsicht, Bundestagswahl

	N	Gültig	Fehlend	
Modus				2590
				928
				1

2590 Personen haben eine gültige Antwort gegeben, von 928 liegt keine Antwort vor. Die meisten der Befragten würden CDU-CSU (Modus=1) wählen.

In der angeforderten Häufigkeitstabelle sind die gültigen und fehlenden Antworten zusammengestellt.

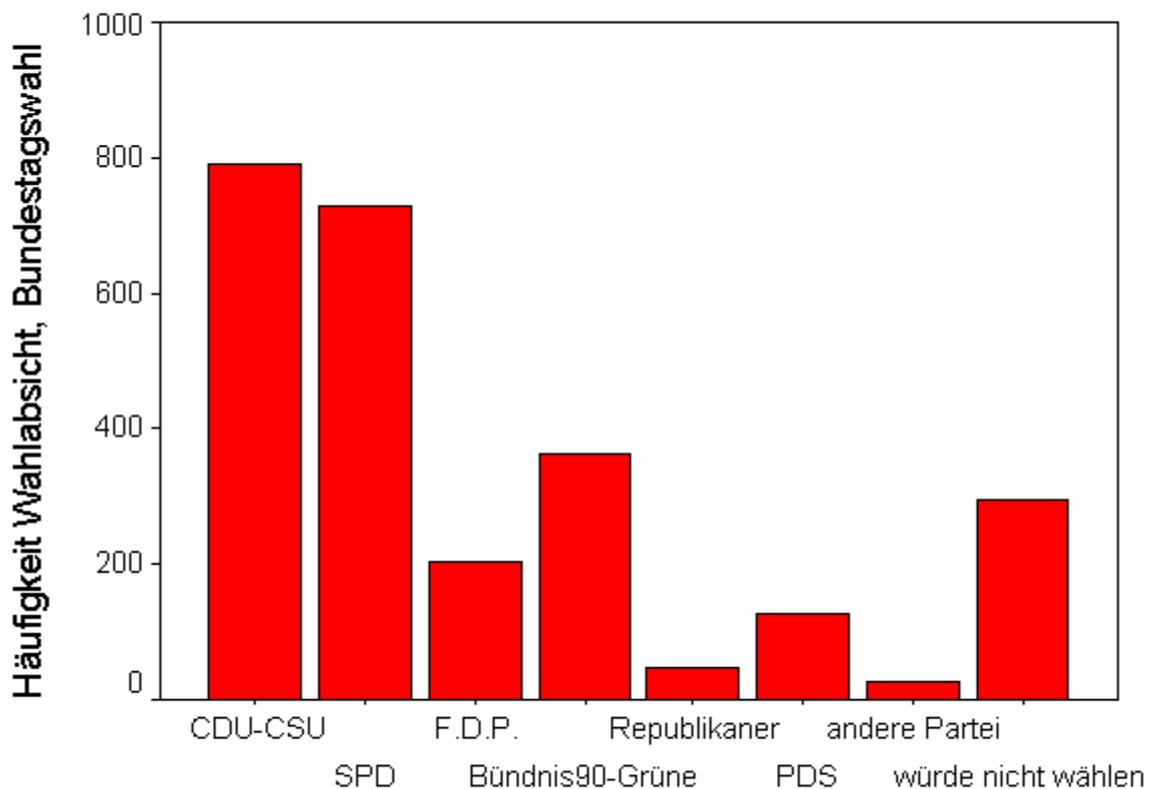
##### Wahlabsicht, Bundestagswahl

		Häufigkeit	Prozent	Gültige Prozen te	Kumulierte Prozen te
Gültig	CDU-CSU	791	22,5	30,5	30,5
	SPD	730	20,8	28,2	58,7
	F.D.P.	206	5,9	8,0	66,7
	Bündnis90-Grüne	362	10,3	14,0	80,7
	Republikaner	49	1,4	1,9	82,5
	PDS	126	3,6	4,9	87,4
	andere Partei	29	,8	1,1	88,5
	würde nicht wählen	297	8,4	11,5	100,0
	Gesamt	2590	73,6	100,0	
	Fehlend	tnz: nicht wahlberechtigt	212	6,0	
vw		234	6,7		
wn		476	13,5		
kA		6	,2		
Gesamt		928	26,4		
Gesamt		3518	100,0		

Wir sehen die absoluten und die prozentualen Häufigkeiten der im ALLBUS-Datensatz vorkommenden Werte der Variablen *wahl\_1*. Häufigkeitstabellen werden ausführlich unter der Prozedur *FREQUENCIES* erläutert. Der Nachbearbeitung und Präsentation von Tabellen widmen wir das gesamte Kapitel 5. Hier werden auch die Pivotisierungsmöglichkeiten, d.h. das Vertauschen von Spalten, Zeilen und Schichten, beschrieben.

Im Gliederungsfenster klicken wir „Balkendiagramm“ an, und sehen nun die absoluten Häufigkeiten der Variablen *wahl\_1* als Balken dargestellt:

## Wahlabsicht, Bundestagswahl



## Wahlabsicht, Bundestagswahl

Präsentationsgrafiken werden ausführlich in [Kapitel 4](#) behandelt.

### 2.2.4 Die Ergebnisse drucken und weiterverarbeiten

Der Drucker wird ähnlich wie in anderen Windows-Programmen bedient. Die Ausgabe oder ausgewählte Objekte daraus werden über das Menü des Ausgabefensters gedruckt:

►Datei ►Drucken  
 Druckbereich:  
 Alle angezeigten Ausgaben  
 oder  
 Auswahl

Die Ausgabeobjekte können mit „Bearbeiten“ und „Kopieren“ in die Windows-Zwischenablage gebracht und von dort von Textverarbeitungs-Systemen wie z.B. Word und WordPerfect mit „Einfügen“ in den Text übernommen werden.

Auch über Dateien ist ein Einbinden in Dokumente möglich. Ein Klick mit der rechten Maustaste auf ein Ausgabeobjekt öffnet ein Menü, das eine Exportfunktion anbietet. Textobjekte und Pivot-Tabellen können als HTML- und Text-Dateien (\*.htm, \*.txt), Diagramme in einem der angebotenen Grafikdatei-Formate exportiert werden.



Liegen die Daten noch nicht in einer Datei vor, empfiehlt sich die Eingabe mit dem  SPSS-Daten-Editor. Dieses ist bequemer und weniger fehleranfällig als die Erfassung mit einem Text-Editor.

### 2.3.3 Fälle, Variablen und Werte

Die Fragen nach Einkommen und Geschlecht z.B. sind für jeden Fragebogen gleich, die Antworten natürlich von Person zu Person verschieden. Ein Fragebogen oder eine Person wird als Fall („case“), eine Frage als Variable („variable“) und eine Antwort als Wert („value“) bezeichnet. Eine Variable hat für jeden Fall genau einen Wert.

## 2.4 Datendateien

Die durch die Kodierung gewonnenen Zahlen und Zeichenfolgen können mit dem SPSS-Daten-Editor erfaßt und als  **SPSS-Datendatei** gespeichert werden. Auch mit Text-Editoren<sup>1</sup>, Textverarbeitungssystemen<sup>2</sup>, Tabellenkalkulations<sup>3</sup>- und Datenbank-Programmen<sup>4</sup> erstellte Datendateien können von SPSS für Windows übernommen werden.

Eine Export-/Import-Funktion ermöglicht den Austausch von  portablen SPSS-Datendateien zwischen verschiedenen Rechnerplattformen, z.B. zwischen Windows- und UNIX-Rechnern.

Ebenso kann SPSS die Datendateien anderer Statistik-Programmsysteme wie z.B. SAS<sup>5</sup> lesen. Näheres dazu findet man in der Online-Hilfe (SPSS-Befehl *GET SAS*) und den Handbüchern.

In unserem Fall wurden die mit dem ALLBUS-Fragebogen erhobenen Daten entsprechend dem zugehörigen Kodierplan umgesetzt und in einer Textdatei, auch ASCII- oder ANSI-Datei<sup>6</sup> genannt, in festem Format abgelegt.

In Abhängigkeit von der Art der Datenerfassung ergeben sich unterschiedliche Vorgehensweisen bei der SPSS-Datenbeschreibung, d.h. bei der Vergabe der **Variablennamen**, der Festlegung des **Typs** und des **Meßniveaus der Variablen** sowie der Definition der **fehlenden Werte**, **Variablen- und Wertetiketten**. Beim Variablentyp ist darauf zu achten, daß SPSS **numerische Variablen**, die nur Zahlen als Datenwerte annehmen dürfen, und **Zeichenketten-Variablen** (auch **alphanumerische Variablen** genannt), deren Werte aus Buchstaben, Ziffern und Sonderzeichen bestehen können, unterscheidet. Die für die Datenbeschreibung erforderlichen Begriffe werden ausführlich in  Kapitel 2.6 („Datenbeschreibung und Datenerfassung“) beschrieben.

So wie SPSS viele verschiedene Datendateien lesen kann, ist es auch möglich, Datendateien für andere Programme und SPSS-Versionen zu schreiben, damit sie von diesen weiterverarbeitet werden können.

### 2.4.1 Erstmalige Datenerfassung

Sind die Daten noch nicht in einer Datei verfügbar, so empfiehlt sich die Datenbeschreibung mit dem SPSS-Daten-Editor.

<sup>1</sup> Windows95/98/NT: ▶Start ▶Programme ▶Zubehör ▶Editor

<sup>2</sup> In Textverarbeitungssystemen wie z.B. WordPad, Word oder WordPerfect muß die Rohdatendatei als Windows-Textdatei abgespeichert werden.

<sup>3</sup> z.B. Excel

<sup>4</sup> z.B. Access, dBASE oder Lotus 1-2-3

<sup>5</sup> SAS: Statistical Analysis System von SAS Inc.

<sup>6</sup> ASCII und ANSI sind amerikanische Normen für Textdateiformate.

Die Daten werden ebenfalls mit dem Daten-Editor eingegeben und zusammen mit den Variableninformationen als SPSS-Datendatei gespeichert.

Die Dateneingabe mit dem SPSS-Daten-Editor ist der Erfassung mit Text-Editoren und Textverarbeitungssystemen vorzuziehen, da sie bequemer und weniger fehleranfällig ist. Der Daten-Editor achtet u. a. auf den richtigen Typ und das richtige Format der eingegebenen Daten.

## 2.4.2 Rohdatendateien lesen

Die Daten befinden sich bereits in einer Rohdatendatei. Sie wurden mit einem Text-Editor oder einem Textverarbeitungssystem in eine ASCII- oder ANSI-Textdatei in festem (wie der ALLBUS-Datensatz), freiem oder Tabulator-getrenntem Format eingegeben. Weitere Rohdatendatei-Formate sind im „SPSS Base Syntax Reference Guide“ bei den Befehlen *DATA LIST*, *READ MATRIX* und *FORMATS* beschrieben.

### 2.4.2.1 Rohdatendateien in festem Format

Festes Format bedeutet, daß die Werte für jeden Fragebogen in denselben Spalten einer Datenzeile untergebracht sind. Festformatierte Daten können auf zwei Arten von SPSS gelesen werden:

1. Die Datenbeschreibung wird im Syntax-Editor-Fenster mit den Befehlen *DATA LIST*, *MISSING VALUES*, *VARIABLE LEVEL*, *VARIABLE LABELS* und *VALUE LABELS* durchgeführt. Es ist auch möglich, zunächst nur die Variablenamen und den Variablentyp mit *DATA LIST* zu definieren und die fehlenden Werte, das Meßniveau und die Etiketten nachträglich mit dem Daten-Editor zu beschreiben.
2. Die Daten werden mit Hilfe des Textimport-Assistenten gelesen und anschließend mit dem Daten-Editor beschrieben:

►Datei ►Texdaten einlesen

Wir empfehlen bei vielen Variablen die erste Methode mit vollständiger Datenbeschreibung über die SPSS-Befehle.

### 2.4.2.2 Rohdatendateien in freiem Format

In einer freiformatierten Datendatei sind die einzelnen Werte hintereinander, nur durch eine oder mehrere Leerstellen voneinander getrennt, in der Datei angeordnet. Die Daten eines Fragebogens müssen nicht jeweils in einer neuen Datenzeile beginnen.

Die Datenbeschreibung wird wie bei festformatierten Dateien mit *DATA LIST* unter Angabe des Schlüsselwortes *FREE* durchgeführt.

### 2.4.2.3 Rohdatendateien mit Tabulator-Trennzeichen

Auch diese sind Textdateien. Die Datenwerte liegen, ähnlich wie bei einem Tabellenkalkulationsblatt, in der Datei untereinander in Spalten, die durch Tabulatoren voneinander getrennt sind. SPSS ordnet beim Einlesen einer Tabulator-Datei jeweils die Werte einer Spalte einer Variablen zu. Enthält die erste Zeile der Datei Variablenamen, so können diese mitübernommen werden. Der Variablentyp richtet sich jeweils nach dem Wert in der ersten Datenzeile. Ist dies eine Zahl, so wird es eine numerische Variable. Enthält das Feld Buchstaben oder Sonderzeichen, wird eine Zeichenketten-Variable definiert.

Rohdatendateien mit Tabulator-Trennzeichen werden im Daten-Editor-Fenster wie folgt gelesen:

►Datei ►Öffnen ►Dateityp: Tabulatorbegrenzt (\*.dat, \*.txt)

### 2.4.3 Tabellenkalkulations- und Datenbank-Dateien lesen

Die Daten wurden mit einem Tabellenkalkulations- oder Datenbank-Programm erfaßt. Daten und auch Variablenamen, sofern sie definiert wurden, können von SPSS im Daten-Editor-Fenster übernommen werden.

Datendateien, die mit einem der folgenden Programme erzeugt wurden, werden mit **►Datei ►Öffnen** unter Auswahl einer der folgenden Dateitypen gelesen:

- Excel (\*.xls) (Excel; bis SPSS 9 Excel-Versionen 4.0 und älter, ab SPSS 10 Dateien aller Excel-Versionen)
- Lotus (\*.w\*) (Lotus 1-2-3; Versionen 3.0, 2.0 und 1A)
- SYLK (\*.slk) (Symbolic-Link-Dateien einiger anderer Tabellenkalkulations-Programme)
- dBASE (\*.dbf) (dBASE; Versionen IV, III, und II)

Die ausgewählte Datei wird im Daten-Editor-Fenster angezeigt, die Datenbeschreibung kann ergänzt werden. Die übernommenen Daten und deren Beschreibung werden in einer SPSS-Datendatei gespeichert.

Für Datendateien, die nicht direkt in SPSS eingelesen werden können, können die Daten über die ODBC-Schnittstelle<sup>1</sup> der verwendeten Programme ausgetauscht werden.

SPSS für Windows kann Daten von allen Datenbanksystemen und allen Programmen, für die ein ODBC-Treiber<sup>2</sup> verfügbar ist, übernommen werden; so auch z.B. von Informix, Oracle und Paradox.

Die ODBC-Schnittstellen der Programme sind im Daten-Editor-Fenster erreichbar über: **►Datei ►Datenbankzugriff** .

Vor der Übernahme einer Datei kann eine Auswahl der von SPSS zu lesenden Tabellen, Felder und Daten getroffen werden. Die ausgewählten Daten werden im Daten-Editor-Fenster angezeigt, die Datenbeschreibung kann ergänzt werden.

SPSS bietet vielfältige Möglichkeiten zur Datenübernahme aus Programmen mit ODBC-Schnittstelle, auf die hier aber nicht ausführlicher eingegangen werden kann. Wir verweisen für weitergehende Informationen auf die Online-Hilfe (SPSS-Befehle *GET TRANSLATE*, *GET CAPTURE*) und die Handbücher.

### 2.4.4 SPSS-Datendateien

SPSS-Datendateien werden ausführlich in Kapitel 2.9 behandelt. Hier soll nur kurz die Menü-Benutzung zum Schreiben und Lesen dieser Dateien beschrieben werden. SPSS hält die Daten und deren Beschreibung (Variablenamen, Meßniveau, fehlende Werte, Variablen- und Wertetiketten) in einer temporären Arbeitsdatei, die als permanente SPSS-Datendatei abgespeichert und in folgenden SPSS-Sitzungen wieder verwendet werden kann.

#### 2.4.4.1 SPSS-Datendateien schreiben

Aus einer temporären Arbeitsdatei wird über das Datei-Menü des Daten-Editors eine permanente SPSS-Datendatei erzeugt mit:

---

<sup>1</sup> Open Database Connectivity (ODBC): genormte Schnittstelle zum Datenaustausch zwischen Programmen.

<sup>2</sup> Programm, das die Verbindung zwischen Datenbanksystem und SPSS herstellt.

►Datei ►Speichern unter ►Dateityp: SPSS (\*.sav)

Diese Menüfolge wird auf den SPSS-Befehl `SAVE` abgebildet.

#### 2.4.4.2 SPSS-Datendateien lesen

Eine permanente SPSS-Datendatei wird über das Datei-Menü des Daten-Editors zu einer temporären Arbeitsdatei mit:

►Datei ►Öffnen ►Dateityp: SPSS (\*.sav)

Diese Menüfolge wird auf den SPSS-Befehl `GET` abgebildet.

#### 2.4.5 Rohdatendateien schreiben

Aus einer Arbeitsdatei können über das Daten-Editor-Menü die Rohdaten in festem Format als Textdatei geschrieben werden:

►Datei ►Speichern unter ►Dateityp: ASCII (feste Formate) (\*.dat)

Die Variablenamen und die Variableninformationen gehen dabei verloren. Freiformatierte Rohdatendateien kann SPSS nicht schreiben, wohl aber Tabulator-getrennte Datendateien (s. o.), in die auf Wunsch neben den Datenwerten auch die Variablenamen übernommen werden können:

►Datei ►Speichern unter ►Dateityp: Tabulator-getrennt (\*.dat)

Weitere Möglichkeiten zum Schreiben von Rohdatendateien bietet der SPSS-Befehl `WRITE`.

#### 2.4.6 Tabellenkalkulations- und Datenbank-Dateien schreiben

Rohdaten und Variablenamen einer Arbeitsdatei lassen sich auch als Dateien für Tabellenkalkulations- und Datenbank-Programme schreiben. Über das Daten-Editor-Menü lassen sich jedoch nur Dateien für ältere Versionen dieser Programme erzeugen. Nach

►Datei ►Speichern unter

ist der gewünschte Dateityp (s. o.) anzugeben. Mit dem SPSS-Befehl `SAVE TRANSLATE` können Datendateien für weitere Programme geschrieben werden.

### 2.5 Die SPSS-Befehlssprache

Auch wenn die Datenanalyse mit SPSS durch übersichtliche Menüs und selbsterklärende Dialogboxen fast zum „Kinderspiel“ geworden ist und sich häufig mit wenigen Mausklicks erledigen läßt, sind zumindest grundlegende Kenntnisse der SPSS-Befehlssprache für ein erfolgreiches Arbeiten mit SPSS unbedingt erforderlich.

Alle datenanalytischen Schritte, die über die Menüs und Dialogboxen durchgeführt werden, bildet das SPSS-Programm auf seine Befehlssprache ab. Die so erzeugten Befehle können auf Wunsch in das Syntaxfenster eingefügt und nach Ausführung im Ausgabefenster protokolliert werden. Fehlermeldungen beziehen sich auf die Befehle und werden in englischer Sprache - auch in deutschen SPSS-Programmversionen - ausgegeben.

Einige Befehle, wie z.B. die zur Datentransformation, lassen sich manuell bequemer und schneller formulieren als menüunterstützt. Außerdem sind nicht alle Statistikverfahren und Hilfsroutinen über die Menüs aufrufbar.

Aus diesen Gründen beschreiben wir hier die SPSS-Befehlssprache im allgemeinen und die Befehle zum Aufruf der verschiedenen Statistikprozeduren in den folgenden Kapiteln jeweils an Ort und Stelle. Zum besseren Verständnis der englischen Fehlermeldungen geben wir auch die englischen Begriffe

der Befehlselemente an.

### 2.5.1 SPSS-Befehlsdateien

Ein SPSS-Programm besteht aus einer Folge von Befehlen („commands“), die an SPSS zur Verarbeitung übergeben werden können. Die Befehle müssen den Regeln (Syntax) der SPSS-Befehlssprache genügen.

Ein SPSS-Programm kann Befehle zur

- Datenbeschreibung („data definition commands“),
- Datentransformation („data transformation commands“),
- Statistikauswertung („procedure commands“)
- und Datendateimanipulation

enthalten.

SPSS-Programme können auf zwei Arten im Syntax-Editor-Fenster - oder auch kurz Syntaxfenster genannt - erstellt werden:

- durch manuelle Eingabe der einzelnen Befehle
- mit Hilfe der Menüs und Dialogboxen

Die so erstellten Programme werden in SPSS-Befehlsdateien - oder auch SPSS-Syntaxdateien genannt - gespeichert.

Beide Methoden werden in den folgenden Kapiteln beschrieben.

#### 2.5.1.1 Ein neues Syntaxfenster öffnen

Beim Start von SPSS wird standardmäßig kein Syntaxfenster, sondern nur ein leeres Datenfenster geöffnet. Die Menüs des Datenfensters bieten folgende Möglichkeiten, ein neues Syntaxfenster zu öffnen:

- über die Menüfolge ▶Datei ▶Neu ▶Syntax
- durch Aufruf einer Statistikprozedur über das Menü, wie z.B.  
▶Analysieren ▶Deskriptive Statistiken ▶Häufigkeiten ▶Einfügen

#### 2.5.1.2 Eine Befehlsdatei speichern und öffnen

Das in einem Syntaxfenster erstellte SPSS-Programm wird als SPSS-Befehlsdatei über das Datei-Menü dieses Fensters gespeichert.

Beispiel:

```
▶Datei ▶Speichern unter  
▶Dateiname: beispiel1.sps ▶Dateityp: SPSS-Syntaxdateien(*.sps)
```

Bestehende Befehlsdateien können zur erneuten Bearbeitung und Ausführung wieder in das Syntaxfenster geladen werden. Dies erfolgt über das Datei-Menü des Datenfensters. Wir öffnen die im obigen Beispiel gespeicherte Datei mit:

```
▶Datei ▶Öffnen  
▶Dateiname: beispiel1.sps ▶Dateityp: Syntax(*.sps)
```

### 2.5.1.3 SPSS-Befehle ausführen

Die im Syntaxfenster erstellten SPSS-Befehle sind zunächst nur Textzeilen ohne jede Beziehung zu SPSS. So können SPSS-Programme auch außerhalb von SPSS mit einem Text-Editor in eine Datei (Typ: \*.sps) geschrieben und in das SPSS-Syntaxfenster geladen werden.

Damit die im SPSS-Programm formulierten Befehle zur Datenbeschreibung und Datentransformation sowie die eigentlichen Statistik- und Grafikprozeduren ausgeführt werden, müssen sie zur Verarbeitung an SPSS übergeben werden. Mit **Ausführen** im Syntaxfenster-Menü können ganze Programme, eine Auswahl von Befehlen und auch einzelne Befehle ausgeführt werden.

SPSS überprüft zunächst nur die Syntax der Befehle. Die Daten werden erst beim Aufruf einer Statistikprozedur verarbeitet. Sollen z.B. Befehle zur Beschreibung, Transformation und Auswahl von Daten sofort auf die Arbeitsdatei bzw. die Daten wirken, muß diesen der Befehl **EXECUTE** folgen.

Die Befehle werden, sofern die von uns vorgeschlagenen SPSS-Programmooptionen (Registerkarte Viewer: Befehle im Log anzeigen) eingestellt wurden, im Ausgabefenster, das von SPSS geöffnet wird, protokolliert. Ebenso werden die Ergebnisse der Statistik- und Grafikprozeduren im Ausgabefenster angezeigt, wo sie mit dem Viewer betrachtet, gedruckt und nachbearbeitet werden können.

### 2.5.2 Befehle

SPSS-Befehle werden in einer Sprache formuliert, die der natürlichen englischen Sprache sehr ähnlich ist. Es werden die Buchstaben a-z (klein und groß, ohne Umlaute und ß), die Ziffern 0-9, die Sonderzeichen wie z.B. + - \* ( ) / . , und das Leerzeichen („blank“) verwendet. Die folgenden Regeln für den Bau und die Gliederung der SPSS-Befehle (Syntax) müssen genau eingehalten werden.

Ein **SPSS-Befehl** („command“) beginnt mit einem **Befehlsnamen** („command name“), dem in der Regel weitere **Angaben** („specifications“) folgen. Die Befehle können beliebig mit Groß- und Kleinbuchstaben geschrieben werden, die folgenden Befehle sind somit gleichwertig:

```
FREQUENCIES VARIABLES=ALTER, B_TAETIG, ...
frequencies variables=alter, b_taetig, ...
Frequencies Variables=Alter, B-Taetig, ...
```

- Jeder Befehl muß in einer neuen Zeile - möglichst in der ersten Spalte - beginnen.
- Jeder Befehl muß mit einem Punkt abgeschlossen werden.
- Jede Zeile darf höchstens 80 Zeichen lang sein.
- Fortsetzungszeilen sollten eingerückt werden.
- Achtung: Leerzeilen wirken wie ein Abschlußpunkt!

Ein **Befehlsname** besteht aus einem oder mehreren fest vorgegebenen Schlüsselwörtern. Mehrere Wörter eines Befehlsnamens werden durch ein oder mehrere Leerzeichen getrennt. Jeder Befehl beginnt auf einer neuen Zeile.

#### Beispiel

```
Spalte 1 2 3 4 5
123456789012345678901234567890123456789012345678901234567890123
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
FREQUENCIES
VARIABLE LABELS
NPAR TESTS
```

**Angaben** können sich aus den folgenden Elementen zusammensetzen.

### 2.5.2.1 Unterbefehle („subcommands“)

- Sie beginnen mit einem fest vorgegebenen Befehlsnamen.
- Sie können nach einem Gleichheitszeichen (=) Angaben haben.
- Mehrere Unterbefehle werden durch einen Schrägstrich (/) getrennt.

Beispiel

```
FREQUENCIES VARIABLES=alter
  /STATISTICS=ALL/HISTOGRAM=NORMAL.
```

*VARIABLES*, *STATISTICS* und *HISTOGRAM* sind Unterbefehle.

### 2.5.2.2 Variablennamen („variable names“)

Zur Kennzeichnung von Variablen werden ihnen eindeutige Namen nach den folgenden Regeln gegeben:

- Länge: max. 8 Zeichen
- 1. Zeichen: Buchstabe (Groß- und Kleinbuchstaben)
- Folgezeichen: Buchstaben, Ziffern, Punkt (.), Unterstreichungszeichen ( \_ ) und die Zeichen @, #, \$
- Letztes Zeichen: nicht Punkt oder Unterstreichungszeichen
- verboten: das Leerzeichen („blank“) und die anderen Sonderzeichen sowie von SPSS verwendete Schlüsselwörter wie z.B. ALL, AND, BY, EQ, TO, WITH

Es wird nicht zwischen Groß- und Kleinbuchstaben unterschieden, die Variablennamen *alter*, *Alter* und *ALTER* z.B. sind für SPSS identisch. In der Regel, wie auch in diesem Buch, werden Variablennamen klein geschrieben.

Beispiel

```
FREQUENCIES VARIABLES=familie, frau_1, b_status.
```

*familie*, *frau\_1* und *b\_status* sind gültige Variablennamen.

### 2.5.2.3 Werte („values“)

Werte in SPSS-Befehlen beziehen sich auf Datenwerte, die die Variablen in der Datendatei annehmen, oder sie werden als Konstanten in Berechnungen verwendet. Es gibt numerische (Zahlen) und alphanumerische Werte (Zeichenketten):

#### **Zahlen** („numbers“)

SPSS kennt ganze Zahlen („integer numbers“) und Dezimalzahlen („decimal numbers“). Zahlen können in den Daten und in SPSS-Befehlen mit und ohne Vorzeichen (+ oder -) angegeben werden. Bei Weglassen des Vorzeichens ist die Zahl positiv.

Bei Dezimalzahlen in SPSS-Befehlen ist - unabhängig von der Ländereinstellung für Zahlen in der Systemsteuerung von Windows - statt des Dezimalkommas immer der Dezimalpunkt zu setzen.

Gültige Zahlen sind:

63 (=+63), -1, +10.5 (=+10,5), -0.75 (= -0,75)

#### Beispiel

```
|SELECT IF(f_stand EQ 5).
```

Dieser SPSS-Befehl wählt nur die Fälle aus, bei denen die Variable *f\_stand* (Familienstand) den Wert 5 (ledig) hat.

#### Zeichenketten („strings“)

Variablen können anstelle von Zahlen auch Zeichenketten (d.h. Folgen von Buchstaben, Ziffern und Sonderzeichen) als Datenwerte haben. SPSS kennt kurze (max. 8 Zeichen) und lange (max. 255 Zeichen) Zeichenketten-Variablen, auch alphanumerische oder String-Variablen genannt. Für die langen Zeichenketten-Variablen gelten besondere Regeln, auf die hier nicht eingegangen werden soll.

Zeichenketten müssen innerhalb von SPSS-Befehlen in Apostroph (') oder Anführungszeichen (") eingeschlossen werden. In den Datendateien werden sie nicht in besonderen Zeichen eingeschlossen. Es muß zwischen Groß- und Kleinbuchstaben unterschieden werden, Leerzeichen sind relevant.

Gültige Zeichenfolgen für kurze Zeichenketten-Variablen sind:

'm', 'W', "-", "Big Mac", "M&M", '4711'

#### Beispiel

```
|SELECT IF(sex EQ 'w').
```

Dieser SPSS-Befehl wählt nur die Fälle aus, bei denen die Variable *sex* (Geschlecht) den Wert *w* (weiblich) hat.

#### 2.5.2.4 Schlüsselwörter („keywords“)

Schlüsselwörter werden zur Benennung von Befehlen, Funktionen, Operatoren und anderen Angaben benötigt. Ihre Schreibweise ist fest vorgegeben.

#### Beispiel

```
|RECODE alter (MISSING = 9)
  (LOWEST THRU 50 = 1) (50 THRU HIGHEST = 2).
```

*RECODE*, *MISSING*, *LOWEST*, *THRU* und *HIGHEST* sind Schlüsselwörter.

#### 2.5.2.5 Zeichenketten („strings“)

Zusätzlich zu den Zeichenketten für die Angabe von Datenwerten in SPSS-Befehlen (s. o.) werden Zeichenketten zur Formulierung von Überschriften und Etiketten benötigt. Auch diese Zeichenketten müssen in Apostroph (') oder Anführungszeichen (") eingeschlossen werden. Es muß zwischen Groß- und Kleinbuchstaben unterschieden werden, Leerzeichen werden mitübernommen. In der deutschen SPSS-Version sind in diesen Zeichenketten auch Umlaute und „ß“ erlaubt.

#### Beispiele

```
|VARIABLE LABELS id "Identifikationsnummer".
VALUE LABELS ost_west 1 'alte Bundesländer'
                2 'neue Bundesländer'.
```

#### 2.5.2.6 Arithmetische Operatoren („arithmetic operators“)

SPSS verwendet die Zeichen + (Addition), - (Subtraktion), \* (Multiplikation), / (Division), \*\* (Expo-

mentation) und = (Wertzuweisung) für arithmetische Operationen.

#### Beispiel

```
COMPUTE
  konindex=(kontakt1+kontakt2+kontakt3+kontakt4)/4.
```

Die Werte der Variablen *kontakt1*, *kontakt2*, *kontakt3*, *kontakt4* werden addiert, die Summe durch 4 geteilt, und der neue Wert der Variablen *konindex* zugewiesen. Runde Klammern werden eingesetzt, um die Abarbeitung des arithmetischen Ausdrucks zu steuern.

#### 2.5.2.7 Trennzeichen („delimiters“)

Trennzeichen dienen zur Trennung von Befehlselementen, wie z.B. Werten und Schlüsselwörtern.

##### Gewöhnliche Trennzeichen

Normalerweise wird das Leerzeichen („blank“) benutzt, um eine Angabe von der anderen zu trennen. Anstelle des Leerzeichens kann auch häufig, wie z.B. bei der Aufzählung von Variablen, das Komma (,) eingesetzt werden. Die beiden folgenden Zeilen sind gleichwertig:

#### Beispiel

```
kontakt1 kontakt2 kontakt3 kontakt4
=
kontakt1,kontakt2,kontakt3,kontakt4
```

##### Spezielle Trennzeichen

(, ), /, ", ' und = haben eine besondere Trennfunktion, vor und nach ihnen können ein oder mehrere Leerzeichen eingefügt werden. Der Schrägstrich (/) dient hauptsächlich der Abtrennung von Unterbefehlen und Variablenlisten. Das Gleichheitszeichen (=) wird zwischen Unterbefehlen und deren Angaben gesetzt.

#### Beispiel

```
FREQUENCIES VARIABLES=alter
  /STATISTICS=MEAN/HISTOGRAM=NORMAL.
```

#### 2.5.3 Variablen („variables“)

Zur eindeutigen Kennzeichnung von Variablen werden ihnen Namen nach den oben beschriebenen Regeln gegeben (Länge: max. 8 Zeichen; 1. Zeichen: Buchstabe; Folgezeichen: Buchstaben oder Ziffern). Jede Variable erhält einen eindeutigen Namen.

Die Namen sollten so gewählt werden, daß sie den Inhalt der Variablen beschreiben, wie z.B.:

Identifikationsnummer des Fragebogens: *id*  
 Lebensalter: *alter*

##### 2.5.3.1 Variablentypen

In Abhängigkeit von den Datenwerten, die eine Variable annehmen kann, unterscheiden wir zwei benutzerdefinierbare Variablentypen:

- **Numerische Variablen**  
 Sie können nur Zahlen als Datenwerte haben.

- **Zeichenketten-Variablen**

(Auch alphanumerische Variablen oder String-Variablen genannt)

Ihnen können beliebige Zeichen in der Datendatei zugeordnet werden.

Der Variablentyp wird bei der Variablendefinition mit dem Befehl *DATA LIST* oder über das Daten-Editor-Menü festgelegt.

Numerische Variablen und Zeichenketten-Variablen werden unterschiedlich von SPSS behandelt. So können auf numerische Variablen Rechenoperationen angewendet werden, während Zeichenketten-Variablen nur ausgezählt und verglichen werden können.

SPSS definiert in Arbeitsdateien noch selbst einige Variablen, die sogenannten Systemvariablen, wie z.B. *\$casenum* (laufende Fallnummer), *\$date* (Datum) und *\$time* (Uhrzeit). Systemvariablen beginnen mit dem Dollarzeichen, sie können nicht in Statistikprozeduren, sondern nur in Datentransformationen angesprochen werden.

### 2.5.3.2 Variablengruppen

Mehrere aufeinanderfolgende Variablen können mit Hilfe des Schlüsselworts *TO* als Variablengruppe definiert werden. So sind die beiden folgenden Variablenangaben gleichwertig:

```
frau_1, frau_2, frau_3, frau_4, frau_5, frau_6
=
frau_1 TO frau_6
```

SPSS erzeugt aus einer Variablengruppe die einzelnen Variablennamen. Der erste Teil des Variablennamens ist ein allgemeingültiger Name, der zweite Teil besteht aus einer oder mehreren Ziffern. Der erste Teil ist für beide Variablen gleich, der Ziffernteil wird als Zahl entsprechend hochgezählt. Der Ziffernteil der ersten Variablen muß eine kleinere Zahl als der der zweiten ergeben. Insgesamt dürfen die sich ergebenden Variablennamen nicht länger als 8 Zeichen sein. Führende Nullen des Ziffernteils sind relevant:

```
frau_1 TO frau_6
=
frau_1, frau_2, frau_3, frau_4, frau_5, frau_6

frau_01 TO frau_06
=
frau_01, frau_02, frau_03, frau_04, frau_05, frau_06
```

### 2.5.3.3 Variablenlisten

Eine Variablenliste besteht aus einem oder mehreren Variablennamen, wobei auch Variablengruppen (TO-Regel) verwendet werden können. Variablenlisten werden überall dort benötigt, wo Variablen definiert oder angesprochen werden.

#### Definierende Variablenlisten

##### Beispiel

```
DATA LIST ...
... familie 6 heirat 7 frau_1 TO frau_6 9-14 ...
```

In diesem Beispiel werden Namen und Reihenfolge für folgende Variablen definiert: *familie*, *heirat*, *frau\_1*, *frau\_2*, *frau\_3*, *frau\_4*, *frau\_5*, *frau\_6*.

Diese Variablen können nun, nachdem sie mit *DATA LIST* definiert wurden, in Statistikprozedurbe-

fehlen und anderen Befehlen in beliebigen Kombinationen und Reihenfolgen verwendet werden.

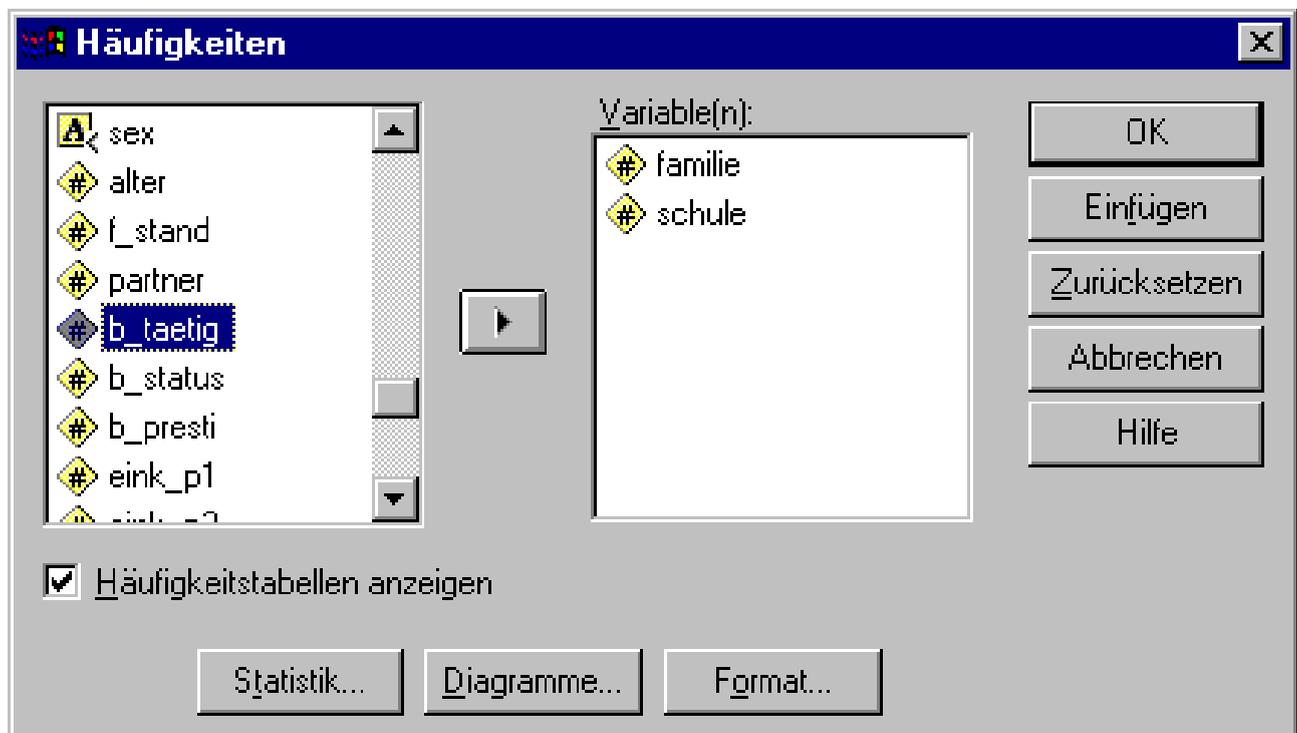
### Variablenlisten in Befehlen

Variablenliste	Angesprochene Variablen
<i>familie</i>	<i>familie</i>
<i>heirat frau_3</i>	<i>heirat, frau_3</i>
<i>frau_2 TO frau_5</i>	<i>frau_2, frau_3, frau_4, frau_5</i>
<i>frau_6 familie TO frau_3</i>	<i>frau_6, familie, heirat, frau_1, frau_2, frau_3</i>

Voraussetzung für die Verwendung von Variablengruppen nach der TO-Regel ist hier lediglich, daß die genannten Variablen hintereinander liegen (*familie TO frau\_3*).

### Variablenlisten in Dialogfeldern

Wird eine Statistik- oder Grafikprozedur über die Menüs aufgerufen, helfen Dialogfelder bei der Zusammenstellung der Variablenlisten.



Im linken Fenster der Dialogbox werden die Variablen der Arbeitsdatei (Quellvariablenliste) angezeigt, im rechten Fenster die Variablen, die in die Datenanalyse eingehen (Zielvariablenliste). Die Variablen werden von einem zum anderen Fenster verschoben, in dem man eine oder mehrere markiert (linke Maustaste festhalten und Mauszeiger ziehen) und dann den Pfeil in der Mitte anklickt. Einzelne Variablen können auch durch einen Doppelklick verschoben werden. Numerische Variablen sind durch das Nummernzeichen (#), alphanumerische durch "A" gekennzeichnet.

Im obigen Beispiel wird eine Häufigkeitsauszählung für die Variablen *familie* und *schule* vorgenommen:

►Analysieren ►Deskriptive Statistiken ►Häufigkeiten

Informationen zu einzelnen Variablen erhält man durch Anklicken mit der rechten Maustaste und

Auswahl von Info zu Variable im eingblendeten Popup-Kontextmenü:

The screenshot shows the SPSS 'Häufigkeiten' (Frequencies) dialog box for the variable 'familie'. The variable list on the left includes 'id', 'familie', 'heirat', and several 'frau\_' variables. The 'familie' variable is selected. The 'Variable(n):' field is empty. The 'Häufigkeitstabellen anzeigen' checkbox is checked. A context menu is open over the 'familie' variable, showing the question 'Braucht man Familie zum Glück?' and a list of value labels: '1 man braucht Familie', '2 ohne F. gleich glücklich', '3 allein glücklicher', '4 unentschieden', and '9 kA'. The '1 man braucht Familie' label is highlighted.

	frau_4	frau_5	frau_6
7	2	2	1
8	1	2	2
9	3	3	1

## 2.5.4 Grundlegende SPSS-Befehle

SPSS kennt verschiedene Arten von Befehlen, die unterschiedliche Funktionen erfüllen. Sie können in folgende Hauptgruppen unterteilt werden (*in Klammern die Befehle, die in dieser Einführung behandelt werden*):

- Hilfsbefehle  
(*COMMENT, EXECUTE*)
- Datenbeschreibung  
(*DATA LIST, MISSING VALUES, VARIABLE LEVEL, VARIABLE LABELS, VALUE LABELS*)
- Datentransformation  
(*TEMPORARY, RECODE, COMPUTE, IF, WEIGHT, COUNT*)
- Datenauswahl  
(*SELECT IF, SAMPLE*)
- SPSS-Datendateien  
(*SAVE, GET, SYSFILE INFO, DISPLAY, LIST, SORT CASES, SPLIT FILE, EXPORT, IMPORT*)
- Statistikprozeduren  
(*CLUSTER, CORRELATIONS, CROSSTABS, DESCRIPTIVES, EXAMINE, FACTOR, FREQUENCIES, NONPAR CORR, NPAR TESTS, ONEWAY, PARTIAL CORR, PROXIMITIES, REGRESSION, RELIABILITY, T-TEST*)
- Grafikprozedur  
(*GRAPH*)
- Tabellenprozedur  
(*TABLES*)

Die SPSS-Befehle können manuell im Syntaxfenster eingegeben oder dort mit Hilfe der Menüs zusammengestellt werden. In den folgenden Kapiteln werden beide Arten der Befehlseingabe beschrieben und an Hand von Beispielen - auf den vorliegenden Datensatz bezogen - erläutert. Bei der allgemeinen Beschreibung eines Befehls, dem Syntaxdiagramm, werden alle Unterbefehle, Schlüsselwörter und sonstigen Angaben aufgeführt, die für den jeweiligen Befehl gültig sind.

Beim Lesen eines Syntaxdiagramms ist folgendes zu beachten:

- Die Schlüsselwörter sind groß geschrieben; im SPSS-Programm können sie groß oder klein geschrieben werden.
- Klein geschriebene Teile müssen durch entsprechende eigene Angaben ersetzt werden.
- Der Ausdruck *varname* steht für eine Variable, *varlist* für eine Variablenliste.
- Die speziellen Trennzeichen (, ), /, ", ' und = sind dort erforderlich, wo sie in den Syntaxdiagrammen angegeben sind.
- Geschweifte Klammern „{}“ zeigen eine Wahlmöglichkeit zwischen mehreren Elementen an.
- Elemente in eckigen Klammern „[]“ sind optional.
- Voreinstellungen sind bei mehreren möglichen Angaben als erste in „[{}]“ aufgeführt.
- Mögliche Wiederholungen von vorangegangenen Teilen werden mit drei Punkten „(...)“ dargestellt.
- SPSS-Befehle müssen mit einem Punkt abgeschlossen werden. In den Syntaxdiagrammen ist der Punkt nicht angegeben.

Eine vollständige Dokumentation der SPSS-Befehlsprache finden Sie im „SPSS Base Syntax Reference Guide“, das auf CD-ROM mitgeliefert wird.

### 2.5.5 Reihenfolge der SPSS-Befehle

Die Befehle eines SPSS-Programms werden von SPSS in der Reihenfolge der Eingabe abgearbeitet. Einige Befehle müssen zwingend an einer bestimmten Stelle eingegeben werden, andere können beliebig positioniert werden. In jedem Fall gilt die Regel, daß Variablen zuerst definiert werden müssen, bevor sie von anderen Befehlen angesprochen werden können.

### 2.5.6 Hilfsbefehle

Die Befehle *COMMENT* und *EXECUTE* sind von allgemeiner Bedeutung für ein SPSS-Programm und werden daher vor den eigentlichen Daten- und Statistikbefehlen beschrieben.

#### 2.5.6.1 COMMENT

Ein SPSS-Programm kann zusätzlich zu den ausführbaren Befehlen auch Kommentare enthalten. Kommentare dienen der Erläuterung eines Programms. Sie haben keine Auswirkung auf das Programm, in der Ausgabe werden sie mit aufgelistet. Kommentare können an beliebiger Stelle eines Programms stehen.

Befehlssyntax für Comment

```
COMMENT text
* text
(Nur über die Syntax verfügbar.)
```

Kommentare beginnen mit *COMMENT* oder \* und müssen wie andere Befehle mit einem Punkt abschließen. Der angegebene *text* wird nicht in Apostroph oder Anführungszeichen eingeschlossen und kann auf beliebig vielen Zeilen fortgesetzt werden.

## Beispiel

```
* Analyse des ALLBUS-Datensatzes.
* Lebensalter: Häufigkeitsverteilung, Histogramm,
  statistische Kennwerte.
FREQUENCIES VARIABLES=alter
  /HISTOGRAM=NORMAL
  /STATISTICS=ALL.
* Lebensalter: Z-Transformation.
DESCRIPTIVES VARIABLES=alter /SAVE.
```

### 2.5.6.2 EXECUTE

Wie bereits unter „SPSS-Befehle ausführen“ beschrieben, überprüft SPSS zunächst nur die Syntax der Befehle. Die Daten werden erst beim Aufruf einer Statistikprozedur verarbeitet. Sollen z.B. Befehle zur Beschreibung, Transformation und Auswahl von Daten sofort auf die Arbeitsdatei bzw. die Daten wirken, muß diesen der Befehl *EXECUTE* folgen.

Befehlssyntax für Execute

```
EXECUTE
```

(Nur über die Syntax verfügbar.)

## Beispiel

```
* Analyse des ALLBUS-Datensatzes.
* Ausländerdiskriminierung: Index-Berechnung.
COMPUTE disindex = SUM.4(fremde_1 TO fremde_4)-4.
VARIABLE LABELS disindex 'Ausländerdiskriminierung'.
VALUE LABELS disindex
  0 'keine'
  24 'starke'.
EXECUTE.
```

Nach Ausführung dieser Befehlssequenz erscheint die Variable *disindex* mit ihren Werten im Datenfenster.

## 2.6 Datenbeschreibung und Datenerfassung

Die sorgfältige Beschreibung der Rohdatenmatrix mit den dafür in SPSS vorgesehenen Befehlen oder Menüs ist Voraussetzung für eine erfolgreiche Datenanalyse:

- Die Variablen erhalten eindeutige Namen.
- Der Variablentyp (numerisch oder Zeichenkette) wird festgelegt.
- Das Meßniveau der Variablen wird bestimmt.
- Den Variablennamen wird ein erläuternder Text (Variablenetikett) zugeordnet.
- Den Variablen werden Datenwerte zugeordnet.
- Den Datenwerten wird ein erläuternder Text (Wertetikett) zugeordnet.
- Fehlende Werte werden als solche gekennzeichnet.

SPSS speichert diese Informationen und die Rohdaten in der Arbeitsdatei („working file“), aus der eine SPSS-Datendatei erzeugt werden kann.

Wie bereits im Kapitel „Datendateien“ erwähnt, gibt es mehrere Möglichkeiten der Datenbeschreibung. Hier beschränken wir uns auf die Datenbeschreibung für Rohdaten, die bereits vorliegen, wie der ALLBUS-Datensatz, und die noch erfaßt werden müssen.

- **Die Daten liegen als festformatierte Rohdatendatei vor:**

- Variablendefinition und Einlesen der Rohdaten mit *DATA LIST*
- die weitere Datenbeschreibung mit den Befehlen *MISSING VALUES*, *VARIABLE LEVEL*, *VARIABLE LABELS* und *VALUE LABELS* oder
- menüunterstützt mit dem Daten-Editor
- **Die Daten sind noch nicht erfasst:**
  - Datenbeschreibung menüunterstützt mit dem Daten-Editor
  - Dateneingabe mit dem Daten-Editor

## 2.6.1 Befehle

Zum besseren Verständnis der Datenbeschreibungsmöglichkeiten werden hier die Befehle *DATA LIST*, *MISSING VALUES*, *VARIABLE LEVEL*, *VARIABLE LABELS* und *VALUE LABELS* vorgestellt und anschließend auf den ALLBUS-Datensatz angewendet. Auch für die menüunterstützte Datenbeschreibung mit dem Daten-Editor sind die im folgenden erläuterten Grundbegriffe eine wichtige Hilfe.

### 2.6.1.1 DATA LIST

*DATA LIST* definiert Variablen, legt ihren Typ fest und ordnet ihnen Werte in der Rohdatendatei zu.

Befehlssyntax für Data List (Auswahl!)

```
DATA LIST [FILE=datei] [{FIXED}] [RECORDS=1{ }]
           {FREE }           {n}
  /1 varlist {spaltenort [(format)] } [varlist ..]
  [/2{ } ...] [/ ...]
```

*FILE=datei* Rohdaten-Dateibezeichnung

vollständige Windows-Dateibezeichnung in Apostroph oder Anführungszeichen: Laufwerksbuchstabe, Verzeichnisname, Dateiname, Dateityp (\*.dat)

- Voreinstellung: keine Datei
- Beispiel: 'c:\allbus\allbus96.dat'

*FIXED* Rohdatenformat

Die Rohdaten sind in festem Format abgelegt, d.h. die Werte einer Variablen befinden sich für alle Fälle in den gleichen Spalten einer Datenzeile.

- *FREE*: freies Format.
- Voreinstellung: *FIXED*

*RECORDS=n* Anzahl der Datenzeilen pro Fall

- Voreinstellung: 1

/1 /2 Datenzeilen-Nummer

Es folgt die Definition der Variablen für die Datenzeilen.

- Voreinstellung: 1
- (Bei einer Datenzeile pro Fall kann die Angabe 1 entfallen.)

*varlist* Variablenliste

ein oder mehrere Variablennamen, auch TO-Variablengruppen

- Voreinstellung: keine (Angabe erforderlich)

*spaltenort* Spalten der Datenzeile

Spalten, in denen die Werte für die Variablen abgelegt sind, in der Form *anfangsspalte-endspalte*; bei einspaltigen Variablen in der Form *spalte*

- Voreinstellung: keine (Angabe erforderlich)

*(format)*      Daten- und Variablentyp

- Numerische Werte
  - ganze Zahlen: *(0)*
  - Dezimalzahlen mit Dezimalpunkt: *(0)*  
Der Zahlenwert wird wegen des gesetzten Dezimalpunkts immer richtig erkannt. Bei Angabe ( $n > 0$ ) werden  $n$  Dezimalstellen im Daten-Editor-Fenster angezeigt, ohne eine Angabe keine.
  - Dezimalzahlen ohne Dezimalpunkt: *(n)*  
Anzahl der nachträglich von SPSS zu definierenden Dezimalstellen. Diese Dezimalstellen werden auch im Daten-Editor-Fenster angezeigt.
- Zeichenketten: *(A)*
- Voreinstellung: *(0)*

SPSS kennt **numerische** und **Zeichenketten-Variablen**. Mit *(format)* wird der Variablentyp festgelegt, numerisch *(n)* oder Zeichenketten *(A)*. Weitere Formate wie z.B. Datums-, Währungs- und FORTRAN<sup>1</sup>-Formate sind im „SPSS Base Syntax Reference Guide“ beschrieben.

Variablen können als Variablengruppe zusammengefaßt werden, wenn der Wert jeder Variablen gleich viele Spalten belegt. Auch mehrspaltige Variablen können auf diese Weise mit *TO* definiert werden.

Ganzzahlige numerische Werte müssen rechtsbündig innerhalb ihrer Spalten in der Datendatei angeordnet sein, da Leerzeichen von SPSS als Nullen interpretiert werden. So würde z.B. das persönliche Nettoeinkommen in DM des ALLBUS-Datensatzes, das in den 5 Spalten 17-21 der zweiten Datenzeile abgelegt ist, wie folgt von SPSS umgesetzt:

Wert in Rohdatendatei	von SPSS gelesen als
624	624
9640	9640
345	3450
624	62400

### 2.6.1.2 MISSING VALUES

*MISSING VALUES* ordnet Variablen fehlende Werte zu. Sie können auch mit dem Daten-Editor definiert werden.

Befehlssyntax für Missing Values

```
MISSING VALUES
{varlist (werteliste) [[/]varlist ...]}
{ALL (wert) }
```

*varlist*      eine oder mehrere Variablen

*werteliste*   ein bis drei Werte, Wertebereich mit *THRU*, *ALL* oder kein Wert; ebenso die Schlüsselwörter *LO*, *LOWEST*, *HI*, *HIGHEST*

Fehlende Werte werden nur für die Variablen definiert, für die laut Kodierplan auch solche auftreten

<sup>1</sup> Programmiersprache zur Lösung numerischer Aufgaben.

können. Fehlende Werte bleiben bei statistischen Berechnungen unberücksichtigt, bei Häufigkeitsauszählungen werden sie entsprechend gekennzeichnet.

### Beispiel

```
MISSING VALUES
  familie (9)
  /heirat TO abbruch7 (8,9)
  /pzahl (99).
```

Wie im Kodierplan festgelegt, werden für die Variablen die entsprechenden fehlenden Werte definiert. Eine neue Variablenliste wird jeweils mit einem Schrägstrich von der vorhergehenden getrennt. Um das SPSS-Programm übersichtlich zu gestalten, wurde mit jeder Variablenliste eine neue Zeile begonnen. In TO-Gruppen muß die mit *DATA LIST* festgelegte Reihenfolge der Variablen eingehalten werden, ansonsten ist sie beliebig.

Schlüsselwörter für die Werteliste sind:

LO, LOWEST für den kleinsten auftretenden Wert

HI, HIGHEST für den größten auftretenden Wert

THRU für einen Wertebereich

Es können maximal drei Werte in einer Werteliste aufgeführt werden, die Anzahl der damit definierbaren fehlenden Werte ist jedoch beliebig:

```
MISSING VALUES wahl_1 (0,97 THRU 99).
```

Für die Variable *wahl\_1* werden 0, 97, 98 und 99 als fehlende Werte erklärt, insgesamt also vier Werte.

Falls die fehlenden Werte für alle Variablen gleich sind, kann anstelle einer Variablenliste das Schlüsselwort *ALL* verwendet werden:

```
MISSING VALUES ALL(9).
```

Damit gilt für alle Variablen 9 als fehlender Wert.

Bei einigen Auswertungen ist es erforderlich, die fehlenden Werte aufzuheben. Dazu bleibt die Werteliste leer:

```
MISSING VALUES ALL().
```

Bei Dezimalzahlen muß der Dezimalpunkt (nicht das Dezimalkomma) bei den fehlenden Werten mit angegeben werden:

```
MISSING VALUES b_presti(0, 99.9).
```

Für Zeichenketten-Variablen müssen die fehlenden Werte in Apostroph oder Anführungszeichen eingeschlossen werden:

```
MISSING VALUES sex('-').
```

### Automatisch definierte fehlende Werte

Zusätzlich zu den mit *MISSING VALUES* definierten fehlenden Werten - auch benutzerdefinierte fehlende Werte genannt - gibt es noch die von SPSS automatisch definierten fehlenden Werte („systemmissing values“).

Variablen erhalten diesen Wert immer dann von SPSS, wenn für sie kein gültiger Wert vorhanden ist. Das ist z.B. der Fall, wenn in der Rohdatendatei für eine numerische Variable eine ungültige Zahl - z.B. statt der Ziffer 0 der Buchstabe O geschrieben wurde - oder überhaupt keine Zahl abgelegt ist.

Die automatisch definierten fehlenden Werte können in einigen SPSS-Befehlen über das Schlüsselwort *SYSMIS* angesprochen werden. Nur numerische Variablen können automatisch definierte fehlende Werte haben, Zeichenketten-Variablen nicht. Der automatisch definierte fehlende Wert wird in der SPSS-Ausgabe, z.B. in Häufigkeitstabellen, unter den fehlenden Werten als Wert "System" aufgeführt.

### 2.6.1.3 VARIABLE LEVEL

*VARIABLE LEVEL* (Variablenmeßniveau) legt das Meßniveau („level“) einer Variablen fest. Es kann auch mit dem Daten-Editor definiert werden.

Befehlssyntax für Variable Level

```
VARIABLE LEVEL varlist ((SCALE ) ) [/varlist...]
                    {ORDINAL}
                    {NOMINAL}
```

<i>varlist</i>	eine oder mehrere Variablen
<i>SCALE</i>	Meßniveau: metrisch (Voreinstellung)
<i>ORDINAL</i>	Meßniveau: ordinal
<i>NOMINAL</i>	Meßniveau: nominal

Die Variablen, deren Meßniveau festgelegt wird, müssen vorher definiert worden sein. Das Meßniveau wird u. a. von den Diagrammprozeduren berücksichtigt.

#### Beispiel

```
VARIABLE LEVEL
  id TO heirat (NOMINAL)
  /frau_1 TO frau_6 (ORDINAL).
```

Die Variablen *id*, *familie*, *heirat* sind nominal, die Variablen *frau\_1* bis *frau\_6* ordinal skaliert.

### 2.6.1.4 VARIABLE LABELS

*VARIABLE LABELS* (Variablenetiketten) ordnet einer Variablen ein Etikett („label“) zu. Sie können auch mit dem Daten-Editor definiert werden.

Befehlssyntax für Variable Labels

```
VARIABLE LABELS varname 'label'
  [/varname...]
```

<i>varname</i>	ein Variablenname
<i>label</i>	Etikett: beliebige Zeichenkette in Apostroph oder Anführungszeichen (empfohlene Länge: max. 40 Zeichen)

Die Variablen, denen ein Etikett zugewiesen wird, müssen vorher definiert worden sein. Jeder Variablen kann nur ein Etikett zugewiesen werden; ein Etikett kann nur zu einer Variablen gehören. Die Variablenetiketten werden mit in die SPSS-Ausgabe geschrieben.

**Beispiel**

```
VARIABLE LABELS
  id      'Identifikationsnummer'
/familie 'Braucht man Familie zum Glück?'
/heirat  'Heirat bei dauerndem Zusammenleben?'
/frau_1  'Berufst. Frau: Herzl. Verhält. z. Kind'
/frau_2  'Frau: Lieber Mann bei Karriere helfen?'
/frau_3  'Frau: Nicht arbeiten bei Kleinkind?'
/frau_4  'Frau: Zuhause bleiben + Kinder versorgen?'
/frau_5  'Frau: Bessere Mutter bei Berufstätigkeit?'
/frau_6  'Frau: Nach Heirat Arbeitsplatz räumen?'
```

Den Variablen *id*, *familie*, *heirat* und *frau\_1* bis *frau\_6* wird je ein Etikett zugewiesen. Variablen, die zum gleichen Fragenkomplex gehören, werden sinnvollerweise Etiketten zugewiesen, aus denen dies ersichtlich ist ("*Frau: ...*").

**2.6.1.5 VALUE LABELS**

*VALUE LABELS* (Wertetiketten) ordnet Werten einer Variablen ein Etikett („label“) zu. Sie können auch mit dem Daten-Editor definiert werden.

Befehlssyntax für Value Labels

```
VALUE LABELS varlist
  wert 'label' wert 'label'...
[/varlist...]
```

*varlist*     eine oder mehrere Variablen

*wert*        ein Wert

*label*       Etikett: beliebige Zeichenkette in Apostroph oder Anführungszeichen (empfohlene Länge: max. 20 Zeichen)

Die Variablen, denen ein Etikett zugewiesen wird, müssen vorher definiert worden sein. Die Werte müssen nicht in den Daten vorkommen. Jedem Wert einer Variablen kann nur ein Etikett zugewiesen werden. Die Wertetiketten werden mit in die SPSS-Ausgabe geschrieben.

**Beispiel**

```
VALUE LABELS
  familie  1 'man braucht Familie'
           2 'ohne F. gleich glücklich'
           3 'allein glücklicher'
           4 'unentschieden'
           9 'kA'.
```

Nicht für alle Variablen werden Wertetiketten vergeben, wie z.B. für die Identifikationsnummer und für das Alter in Jahren. Fehlende Werte dieser Art von Variablen, wie z.B. das Alter mit dem Wert 999, sollten jedoch mit einem Etikett versehen werden.

Beim Einkommen in DM definiert man ebenfalls nur die Etiketten für die fehlenden Werte:

```
VALUE LABELS
  eink_p1  0 'kein Einkommen'
           99997 'vw'
           99998 'wn'
           99999 'kA'.
```

Die Werte des gruppierten Einkommens erhalten alle ein Etikett:

```

VALUE LABELS
  eink_p2      1 'unter 400 DM'
              2 '400 - 599 DM'
              3 '600 - 799 DM'
              .....
              21 '10000 - 14999 DM'
              22 '15000 DM und mehr'
              95 'Angabe schon da'
              96 'kein Einkommen'
              97 'vw'
              98 'wn'
              99 'kA'.

```

Auch werden nicht unbedingt alle Werte einer Variablen mit Wertetiketten versehen. Häufig reicht die Vergabe eines Etiketts für den kleinsten, den größten Wert und die fehlenden Werte:

```

VALUE LABELS
  fremde_1 TO fremde_4
              0 'tnz: keine dt. Staatsbürgerschaft'
              1 'stimme gar nicht zu'
              7 'stimme voll und ganz zu'
              9 'kA'.

```

Die Werte von Zeichenketten-Variablen müssen in Apostroph oder Anführungszeichen eingeschlossen werden:

```

VALUE LABELS
  sex  "m"  "männlich"
       "w"  "weiblich"
       "-"  "kA".

```

## 2.6.2 ALLBUS-Datensatz

Zunächst werden die Variablen mit *DATA LIST* definiert, anschließend die fehlenden Werte beschrieben, das Meßniveau der Variablen festgelegt und die Variablen- und Wertetiketten hinzugefügt.

### 2.6.2.1 DATA LIST

Wir definieren die Variablen des ALLBUS-Fragebogens an Hand des vorliegenden Kodierplans mit *DATA LIST* und fügen als zweiten Befehl  *FREQUENCIES* zum Auszählen der Variablen *familie* hinzu. Dazu rufen wir SPSS auf und öffnen ein neues Syntaxfenster über das Daten-Editor-Menü:

►Datei ►Neu ►Syntax.

In dieses Fenster schreiben wir die beiden Befehle und schließen sie jeweils mit einem Punkt ab:

The screenshot shows the SPSS Syntax Editor window titled "allbus96.sps - SPSS für Windows Syntax-Editor". The menu bar includes "Datei", "Bearbeiten", "Ansicht", "Analysieren", "Grafiken", "Extras", "Ausführen", "Fenster", and "Hilfe". The toolbar contains icons for file operations and execution. The main text area contains the following syntax:

```
DATA LIST FILE='c:\allbus\allbus96.dat' RECORDS=2
/1
  id 1-4 familie 6 heirat 7 frau_1 to frau_6 9-14
  anomia_1 to anomia_4 16-19 abbruch1 to abbruch7 21-27
  fremde_1 to fremde_4 29-32 kontakt1 to kontakt4 33-36
  pm_1 to pm_4 38-41 pm_index 42
  politik 44 wahl_1 45-46 wahl_2 47 wahl_3 48-49
  li_re 51-52 anteil 54 schicht 55
/2
  sex 1 (A) alter 2-4 f_stand 5 partner 6 schule 7
  b_taetig 9 b_status 10-11 b_presti 13-15 (1)
  eink_p1 17-21 eink_p2 22-23
  pzahl 25-26
  eink_hh1 28-32 eink_hh2 33-34
  religion 36 kirche 37
  ost_west 39 pass 40
  gewicht 42-47(5)
.
FREQUENCIES VARIABLES=familie.
```

The status bar at the bottom right indicates "SPSS für Windows Prozessor ist bereit".

Den Inhalt des Syntaxfensters speichern wir in einer Befehlsdatei ab:

- ▶Datei ▶Speichern unter
- ▶Speichern in: c:\allbus (Laufwerk und Verzeichnis auswählen)
- ▶Dateityp: SPSS-Syntaxdateien(\*.sps)
- ▶Dateiname: allbus96.sps

Die Daten der ersten drei Fragebögen sind (je Fragebogen zwei Datenzeilen):

```
m 1 11 242221 1211 1112128 21111111 21342 3 11 1 3 32
m 35105 161 329 400095 5 500095 14 11 1,1843
  2 13 111122 1211 1111111 53412111 32411 4 21 2 5 32
m 28512 162 428 230095 2 380095 25 11 1,1843
  3 12 132331 1121 1111111 41112111 21342 2 21 2 6 33
m 19522 171 344 120095 1 9999696 16 11 1,1843
```

Die Rohdaten sind in der Datei *FILE='c:\allbus\allbus96.dat'* abgelegt. Pro Fall gibt es zwei Datenzeilen (*RECORDS=2*). Die Daten sind festformatiert (*FIXED* als Voreinstellung).

Die erste Datenzeile (/1) enthält die Identifikationsnummer (*id*), die in den Spalten 1-4 abgespeichert ist. Es folgen die einspaltigen Variablen *familie*, *heirat* und die Variablengruppe *frau\_1 to frau\_6*, deren Werte in den Spalten 9-14 untergebracht sind. Der Wert jeder der Variablen *frau\_1* bis *frau\_6* belegt genau eine Spalte.

Alle folgenden Variablen, auch die für die zweite Datenzeile (/2), werden entsprechend definiert.

An dieser Stelle sei daran erinnert, daß ganzzahlige numerische Werte rechtsbündig innerhalb ihrer Spalten in der Datendatei angeordnet sein müssen, da Leerzeichen von SPSS als Nullen interpretiert werden.

In den Daten sind auch Dezimalzahlen enthalten:

- Ohne Dezimalkomma sind in der zweiten Datenzeile die Skalenwerte für das Berufsprestige angegeben: *b\_presti 13-15 (1)*. Das Dezimalkomma wird nachträglich beim Einlesen der Daten gesetzt, z.B. wird aus 329 die Dezimalzahl 32,9, die so im Daten-Editor-Fenster angezeigt wird. Die Zahlen müssen rechtsbündig in der Rohdatendatei angeordnet sein.
- Die Werte für die personenbezogene Ost-West-Gewichtung (*gewicht 42-47 (5)*) enthalten ein Dezimalkomma. Sie können daher beliebig in den Spalten 42-47 eingegeben werden. SPSS erkennt wegen des Dezimalkommas die richtigen Werte. Die Formatangabe (5) bewirkt lediglich die Anzeige der fünf Dezimalstellen im Daten-Editor-Fenster.

Bis auf die Variable für das Geschlecht haben wir es nur mit numerischen Variablen zu tun. Da deren Werte Buchstaben und Zeichen sind (*m, w, -*), wurde sie als Zeichenketten-Variable definiert: *sex 1 (A)*.

Unser erstes SPSS-Programm, das bis jetzt nur aus zwei Befehlen besteht, lassen wir von SPSS ausführen. Dazu klicken wir im Syntax-Editor-Fenster die Menüfolge ▶Ausführen ▶Alles an.

Die Rohdaten und Variableninformationen werden gelesen und in der temporären Arbeitsdatei abgelegt. Im Daten-Editor-Fenster erscheinen die Variablennamen und die Werte, im Ausgabefenster werden die ausgeführten Befehle aufgelistet.

Die Befehle werden nur dann im Ausgabefenster gelistet, wenn die von uns empfohlenen Einstellungen bei den  Programmoptionen vorgenommen wurden. Die Auflistung ist hilfreich bei der Fehlersuche, da sich eventuelle Fehlermeldungen auf den ausgeführten Befehl beziehen.

SPSS liest die Rohdaten erst dann ein, wenn sie für eine Statistikprozedur, wie hier für *FREQUENCIES*, benötigt werden. Datenfehler werden im „Log“ und unter „Warnungen“ gemeldet und beschrieben:

```
>Command line: 5 Current case: 966 Current splitfile group: 1
>Field contents: ' O '
>Record number: 1931 Starting column: 45 Record length: 8192
```

### Warnungen

An invalid numeric field has been found. The result has been set to the system-missing value.

Im ALLBUS-Datensatz ist für eine numerische Variable der Buchstabe „O“ anstelle einer Ziffer (vielleicht Null) eingegeben worden. SPSS kann keine Zahl bilden und verwendet statt dessen den automatisch definierten fehlenden Wert ( *MISSING VALUES*). Datenfehler und Möglichkeiten ihrer Bereinigung sind im Kapitel „Datenprüfung und Datenbereinigung“ beschrieben.

Da wir noch keine fehlenden Werte definiert haben, sind alle Werte gültig. Tatsächlich treten aber fehlende Werte auf (siehe den Wert „9“ in der folgenden Tabelle).

Zum Schluß wird die erwünschte Tabelle mit der Auszählung der Variablen *familie* zusammengestellt. Wir sehen die absoluten und die prozentualen Häufigkeiten der im ALLBUS-Datensatz vorkommenden Werte der Variablen *familie*. Eine ausführliche Beschreibung der Tabelleninhalte findet sich unter der Prozedurbeschreibung von *FREQUENCIES*.

		FAMILIE			
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	1	2570	73,1	73,1	73,1
	2	586	16,7	16,7	89,7
	3	75	2,1	2,1	91,8
	4	281	8,0	8,0	99,8
	9	6	,2	,2	100,0
<b>Gesamt</b>		3518	100,0	100,0	
<b>Gesamt</b>		3518	100,0		

Wir beenden SPSS über das Daten-Editor-Menü durch Anklicken des Symbols  oder mit **Datei ▶ Beenden**. SPSS fragt, ob wir den Inhalt der Fenster speichern möchten. Bis auf den Inhalt des Syntaxfensters, nämlich unser SPSS-Programm, speichern wir noch nichts ab. Bei Bedarf kann die Ausgabe oder eine Auswahl davon vor Beenden des Programms gedruckt werden.

### 2.6.2.2 Fehlende Werte, Meßniveau und Etiketten

Den mit *DATA LIST* definierten Variablen ordnen wir fehlende Werte zu, in dem wir diese mit *MISSING VALUES* definieren. Es folgen die Angaben zum Meßniveau der Variablen und die Etiketten für die Variablen und Werte. Die fehlenden Werte sind im Kodierplan angegeben, zur Formulierung der Variablen- und Wertetiketten werden die Texte des Fragebogens als Vorlage genommen.

Wir rufen SPSS auf, öffnen die bisher erstellte Befehlsdatei über das Daten-Editor-Menü und fügen die Befehle *MISSING VALUES*, *VARIABLE LEVEL*, *VARIABLE LABELS* und *VALUE LABELS* ein:

```

▶Datei ▶Öffnen
▶Suchen in: c:\allbus
▶Dateityp: Syntax(*.sps)
▶Dateiname: allbus96.sps

```

Das folgende Syntax-Fenster enthält alle Befehle für die Datenbeschreibung des ALLBUS-Datensatzes, die hier nur skizziert dargestellt sind. Das vollständige SPSS-Programm ist im Anhang abgedruckt.

```

allbus96.sps - SPSS für Windows Syntax-Editor
Datei Bearbeiten Ansicht Analysieren Grafiken Extras Ausführen Fenster Hilfe

DATA LIST FILE='c:\allbus\allbus96.dat' RECORDS=2
  /1
    id 1-4 familie 6 heirat 7 frau_1 to frau_6 9-14
  ...
  .
MISSING VALUES
  familie (9)
  ...
  .
VARIABLE LEVEL
  id TO heirat (NOMINAL)
  ...
  .
VARIABLE LABELS
  id      'Identifikationsnummer'
  ...
  .
VALUE LABELS
  familie      1 'man braucht Familie'
  ...
  .
FREQUENCIES VARIABLES=familie.
  
```

Den Inhalt des Syntaxfensters speichern wir wieder in der Befehlsdatei `allbus96.sps` ab und lassen das Programm von SPSS ausführen.

In der Häufigkeitstabelle der Variablen *familie* sind die fehlenden Werte (9) nun als solche gekennzeichnet und bei den prozentualen Häufigkeiten entsprechend berücksichtigt. Ebenso sind das Variablenetikett („Braucht man Familie zum Glück?“) und die Wertetiketten mit ausgegeben.

#### Braucht man Familie zum Glück?

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	man braucht Familie	2570	73,1	73,2	73,2
	ohne F. gleich glücklich	586	16,7	16,7	89,9
	allein glücklicher	75	2,1	2,1	92,0
	unentschieden	281	8,0	8,0	100,0
	Gesamt	3512	99,8	100,0	
Fehlend	kA	6	,2		
Gesamt		3518	100,0		

Die Datenbeschreibung ist nun vollständig, und wir sollten die im Daten-Editor-Fenster angezeigten Daten und die dazugehörigen Informationen (Variablennamen, fehlende Werte, Meßniveau, Variablen- und Wertetiketten), die sich z. Zt. in der temporären Arbeitsdatei befinden, in einer permanenten SPSS-Datendatei abspeichern:

- ▶Datei ▶Speichern unter
- ▶Speichern in: c:\allbus
- ▶Dateityp: SPSS(\*.sav)
- ▶Dateiname: allbus96.sav

In den folgenden Sitzungen brauchen wir dann nur noch die SPSS-Datendatei zu öffnen und können mit der eigentlichen Datenanalyse beginnen:

- ▶Datei ▶Öffnen
- ▶Suchen in: c:\allbus
- ▶Dateityp: SPSS(\*.sav)
- ▶Dateiname: allbus96.sav

### 2.6.3 Daten-Editor

Der Daten-Editor ermöglicht die menüunterstützte Datenbeschreibung, also die Vergabe der Variablennamen, die Festlegung des Typs und des Meßniveaus der Variablen, die Erklärung der fehlenden Werte und die Etikettierung der Variablen und Werte. Nach der Variablendefinition werden die Daten eingegeben, wobei der Daten-Editor auf das richtige Datenformat achtet. Die Arbeitsweise ähnelt der mit Tabellenkalkulations-Programmen.

Für bestehende Rohdatendateien, wie für den ALLBUS-Datensatz, kann die Datenbeschreibung nicht mit dem Daten-Editor durchgeführt werden. Zumindest müssen in diesem Fall mit *DATA LIST* die Namen und der Typ der Variablen definiert und die Rohdaten eingelesen werden. Die Informationen über die fehlenden Werte, das Meßniveau und die Etiketten könnten anschließend mit dem Daten-Editor hinzugefügt werden.

Um die Datenbeschreibung mit dem Daten-Editor an Hand eines konkreten Beispiels vorführen zu können, gehen wir davon aus, daß die Daten zum ALLBUS-Fragebogen noch nicht erfaßt wurden.

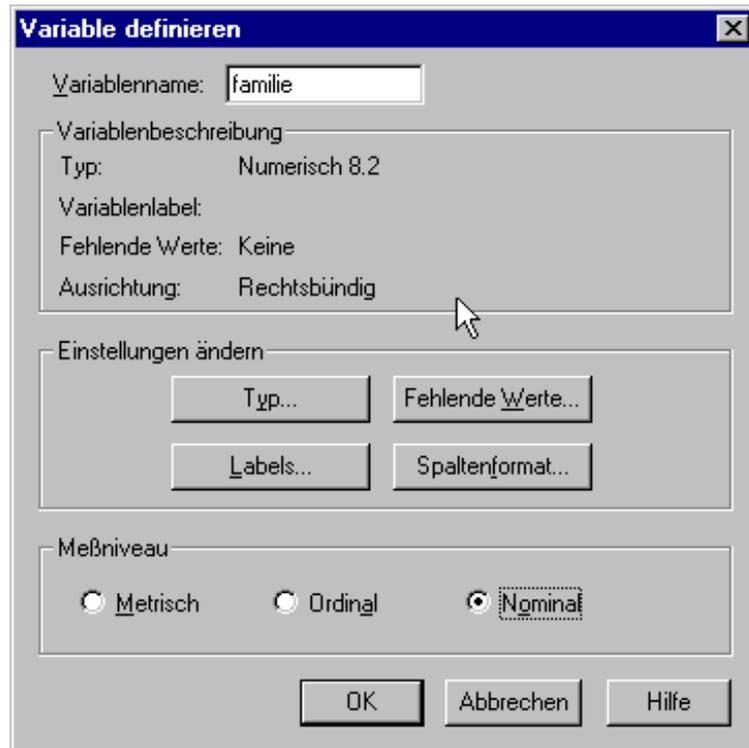
Grundlage für die Datenbeschreibung des ALLBUS-Fragebogens ist der vorliegende Kodierplan. Der Reihe nach müssen nun alle Variablen des Fragebogens definiert werden. Am Beispiel der numerischen Variablen *familie* führen wir die Datenbeschreibung mit dem Daten-Editor in SPSS 9 und SPSS 10 vor.

#### 2.6.3.1 Datenbeschreibung in SPSS 9

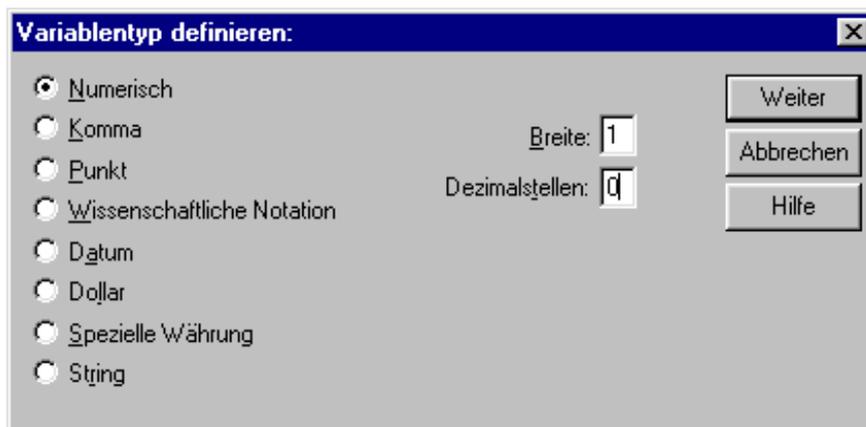
Wir rufen SPSS 9 (oder SPSS 8) auf und beginnen im leeren Daten-Editor-Fenster mit der Variablendefinition:



Stellvertretend für die zu vergebenen Variablenamen ist jede Datenspalte mit „var“ überschrieben. Nach einem Doppelklick mit der linken Maustaste oder einem einfachen Klick mit der rechten Maustaste und „Variable definieren“ auf die entsprechenden Namensfelder öffnet sich ein Fenster für die Variablenbeschreibung:



Den vorgeschlagenen Variablennamen *VAR0000X* überschreiben wir mit dem von uns vorgesehenen, also *familie*. Da diese Variable laut Kodierplan nur ganze, 1-stellige Zahlen annimmt, belassen wir den Variablentyp „Numerisch“. Das Meßniveau wird auf „Nominal“ gesetzt, die Anzahl der anzuzeigenden Ziffern (Breite) und Dezimalstellen wird angepaßt. Mit „Weiter“ wird der definierte Variablentyp übernommen:

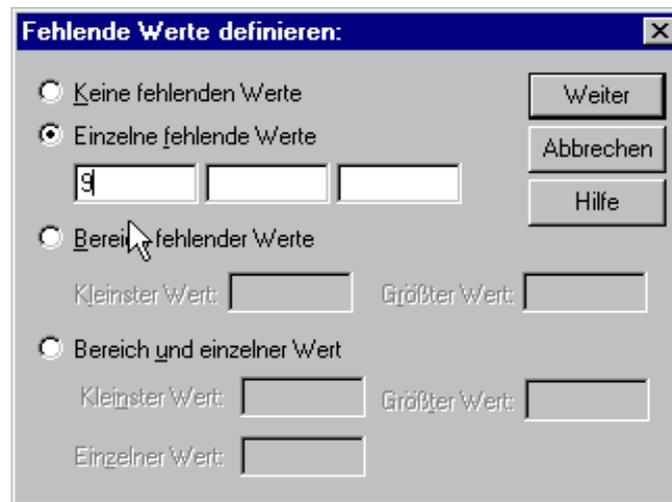


Der Variablentyp „Numerisch“ erlaubt die Eingabe beliebiger ganzer Zahlen und Dezimalzahlen mit Komma. Das Komma muß bei Dezimalzahlen immer mit angegeben werden. Die Angaben zu „Breite“ und „Dezimalstellen“ wirken sich nur auf die Darstellung im Datenfenster aus. Unabhängig von der angezeigten Zahl arbeitet SPSS immer mit dem exakten Zahlenwert. Wir hätten die vorgeschlagenen Einstellungen auch übernehmen können. Damit die Dateneingabe weniger fehleranfällig ist, empfehlen wir, die Formate an das tatsächliche Zahlenformat anzupassen.

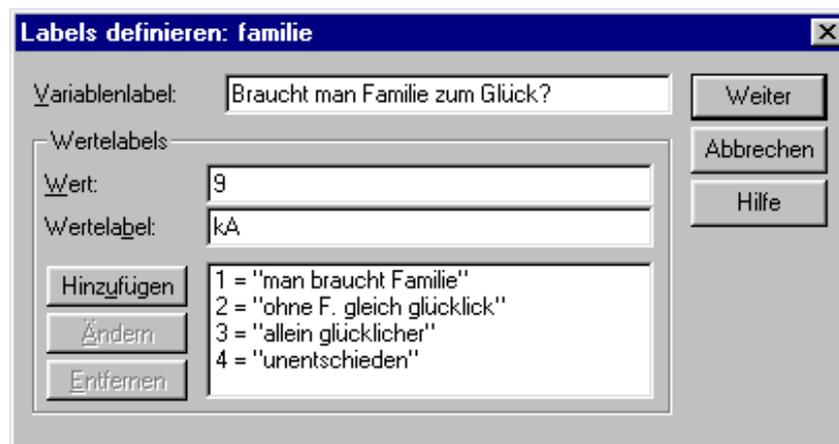
Für Zeichenketten-Variablen (z.B. *sex*) ist der Typ „String“ zu wählen.

Im nächsten Schritt werden der Variablen *familie* die fehlenden Werte zugeordnet. Wie bei dem Be-

fehl *MISSING VALUES* können bis zu drei einzelne Werte, ein Wertebereich (*THRU*) oder ein Wertebereich mit einem Einzelwert angegeben werden. Dezimalzahlen müssen anders als beim Befehl ein Komma statt eines Punktes enthalten.



Über „Labels“ gelangen wir zum Etiketten-Eingabefenster. Nach dem Variablenetikett werden die einzelnen Werte und deren Etiketten eingegeben. Jeder Wert und jedes Etikett wird einzeln über die Eingabemaske mit „Hinzufügen“ übernommen. Apostroph oder Anführungszeichen entfallen bei der Eingabe, SPSS setzt die Etiketten in Anführungszeichen:



Über die Schaltfläche „Spaltenformat“ können wir noch die „Spaltenbreite“ und die „Textausrichtung“ bestimmen. Zahlen werden immer rechtsbündig abgelegt. Die „Spaltenbreite“ ist unabhängig vom oben definierten Zahlenformat und kann auch nachträglich durch Ziehen des Spaltenrahmens verändert werden. Für die Anzeige der einstelligen Zahlen würde die vorgeschlagene Breite 1 ausreichen, nicht jedoch für den Variablennamen:



Die Variable *familie* ist vollständig beschrieben. Nach Verlassen des Fensters „Variable definieren“ mit „OK“ erscheint der vergebene Variablenname in der ersten Spalte des Daten-Editor-Fensters. Die Beschreibung einer Variablen kann jederzeit geändert bzw. ergänzt werden.

SPSS bildet die mit dem Datenditor zusammengestellten Variableninformationen auf die Datenbeschreibungsbefehle *FORMATS* (ähnlich *DATA LIST*), *MISSING VALUES*, *VARIABLE LEVEL*, *VARIABLE LABELS* und *VALUE LABELS* ab, führt sie aus und protokolliert sie im Ausgabefenster.

SPSS bietet Hilfen für die wiederholte Definition gleicher oder ähnlicher Variablen:

### Variablendefinitionen duplizieren

Mit den Funktionen „Kopieren“ und „Einfügen“ des Daten-Editor-Menüs „Bearbeiten“ können Variablendefinitionen bequem dupliziert werden:

- Variable (z.B. *frau\_1*) definieren
- Variable kopieren:
  - ▶ *frau\_1* mit linker Maustaste anklicken
  - ▶ Bearbeiten ▶ Kopieren
- Variablen einfügen:
  - 5 leere Variablenfelder für *frau\_2* bis *frau\_6* markieren (durch Anklicken und Ziehen mit der linken Maustaste)
  - ▶ Bearbeiten ▶ Einfügen
  - Variablennamen, Meßniveau und Variablenetiketten individuell vergeben

### Variablenvorlagen

Vorlagen ermöglichen ein ähnliches Vorgehen. Wir definieren eine für die Variablen *frau\_1* bis *frau\_6* passende Vorlage:

- ▶ Daten ▶ Vorlagen ▶ Definieren ▶ Name: *FRAU*  
 ▶ Zuweisen (alle Angaben) ▶ Vorlage definieren (wie Variable)  
 ▶ Hinzufügen ▶ OK
- 6 leere Variablenfelder markieren
- die Variablendefinition *FRAU* zuweisen: ▶ Daten ▶ Vorlagen
- Variablennamen, Meßniveau und Variablenetiketten individuell vergeben

### SPSS-Datendatei

Wenn alle Variablen definiert sind, speichern wir den Inhalt des Datenfensters - d.h. die Arbeitsdatei - als SPSS-Datendatei ab:

```
▶ Datei ▶ Speichern unter
▶ Speichern in: c:\allbus
▶ Dateityp: SPSS (*.sav)
▶ Dateiname: allbus96.sav
```

### 2.6.3.2 Datenbeschreibung in SPSS 10

Der Daten-Editor in SPSS 10 ermöglicht zwei Ansichten:

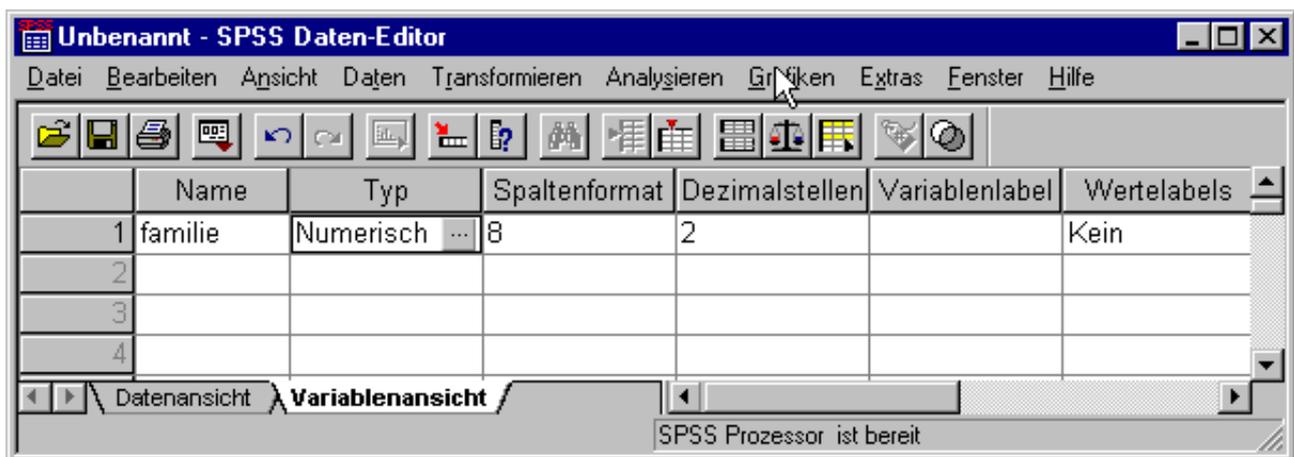
- die Datenansicht zum Anzeigen, Ändern und Eingeben der Daten,
- die Variablenansicht zum Anzeigen, Ändern und Definieren der Variablen, d.h. deren Namen und Attribute.

Wir rufen SPSS 10 auf und beginnen im leeren Daten-Editor-Fenster mit der Variablendefinition. Dazu wählen wir die Variablenansicht:



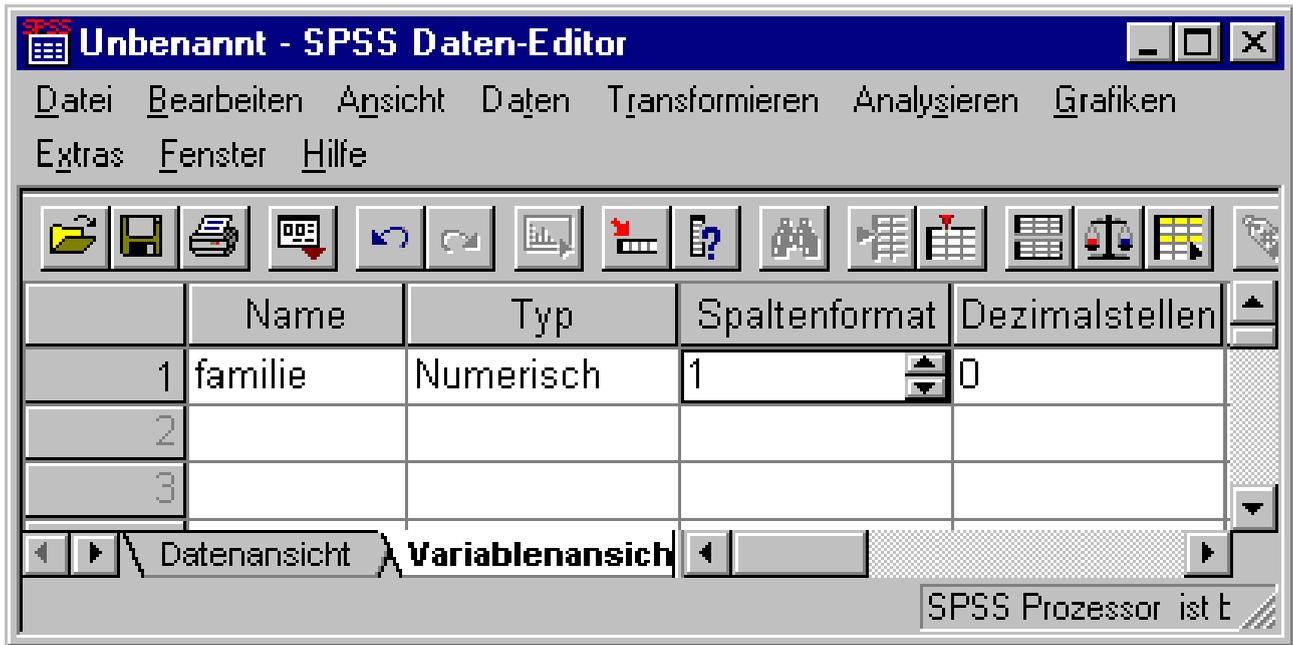
Die Variablen werden zeilenweise definiert. Nach dem Namen in der ersten Spalte folgen die Attribute der jeweiligen Variablen.

Nach Eingabe des von uns vorgesehenen Variablennamens *familie* setzt SPSS Standardwerte für deren Attribute:



Durch Anklicken der Punkte (Numerisch ...) bietet sich die Möglichkeit, andere Variablentypen zu definieren.

Da diese Variable laut Kodierplan nur ganze, 1-stellige Zahlen annimmt, belassen wir den Variablentyp „Numerisch“. Die Anzahl der anzuzeigenden Ziffern (Spaltenformat) und Dezimalstellen wird angepaßt:



Der Variablentyp „Numerisch“ erlaubt die Eingabe beliebiger ganzer Zahlen und Dezimalzahlen mit Komma. Das Komma muß bei Dezimalzahlen immer mit angegeben werden. Die Angaben zu „Spaltenformat“ und „Dezimalstellen“ wirken sich nur auf die Darstellung im Datenfenster aus. Unabhängig von der angezeigten Zahl arbeitet SPSS immer mit dem exakten Zahlenwert.

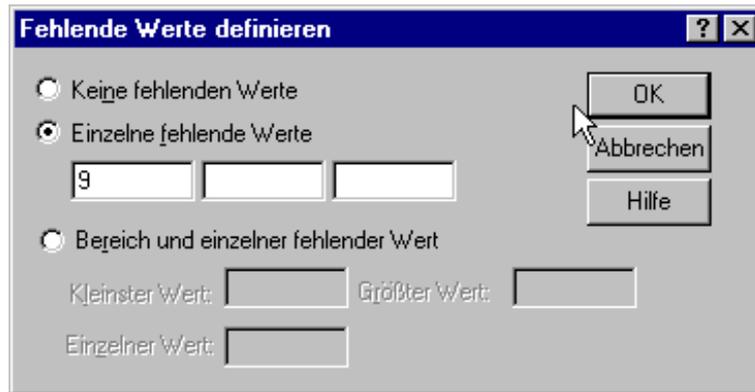
Wir hätten die vorgeschlagenen Einstellungen auch übernehmen können. Damit die Dateneingabe weniger fehleranfällig ist, empfehlen wir, die Formate an das tatsächliche Zahlenformat anzupassen.

Für Zeichenketten-Variablen (z.B. *sex*) ist der Typ „String“ zu wählen.

Nach dem Variablenlabel werden die Wertelabels eingegeben. Über die Punkte (Kein ...) gelangen wir zum Etiketten-Eingabefenster. Jeder Wert und jedes Etikett wird einzeln über die Eingabemaske mit „Hinzufügen“ übernommen. Apostroph oder Anführungszeichen entfallen bei der Eingabe, SPSS setzt die Etiketten in Anführungszeichen:

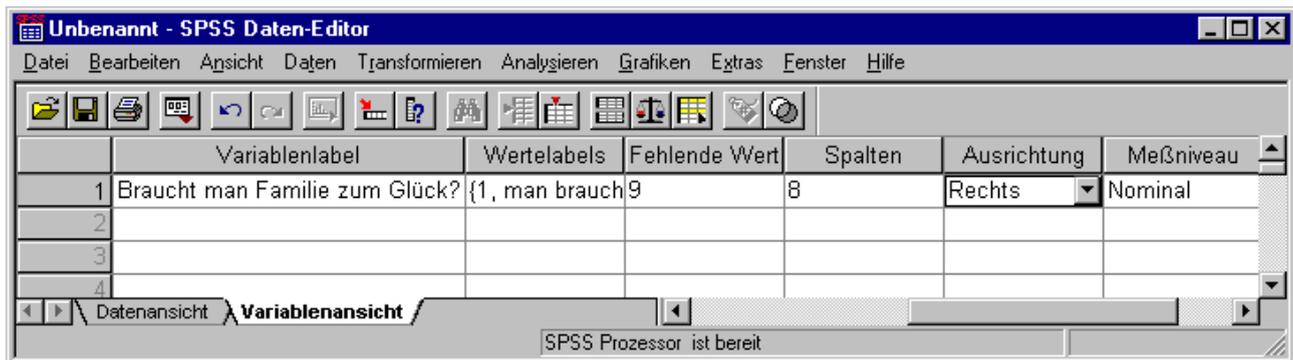


Im nächsten Schritt werden der Variablen *familie* die fehlenden Werte zugeordnet. Wie bei dem Befehl *MISSING VALUES* können bis zu drei einzelne Werte, ein Wertebereich (*THRU*) oder ein Wertebereich mit einem Einzelwert angegeben werden. Dezimalzahlen müssen anders als beim Befehl ein Komma statt eines Punktes enthalten:



In „Spalten“ legen wir die Spaltenbreite der Variablen *familie* für das Daten-Editor-Fenster fest. Die Spaltenbreite ist unabhängig vom oben definierten Spaltenformat und kann auch nachträglich durch Ziehen des Spaltenrahmens verändert werden. Wir übernehmen die vorgeschlagene Spaltenanzahl 8. Da Zahlen immer rechtsbündig abgelegt werden, belassen wir die „Ausrichtung“ mit „Rechts“.

Das „Meßniveau“ wird auf „Nominal“ gesetzt.



Die Variable *familie* ist somit vollständig beschrieben. Die Beschreibung einer Variablen kann jederzeit geändert bzw. ergänzt werden.

SPSS bietet Hilfen für die wiederholte Definition gleicher oder ähnlicher Variablen:

### Variablendefinitionen duplizieren

Mit den Funktionen „Kopieren“ und „Einfügen“ des Daten-Editor-Menüs „Bearbeiten“ oder des Kontextmenüs der linken Maustaste können Variablendefinitionen bequem in der „Variablenansicht“ dupliziert werden:

- Variable (z.B. *frau\_1*) definieren
- Variable kopieren:
  - ▶ Zeilennummer von *frau\_1* mit rechter Maustaste anklicken
  - ▶ Kopieren
- Variablen einfügen:
  - 5 leere Variablenzeilen für *frau\_2* bis *frau\_6* markieren

(durch Anklicken und Ziehen mit der rechten Maustaste)

- Einfügen
- Variablennamen und Variablenetiketten individuell vergeben

### SPSS-Datendatei

Wenn alle Variablen definiert sind, speichern wir den Inhalt des Datenfensters - d.h. die Arbeitsdatei - als SPSS-Datendatei ab:

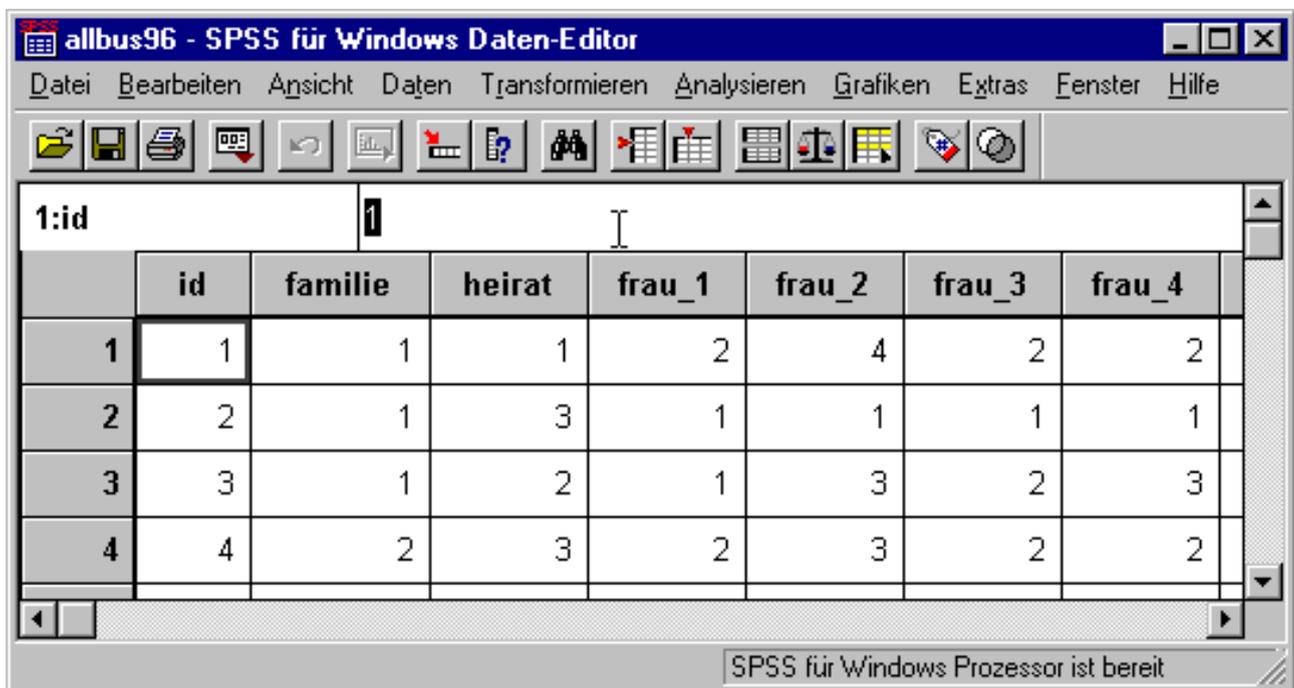
►Datei ►Speichern unter  
 ►Speichern in: c:\allbus  
 ►Dateityp: SPSS(\*.sav)  
 ►Dateiname: allbus96.sav

### 2.6.3.3 Dateneingabe

Wir öffnen die SPSS-Datendatei, die bisher nur die Variablenbeschreibung enthält, und beginnen mit der Dateneingabe:

►Datei ►Öffnen  
 ►Suchen in: c:\allbus  
 ►Dateityp: SPSS(\*.sav)  
 ►Dateiname: allbus96.sav

In SPSS 10 wählen wir die Datenansicht. Die Dateneingabe ist ähnlich der von Tabellenkalkulations-Programmen. Die Daten können in beliebiger Reihenfolge eingegeben werden. Für den ALLBUS-Fragebogen würden wir die fallweise Dateneingabe in der Reihenfolge der Fragebogennummern empfehlen.



The screenshot shows the SPSS Data Editor window titled 'allbus96 - SPSS für Windows Daten-Editor'. The menu bar includes 'Datei', 'Bearbeiten', 'Ansicht', 'Daten', 'Transformieren', 'Analysieren', 'Grafiken', 'Extras', 'Fenster', and 'Hilfe'. The toolbar contains various icons for file operations and data manipulation. The main window displays a data grid with the following content:

	id	familie	heirat	frau_1	frau_2	frau_3	frau_4
1	1	1	1	2	4	2	2
2	2	1	3	1	1	1	1
3	3	1	2	1	3	2	3
4	4	2	3	2	3	2	2

The status bar at the bottom indicates 'SPSS für Windows Prozessor ist bereit'.

Die Werte werden nicht direkt in die einzelnen Felder sondern erst in eine Zeile oberhalb der Variablennamen geschrieben:

- Gewünschtes Datenfeld anklicken.

- Fallnummer und Variablenname des Datenfeldes werden angezeigt.
- Wert eingeben und mit der Tabulatortaste in das Datenfeld einfügen.

Durch das Drücken der Tabulatortaste wird der Wert in das Datenfeld eingefügt und das nächste Feld der Zeile für die Eingabe markiert. Auch durch Drücken der Eingabetaste würde der Wert eingefügt, die Eingabemaske jedoch nach unten bewegt.

Bei der Dateneingabe achtet SPSS auf das richtige Eingabeformat:

- Für numerische Variablen können nur Ziffern eingegeben werden. Die Zahlen werden, wie vereinbart, rechtsbündig abgelegt.
- Für Zeichenketten-Variablen sind nicht mehr Zeichen erlaubt als vereinbart.

Nach Eingabe aller Daten erstellen wir wieder eine SPSS-Datendatei.

Der Daten-Editor bietet viele Funktionen zur Bearbeitung von Werten und zur Manipulation der Datendatei wie z.B. das Einfügen von Variablen und Fällen und das Sortieren der Fälle in Abhängigkeit von Variablen. Die hier verfügbaren Möglichkeiten der Datentransformation und -selektion werden in den entsprechenden Kapiteln dieses Buches gesondert beschrieben.

Erwähnenswert sind hier noch die Positionierungsfunktion:

►Daten ►Gehe zu Fall

und die Suchfunktion:

►Bearbeiten ►Suchen

## 2.7 Datentransformation

SPSS bietet Möglichkeiten, die Datenmatrix nachträglich zu ändern. So können Werte von bekannten Variablen umkodiert und neu berechnet werden; es können aber auch neue Variablen definiert und ihnen Werte zugewiesen werden. Auch die Gewichtung von Fällen ist möglich.

Datentransformationsbefehle werden erst mit der nächsten Statistikprozedur ausgeführt. Sollen die Transformationen ohne Statistikprozedur vorgenommen werden, muß der Befehl *EXECUTE* folgen.

Die Befehle zur Datentransformation können manuell oder über das Daten-Editor-Menü „Transformieren“ erstellt werden.

Es gibt permanente und temporäre Datentransformationen. Temporäre Transformationen bleiben nur bis zum nächsten Statistikprozedurbefehl erhalten, danach sind die ursprünglichen Variablen und Werte wieder in der Arbeitsdatei vorhanden. Transformationen sind temporär, wenn ihnen der *TEMPORARY*-Befehl vorangeht.

Befehlssyntax für Temporary

TEMPORARY

(Nur über die Syntax verfügbar.)

### Beispiel

```
TEMPORARY.  
RECODE b_taetig (1 THRU 3=1).  
FREQUENCIES VARIABLES=b_taetig.  
FREQUENCIES VARIABLES=b_taetig.
```

Der erste *FREQUENCIES*-Befehl erstellt eine Häufigkeitstabelle von der geänderten, der zweite von der ursprünglichen Variablen *b\_tae*tig.

Je nach Änderung der Werte und in Abhängigkeit von der Aufgabenstellung sind auch die Variablen- und Wertetiketten neu zu definieren. Werden fehlende Werte geändert, so verlieren sie die Eigenschaft „fehlend“ und müssen ebenfalls neu definiert werden. Alle Änderungen werden von SPSS in der Arbeitsdatei durchgeführt, die bei Bedarf als neue SPSS-Datendatei gespeichert werden kann.

### 2.7.1 RECODE

*RECODE* ändert die Werte bekannter Variablen. Der *TEMPORARY*-Befehl kann vorangehen.

Befehlssyntax für Recode (Auswahl!)

```
Bei numerischen Variablen:
RECODE varlist (werteliste=wert)
      [(werteliste=wert)...]
      [INTO varlist]
[/varlist...]

Bei String-Variablen:
RECODE varlist(wert,[wert...]=wert)
      [(wert,[wert...]=wert)]
      [INTO varlist]
[/varlist...]
```

<i>varlist</i>	eine oder mehrere bekannte Variablen
<i>werteliste</i>	ein bekannter oder mehrere bekannte Werte und Schlüsselwörter
<i>wert</i>	ein neuer Wert
<i>INTO varlist</i>	eine oder mehrere neue Variablen

Die Werte (*werteliste*) der angegebenen Variablen werden in den neuen Wert (*wert*) geändert. SPSS führt die Änderungen fallweise für den gesamten Datensatz durch. Die Wertelisten werden von links nach rechts abgearbeitet. Die Änderungen werden in der Arbeitsdatei gespeichert und stehen für darauffolgende statistische Auswertungen zur Verfügung.

Je nach Änderung der Werte und in Abhängigkeit von der Aufgabenstellung sind auch die Variablen- und Wertetiketten neu zu definieren. Werden fehlende Werte geändert, so verlieren sie diese Eigenschaft und müssen ebenfalls neu definiert werden.

Mit *INTO* können auch neue Variablen definiert und der Variablentyp geändert werden. So kann z.B. eine Zeichenketten-Variable in eine numerische Variable umgewandelt werden.

Schlüsselwörter für *RECODE*-Wertangaben sind:

LO, LOWEST	für den niedrigsten Wert
HI, HIGHEST	für den höchsten Wert
THRU	für einen Wertebereich
MISSING	für die benutzerdefinierten fehlenden Werte
SYSMIS	für die automatisch definierten fehlenden Werte
ELSE	für alle nicht angegebenen Werte

Für numerische Variablen sind alle Schlüsselwörter verwendbar, für Zeichenketten-Variablen nur *ELSE*.

### Beispiel

```
RECODE schicht (1,2=1) (3=2) (4,5=3).
VALUE LABELS  schicht
              1 "Unterschicht"
              2 "Mittelschicht"
              3 "Oberschicht"
              6 "keine der Schichten"
              7 "vw"
              8 "wn"
              9 "kA".
EXECUTE.
```

Die Werte der Variablen *schicht* werden wie folgt geändert:

```
1 → 1    2 → 1
3 → 2
4 → 3    5 → 3
6 = 6    7 = 7    8 = 8    9 = 9
```

Da sich die Bedeutung der Werte geändert hat, sind neue Werteetiketten erforderlich, der Wert 6 und die fehlenden Werte (7, 8, 9) bleiben unverändert.

### Beispiel

```
RECODE schicht (1 THRU 5=1) (6=2) (ELSE=9).
MISSING VALUES schicht (9).
VALUE LABELS  schicht
              1 "Schicht"
              2 "keine Schicht"
              9 "kA" .
EXECUTE.
```

Die Variable *schicht* hat nach dem *RECODE*-Befehl nur noch die Werte 1, 2 und 9. Die Eigenschaft „fehlend“ wird nicht auf den Wert 9 übertragen. Der fehlende Wert muß daher neu definiert werden. Werteetiketten wurden ebenfalls neu vergeben.

An Hand des obigen Beispiels zeigen wir, wie der *RECODE*-Befehl über das Daten-Editor-Menü erstellt und ausgeführt werden kann:

```
►Transformieren ►Umkodieren ►In dieselben Variablen
Variable schicht auswählen
Alte und neue Werte
Alter Wert ○Bereich: 1 bis 5
Neuer Wert ○Wert: 1
►Hinzufügen
Alter Wert ○Wert:6
Neuer Wert ○Wert: 2
►Hinzufügen
Alter Wert ○Alle anderen Werte
Neuer Wert ○Wert:9
►Hinzufügen
►Weiter
►Einfügen
```

Der *RECODE*-Befehl wird im Syntaxfenster zusammengestellt. Damit die Daten sofort rekodiert werden, fügt SPSS den Befehl *EXECUTE* hinzu:



Mit „Ausführen“ werden die beiden Befehle von SPSS abgearbeitet und die Werte der Variablen *schicht* rekodiert. Jetzt müssen noch der fehlende Wert und die Variablenetiketten neu definiert werden:

- ▶ Doppelklick auf den Variablennamen *schicht* im Daten-Editor-Fenster
- ▶ Fehlende Werte ☺ Fehlende Einzelwerte: 9 ▶ Weiter
- ▶ Labels ▶ Wertelabels:
  - alle bis auf 9 entfernen (▶ Anklicken ▶ Entfernen)
  - 1 und 2 neu definieren
- ▶ Weiter ▶ OK

Häufig ergibt es sich, daß eine kontinuierliche Variable mit ihren vielen Ausprägungen für eine bestimmte Aufgabenstellung nicht brauchbar ist.

```
RECODE alter (MISSING=9)
      (LO THRU 20=1) (20 THRU 30=2)
      (30 THRU 40=3) (40 THRU 50=4)
      (50 THRU 60=5) (60 THRU HI=6)
      INTO altgrup.
MISSING VALUES altgrup (9).
VARIABLE LABELS altgrup 'Alter in Gruppen'.
VALUE LABELS altgrup
  1 'bis 20 Jahre'
  2 '21 bis 30 Jahre'
  3 '31 bis 40 Jahre'
  4 '41 bis 50 Jahre'
  5 '51 bis 60 Jahre'
  6 'aelter 60 Jahre'
  9 'kA'.
FREQUENCIES VARIABLES=alter altgrup.
```

Die Werte der Variablen *alter* werden in Gruppen zusammengefaßt und die neuen Werte der Variablen *altgrup* zugewiesen. Die Variable *altgrup* wird neu definiert und an die bereits bekannten Variablen angehängt. Die fehlenden Werte 97 und 99 werden zu dem neuen fehlenden Wert 9 zusammengefaßt. Für *altgrup* werden der fehlende Wert und die Variablen- und Wertetiketten definiert, anschließend werden beide Variablen mit *FREQUENCIES* ausgezählt.

Einige Aufgabenstellungen erfordern ein Umordnen der Werte:

```
RECODE kontakt1 TO kontakt4 (2=1)(1=2).
VALUE LABELS kontakt1 TO kontakt4
  0 'tnz: keine dt. Staatsbürgerschaft'
  1 'nein'
  2 'ja'
  9 'kA' .
EXECUTE.
```

Die Werte 1 und 2 der Variablen *kontakt1* bis *kontakt4* werden zu 2 und 1 umgepolt, d.h. ja und nein werden vertauscht

Zeichenketten-Variablen können mit *INTO* in numerische Variablen umgewandelt werden:

```
RECODE sex ('m'=1) ('w'=2) ('-'=9) INTO sexn.
MISSING VALUES sexn (9).
VARIABLE LABELS sexn 'Geschlecht numerisch'.
VALUE LABELS sexn 1 'maennlich' 2 'weiblich'
              9 'keine Angabe' .
EXECUTE.
```

## 2.7.2 COMPUTE

*COMPUTE* berechnet Variablen. Der *TEMPORARY*-Befehl kann vorangehen.

Befehlssyntax für Compute

```
COMPUTE zielvariable = ausdruck
```

*zielvariable* eine bekannte oder neue Variable

*ausdruck* ein arithmetischer Ausdruck

*COMPUTE* weist der Zielvariablen den Wert des arithmetischen Ausdrucks zu. Ist die Zielvariable bekannt, so werden ihre Werte entsprechend geändert. Ist sie unbekannt, so wird sie zunächst definiert und nimmt dann die Werte des arithmetischen Ausdrucks an. Eine neue Variable wird in der Arbeitsdatei hinten an den bekannten Variablensatz angefügt.

Ein arithmetischer Ausdruck kann bekannte Variablen, Konstanten, arithmetische Operatoren, Funktionen und runde Klammern enthalten.

Das Ergebnis eines arithmetischen Ausdrucks wird auf den automatisch definierten fehlenden Wert (*SYSMIS*) gesetzt, falls

- eine der Variablen einen fehlenden Wert hat,
- eine arithmetische Operation einen nicht definierten Wert liefert, wie z.B. die Division durch Null,
- eine Funktion einen nicht definierten Wert liefert, wie z.B. die Quadratwurzel aus einer negativen Zahl.

In Abhängigkeit vom Inhalt der Zielvariablen können Variablen- und Wertetiketten folgen.

### Beispiel

```
COMPUTE f_idx=(fremde_1+fremde_2+fremde_3+fremde_4)/4.
VARIABLE LABELS f_idx 'Fremdenfeindlichkeit'.
VALUE LABELS f_idx 1 'keine'
               7 'starke'.
EXECUTE.
```

Die Werte der Variablen *fremde\_1* bis *fremde\_4* werden addiert. Damit das Ergebnis im gleichen Wertebereich (1 bis 7) wie die Werte der Ausgangsvariablen liegt, wird die Summe durch die Anzahl der addierten Variablen (4) geteilt. Für die richtige Abarbeitung des arithmetischen Ausdrucks sorgen runde Klammern. Der errechnete Wert wird der neuen Variablen *f\_idx* zugewiesen, für die anschließend ein Variablenetikett und Wertetiketten vergeben werden.

SPSS berechnet der Reihe nach, beginnend mit dem ersten Fragebogen, die Variable *f\_idx* für jeden Fragebogen (jeden Fall). Die neue Variable wird mit den berechneten Werten an die bestehende Da-

tenmatrix, die in der Arbeitsdatei abgespeichert ist, angehängt und ist für nachfolgende Auswertungen innerhalb desselben SPSS-Programms verfügbar.

Soll auf die Variable  $f\_idx$  in nachfolgenden SPSS-Programmläufen zugegriffen werden können, muß eine neue SPSS-Datendatei erzeugt werden.

Die Variable  $f\_idx$  erhält bei all den befragten Personen den automatisch definierten fehlenden Wert (*SYSMIS*), bei denen eine oder mehrere der Variablen  $fremde\_1$  bis  $fremde\_4$  einen fehlenden Wert (0 oder 9) haben.

Die Befehlssequenz des obigen Beispiels kann auch über das Daten-Editor-Menü durchgeführt werden:

►Transformieren ►Berechnen  
Zielvariable:  $f\_idx$   
►Typ und Label Label: *Fremdenfeindlichkeit* Numerisch ►Weiter  
Numerischer Ausdruck:  
*Benötigte Variablenamen und Symbole durch Klicken zusammenstellen.*  
►Einfügen



Die Variablenbeschreibung wird durch die Vergabe der Wertetiketten vervollständigt (Doppelklick auf  $f\_idx$ ).

### Arithmetische Operatoren

- + Addition
- Subtraktion
- \* Multiplikation
- / Division
- \*\* Exponentiation

### Funktionen

SPSS kennt eine Vielzahl von Funktionen, hier sollen nur die wichtigsten aufgeführt werden. Bei der Schreibweise ist zu beachten, daß die Angaben zu den Funktionen - Argumente genannt - in Klammern übergeben und durch Kommata voneinander getrennt werden.

Arithmetische Funktionen:

ABS (arg)	absoluter Wert
EXP (arg)	potenzieren zur Basis e ( $e^{arg}$ )
LG10 (arg)	dekadischer Logarithmus
LN (arg)	natürlicher Logarithmus

MOD( <i>arg1</i> , <i>arg2</i> )	Rest von <i>arg1</i> geteilt durch <i>arg2</i>
RND( <i>arg</i> )	kaufmännisch gerundeter Wert
SQRT( <i>arg</i> )	Quadratwurzel
TRUNC( <i>arg</i> )	ganzzahliger Teil des Wertes

Alle arithmetischen Funktionen, mit Ausnahme der *MOD*-Funktion, haben nur ein Argument.

### Beispiel

```
COMPUTE
  f_idx=RND((fremde_1+fremde_2+fremde_3+fremde_4)/4).
VARIABLE LABELS f_idx 'Fremdenfeindlichkeit'.
VALUE LABELS f_idx 1 'stark'
                7 'schwach'.
EXECUTE.
```

Die Werte der Variablen *f\_idx* werden zusätzlich mit der *RND*-Funktion zu ganzen Zahlen gerundet, es werden Variablen- und Wertetiketten vergeben.

### Statistische Funktionen:

SUM[.n] ( <i>arglist</i> )	Summe
MEAN[.n] ( <i>arglist</i> )	arithmetisches Mittel
SD[.n] ( <i>arglist</i> )	Standardabweichung
VARIANCE[.n] ( <i>arglist</i> )	Varianz

Die statistischen Funktionen können beliebig viele Argumente haben. Sind die Argumente Variablen, so können auch Variablengruppen (TO-Regel) angegeben werden.

*.n* gibt die Mindestanzahl der Argumente an, die pro Fall einen gültigen Wert haben müssen, damit die Funktion einen gültigen Wert liefert. Da die Voreinstellungen von *n* bei den einzelnen Funktionen unterschiedlich sind, sollte *n* immer angegeben werden.

### Beispiel

```
COMPUTE f_idx=RND(SUM.4(fremde_1 TO fremde_4)/4).
VARIABLE LABELS f_idx 'Fremdenfeindlichkeit'.
VALUE LABELS f_idx 1 'stark'
                7 'schwach'.
EXECUTE.
```

Es wird die Summenfunktion anstelle der arithmetischen Addition (+) verwendet. Der Variablen *f\_idx* wird auch hier nur dann ein gültiger Wert zugewiesen, falls alle vier Variablen *fremde\_1* bis *fremde\_4* eines Fragebogens einen gültigen Wert haben (*SUM.4*).

Zusätzlich zu den oben aufgeführten gibt es noch weitere statistische Funktionen, trigonometrische Funktionen (z.B. Sinus), Funktionen für fehlende Werte, fallübergreifende Funktionen, Zeichenkettenfunktionen sowie Datums- und Zeitfunktionen. Mit letzteren kann z.B. aus zwei Datumsangaben die Differenz in Tagen berechnet werden.

### Abarbeitungsreihenfolge

Arithmetische Ausdrücke werden von SPSS wie folgt abgearbeitet:

1. Klammern
2. Funktionen
3. \*\*
4. \*, /

5. +, -

Operatoren gleicher Priorität werden von links nach rechts abgearbeitet.

### 2.7.3 IF

*IF* berechnet Variablen in Abhängigkeit von einer Bedingung. Der *TEMPORARY*-Befehl kann vorgehen.

Befehlssyntax für If

```
IF [(logischer ausdruck)] zielvariable=ausdruck
```

- (logischer ausdruck)*      eine Bedingung
- zielvariable*                eine bekannte oder neue Variable
- ausdruck*                    ein arithmetischer Ausdruck

*IF* weist, wie *COMPUTE*, der Zielvariablen den Wert des arithmetischen Ausdrucks zu. Die Zuweisung erfolgt jedoch nur dann, wenn die mit einem logischen Ausdruck formulierte Bedingung erfüllt ist.

Ein logischer Ausdruck besteht meistens aus bekannten Variablen und relationalen und logischen Operatoren. Er kann zusätzlich - wie ein arithmetischer Ausdruck - Konstanten, arithmetische Operatoren, Funktionen und runde Klammern enthalten. Das Ergebnis eines logischen Ausdrucks ist wahr („true“) oder falsch („false“), d.h. die Bedingung ist erfüllt oder nicht erfüllt. Für die Wertzuweisung an die Zielvariable gilt daher:

Zielvariable	Ist die Bedingung erfüllt?	
	ja	nein
Bekannte Variable =	neuer Wert	alter Wert
Neue Variable =	Wert	<i>SYSMIS</i>

Für die Berechnung des arithmetischen Ausdrucks gelten die gleichen Regeln wie für *COMPUTE*, ebenso für die Zuweisung des automatisch definierten fehlenden Wertes (*SYSMIS*).

In Abhängigkeit vom Inhalt der Zielvariablen können Variablen- und Wertetiketten folgen.

#### Beispiel

```
IF (f_stand EQ 1) eink_hh1=eink_hh1+400.
EXECUTE.
```

Falls der Wert der Variablen *f\_stand* (Familienstand) gleich (*EQ*) 1 (verheiratet zusammenlebend) ist, wird zu dem alten Wert der Variablen *eink\_hh1* (Monatliches Haushalts-Nettoeinkommen in DM) der Betrag von 400 (DM) hinzuaddiert. Nach Abarbeitung des *IF*-Befehls haben alle verheiratet zusammenlebenden Befragten ein um 400 DM höheres Einkommen. Das Einkommen aller anderen Personen hat sich nicht geändert.

#### Relationale Operatoren (Vergleichsoperatoren)

Relationale Operatoren vergleichen Werte und liefern als Ergebnis die logischen Werte wahr („true“) oder falsch („false“). SPSS kennt folgende relationalen Operatoren (in Klammern die gleichwertigen Symbole):

EQ (=)	<i>Equal to</i>	wahr, falls linker Wert gleich rechtem Wert
NE (~=)	<i>Not Equal to</i>	wahr, falls linker Wert nicht gleich rechtem Wert
LT (<)	<i>Less Than</i>	wahr, falls linker Wert kleiner als rechter Wert
LE (<=)	<i>Less than or Equal to</i>	wahr, falls linker Wert kleiner oder gleich rechtem Wert
GT (>)	<i>Greater Than</i>	wahr, falls linker Wert größer als rechter Wert
GE (>=)	<i>Greater than or Equal to</i>	wahr, falls linker Wert größer oder gleich rechtem Wert

### Logische Operatoren

Die mit relationalen Operatoren formulierten Vergleiche können durch logische Operatoren miteinander verknüpft werden. Das Ergebnis ist wieder wahr oder falsch. Die logischen Operatoren sind:

AND (&)	<i>logisches Und</i>	wahr, falls linker und rechter Wert wahr sind
OR ( )	<i>logisches Oder</i>	wahr, falls linker, rechter oder beide Werte wahr sind
NOT (~)	<i>logisches Nicht</i>	kehrt einen logischen Wert um, von wahr nach falsch und umgekehrt

### Abarbeitungsreihenfolge

Logische Ausdrücke werden von SPSS wie folgt abgearbeitet:

1. Relationale Operatoren
2. *NOT*
3. *AND*
4. *OR*

Arithmetische Ausdrücke werden nach den unter *COMPUTE* beschriebenen Regeln vorher abgearbeitet.

Die relationalen und logischen Operatoren müssen durch mindestens ein Leerzeichen von den links- und rechtsstehenden Ausdrücken getrennt werden. Die Abarbeitungsfolge kann durch runde Klammern gesteuert werden.

### Beispiel

```
IF ((schule EQ 1 OR schule EQ 2) AND
    (eink_hh1 LE 3000))  osi=1.
IF ((schule EQ 3) AND
    (eink_hh1 GT 3000 AND eink_hh1 LE 4500))  osi=2.
IF ((schule EQ 4 OR schule EQ 5) AND
    (eink_hh1 GT 4500))  osi=3.
FREQUENCIES VARIABLES=osi.
```

Aus den Variablen *schule* (Allgemeiner Schulabschluß) und *eink\_hh1* (Monatliches Haushaltsnettoeinkommen) wird die neue Variable *osi* (Objektiver Schichtindex) berechnet und anschließend mit *FREQUENCIES* ausgezählt. Sie ist für alle befragten Personen definiert, die eine der drei mit *IF* formulierten Bedingungen erfüllen, für alle anderen nicht (*SYSMIS*).

SPSS arbeitet die *IF*-Befehle in der gegebenen Reihenfolge für jede befragte Person ab, d.h. nacheinander den ersten, den zweiten und den dritten Befehl für die erste Person, für die zweite Person usw.

Das Haushaltsnettoeinkommen wurde vorher mit

```
FREQUENCIES VARIABLES=eink_hh1/NTILES=3
```

in drei gleich große Bereiche bezüglich der Anzahl der Fälle (3-er-Quantile) aufgeteilt.

Die Formulierung von *IF*-Befehlen über das Daten-Editor-Menü ist ähnlich wie beim *COMPUTE*-Befehl. Nach „Transformieren“ und „Berechnen“ wird vor dem „Numerischen Ausdruck“ über

„Falls“ die Bedingung „zusammengeklickt“.

### 2.7.4 WEIGHT

*WEIGHT* gewichtet Fälle in Abhängigkeit von den Werten einer Variablen. Der *TEMPORARY*-Befehl wirkt hier nicht, statt dessen ist der Befehl „*WEIGHT OFF*“ zu geben.

Befehlssyntax für Weight

```
WEIGHT {BY varname}
       {OFF      }
```

*varname*      eine bekannte numerische Variable  
*OFF*            Gewichtung zurücknehmen

#### Beispiel

```
FREQUENCIES VARIABLES=sex.
DESCRIPTIVES VARIABLES=alter.
RECODE sex (MISSING=9) ('m'=1) ('w'=2) INTO sexn.
MISSING sexn (9).
WEIGHT BY sexn.
FREQUENCIES VARIABLES=sex.
DESCRIPTIVES VARIABLES=alter.
```

In diesem Beispiel werden zum Vergleich statistische Kennwerte der Variablen *alter* ohne und mit Gewichtung berechnet. Als Gewichtungsvariable wird die Geschlechtszugehörigkeit *sex*, die erst in die numerische Variable *sexn* umgewandelt werden muß, benutzt. Durch die Gewichtung hat sich die Anzahl der Frauen (*sexn*=2) verdoppelt. Zur Kontrolle werden die Häufigkeitstabellen von *sex* vor und nach der Gewichtung ausgegeben.

Der *WEIGHT*-Befehl kann auch über das Daten-Editor-Menü „Daten“ mit „Fälle gewichten“ erstellt werden.

### 2.7.5 COUNT

*COUNT* zählt Werte von Variablen. Der *TEMPORARY*-Befehl kann vorangehen.

Befehlssyntax für Count

```
COUNT varname=varlist(werteliste)
     [/varname=...]
```

*varname*      eine bekannte oder neue (Ziel-)Variable  
*varlist*      eine oder mehrere bekannte Variablen  
*werteliste*    ein Wert oder mehrere Werte und Schlüsselwörter

Hat eine der angegebenen Variablen einen der aufgeführten Werte, so wird die Zielvariable um eins hochgezählt. Die Zählung erfolgt fallweise. Wird in einem Fall kein Wert gefunden, so erhält die Zielvariable für diesen Fall den Wert Null.

Die Werteliste kann beliebig viele Werte, die durch Kommata voneinander getrennt werden, enthalten. Schlüsselwörter für die Werteliste sind:

LO, LOWEST    für den niedrigsten Wert  
HI, HIGHEST   für den höchsten Wert  
THRU          für einen Wertebereich  
MISSING      für die benutzer- und automatisch definierten fehlenden Werte  
SYSMIS       für die automatisch definierten fehlenden Werte

Diese Schlüsselwörter gelten nur für numerische Variablen, für Zeichenketten-Variablen sind keine vorgesehen.

**Beispiel**

```

COUNT abbruch = abbruch1 TO abbruch7 (1).
VARIABLE LABELS
  abbruch 'Abbruchsbejahungen pro Person'.
FREQUENCIES VARIABLES=abbruch.

```

Die Variable *abbruch* wird neu definiert und erhält für jede befragte Person als Wert die Anzahl der Abbruchsbejahungen. Die Variable wird etikettiert und mit *FREQUENCIES* ausgezählt.

Bei Bedarf können für die Zielvariable Variablen- und Wertetiketten vergeben werden.

Der *COUNT*-Befehl kann auch über das Daten-Editor-Menü „Transformieren“ mit „Zählen“ erstellt werden.

**2.8 Datenauswahl**

SPSS speichert die gesamte Datenmatrix in der Arbeitsdatei ab, d.h. es werden alle Fälle in die statistischen Berechnungen einbezogen. Für die Auswahl von Fällen aus der Arbeitsdatei bietet SPSS mehrere Möglichkeiten, von denen hier zwei beschrieben werden.

Eine Datenauswahl reduziert das Datenmaterial und damit die Rechenzeit. Für das Testen von umfangreichen Auswertungen ist daher - vor allem dann, wenn sie auf nicht sehr leistungsstarken Rechnern durchgeführt werden - eine Datenauswahl mit z.B. *SAMPLE* zu empfehlen. Für die endgültige Auswertung muß allerdings der gesamte Datensatz verwendet werden.

Die Datenauswahl kann wie die Datentransformationen permanent oder temporär durchgeführt werden. Für die temporäre Auswahl ist der *TEMPORARY*-Befehl voranzustellen.

Soll die Datenauswahl ohne Statistikprozedur vorgenommen werden, muß der Befehl *EXECUTE* folgen.

Die Befehle zur Datenauswahl können manuell oder über das Daten-Editor-Menü „Daten“ erstellt werden.

**2.8.1 SELECT IF**

*SELECT IF* wählt in Abhängigkeit von einer Bedingung Fälle aus dem Datensatz aus. Der *TEMPORARY*-Befehl kann vorangehen.

Befehlssyntax für Select If

```
SELECT IF [(logischer ausdruck)]
```

(*logischer ausdruck*) Bedingung

Die Bedingung wird als logischer Ausdruck formuliert, es gelten die gleichen Regeln wie beim *IF*-Befehl. Ist das Ergebnis des logischen Ausdrucks wahr, wird der Fall für die Bearbeitung ausgewählt, anderenfalls wird er in der Arbeitsdatei gelöscht. Für eine temporäre Auswahl muß der *TEMPORARY*-Befehl vorangehen.

**Beispiel**

```

SELECT IF (sex EQ 'w' AND alter GT 50).
FREQUENCIES VARIABLES=kirche.

```

In der Auswertung mit *FREQUENCIES* werden nur Frauen, die älter als 50 Jahre sind, untersucht.

Eine bedingte Datenauswahl kann auch über das Daten-Editor-Menü durchgeführt werden:

►Daten ►Fälle auswählen: Auswählen:  Falls Bedingung zu trifft ►Falls

Es wird auf diese Weise allerdings nicht der entsprechende *SELECT-IF*-Befehl erstellt, sondern eine Befehlssequenz, die mit einer Filtervariablen und dem Befehl *FILTER* (siehe „Hilfe“) arbeitet.

## 2.8.2 SAMPLE

*SAMPLE* wählt eine zufällige Anzahl von Fällen aus. Der *TEMPORARY*-Befehl kann vorangehen.

Befehlssyntax für Sample

```
SAMPLE {prozentsatz}
      {n FROM m }
```

*prozentsatz*    prozentualer Anteil auszuwählender Fälle (0.0 - 1.0)

*n FROM m*      Auswahl von genau *n* aus den ersten *m* Fällen

### Beispiel

```
SAMPLE 0.1 .
FREQUENCIES VARIABLES=frau_1 TO frau_6.
```

Nur ca. 10% der Fälle bleiben in der Arbeitsdatei erhalten und gehen in die Häufigkeitsauszählung ein.

Eine zufällige Datenauswahl über das Daten-Editor-Menü erfolgt mit:

►Daten ►Fälle auswählen: Auswählen:  Zufallsstichprobe ►Stichprobe

Es wird eine *FILTER*-Befehlssequenz zusammengestellt.

## 2.9 SPSS-Datendateien

In Kapitel 2.4 („Datendateien“) wurde kurz beschrieben, wie SPSS-Datendateien über das Menü gelesen und geschrieben werden können. Hier werden nun ausführlich ihr Inhalt und Format sowie die Befehle zum Lesen, Schreiben und Manipulieren von SPSS-Datendateien besprochen. Es wird auch gezeigt, wie man Informationen über den Inhalt der Dateien abrufen und sich die Daten auflisten lassen kann.

SPSS legt nach dem Start eine temporäre **Arbeitsdatei** („working file“) an, in der die Daten, die Variablenamen und die Variableninformationen gehalten werden. Die Arbeitsdatei wird bei Sitzungsende wieder gelöscht, wenn sie nicht vorher als permanente SPSS-Datendatei („SPSS data file“) gespeichert wurde.

Sind die Rohdaten vollständig erfaßt und die Variablen richtig definiert und beschrieben, ist es sinnvoll, möglichst bald, d.h. nach einer ersten inhaltlichen Datenprüfung, eine SPSS-Datendatei zu erzeugen. In der sich anschließenden Datenanalyse hat man dann einen schnellen und bequemen Zugriff auf seine Variablen und deren Werte, nachdem SPSS diese Datendatei gelesen und daraus wieder eine Arbeitsdatei generiert hat.

Eine SPSS-Datendatei enthält zusätzlich zu den Rohdaten folgende Informationen:

- die Variablenamen
- den Variablentyp
- das Meßniveau der Variablen

- die für eine Variable möglichen fehlenden Werte
- die Variablen- und Werteetiketten
- das Datum und die Uhrzeit der Dateierstellung
- die Gesamtanzahl der Fälle und Variablen

Diese Informationen werden im Datenlexikon („dictionary“) einer SPSS-Datendatei zusammengefaßt. Wurden vor dem Erstellen einer SPSS-Datendatei permanente Datentransformationen, Datenselektionen oder eine Gewichtung durchgeführt, so werden die geänderten und neuen Variablen mit ihren Werten in die Datei aufgenommen.

SPSS-Datendateien enthalten **nicht** die Datenbeschreibungsbefehle wie *DATA LIST* und *MISSING VALUES* selbst, sondern nur die damit vermittelten Informationen. Auch werden **keine** Tabellen, Grafiken und statistischen Kennwerte darin abgespeichert.

SPSS-Datendateien (\* .sav) sind im Gegensatz zu Rohdatendateien (\* .dat), die als Textdateien angelegt werden, Binärdateien. Diese können nur von SPSS, und zwar nur unter dem gleichen Betriebssystem, gelesen werden. SPSS für Windows z.B. „versteht“ keine SPSS-Datendateien, die von SPSS für UNIX erzeugt wurden. Für den Austausch von SPSS-Datendateien zwischen verschiedenen Rechnerplattformen verwenden wir  portable SPSS-Datendateien.

Auf keinen Fall sollten SPSS-Datendateien mit einem Text-Editor oder Textverarbeitungsprogramm angesehen werden; der Dateiinhalt würde beim Zurückspeichern der Datei zerstört werden. Auch sei davor gewarnt, solche Binärdateien zu drucken, da ein Papierstapel unleserlicher Hieroglyphen produziert werden würde.

### 2.9.1 SPSS-Datendateien schreiben (SAVE)

*SAVE* erzeugt aus einer temporären Arbeitsdatei eine permanente SPSS-Datendatei.

Befehlssyntax für Save (Auswahl!)

```
SAVE OUTFILE=datei
  [/KEEP={ALL    }] [/DROP=varlist]
    {varlist}
  [/RENAME=(alte varlist=neue varlist)...]
  [/MAP] [/{COMPRESSED  }]
    {UNCOMPRESSED}
```

*OUTFILE=datei* SPSS-Datendatei-Bezeichnung

vollständige Windows-Dateibezeichnung in Apostroph oder Anführungszeichen: Laufwerksbuchstabe, Verzeichnisname, Dateiname, Dateityp (\* .sav)

- Voreinstellung: keine Datei
- Beispiel: "c:\allbus\allbus96.sav"

*KEEP=varlist*

Nur die hier aufgeführten Variablen werden in die SPSS-Datendatei übernommen.

- Voreinstellung: *ALL*, d.h. alle Variablen

*DROP=varlist*

Die hier aufgeführten Variablen werden nicht übernommen.

*RENAME(alte varlist=neue varlist)*

Die Variablen erhalten in der SPSS-Datendatei einen neuen Namen.

*MAP*

Die Namen der übernommenen Variablen werden im Ausgabefenster aufgelistet.

*COMPRESSED* | *UNCOMPRESSED*

Die SPSS-Datendatei wird komprimiert bzw. nicht komprimiert.

- Voreinstellung: *COMPRESSED*

Statt mit dem *SAVE*-Befehl kann über das Datei-Menü des Daten-Editors eine permanente SPSS-Datendatei erzeugt werden:

►Datei ►Speichern unter ►Dateityp: SPSS (\*.sav)

### 2.9.2 SPSS-Datendateien lesen (GET)

*GET* liest eine permanente SPSS-Datendatei und stellt eine temporäre Arbeitsdatei bereit.

Befehlssyntax für *Get*

```
GET FILE=datei
  [/KEEP={ALL    }] [/DROP=varlist]
    {varlist}
  [/RENAME=(alte varlist=neue varlist)...]
  [/MAP]
```

*FILE=datei*      SPSS-Datendatei-Bezeichnung

vollständige Windows-Dateibezeichnung in Apostroph oder Anführungszeichen: Laufwerksbuchstabe, Verzeichnisname, Dateiname, Dateityp (\*.sav);

- Voreinstellung: keine Datei
- Beispiel: "c:\allbus\allbus96.sav"

*KEEP=varlist*

Nur die hier aufgeführten Variablen werden in die Arbeitsdatei übernommen.

- Voreinstellung: *ALL*, d.h. alle Variablen

*DROP=varlist*

Die hier aufgeführten Variablen werden nicht übernommen.

*RENAME(alte varlist=neue varlist)*

Die Variablen erhalten in der Arbeitsdatei einen neuen Namen.

*MAP*

Die Namen der übernommenen Variablen werden im Ausgabefenster aufgelistet.

Eine permanente SPSS-Datendatei wird über das Datei-Menü des Daten-Editors zu einer temporären Arbeitsdatei mit:

►Datei ►Öffnen ►Dateityp: SPSS (\*.sav)

### 2.9.3 Variableninformationen (SYSDATA LIST, DISPLAY)

Die im Datenlexikon einer SPSS-Datendatei zusammengefaßten Variableninformationen sind über Befehle und Menüs abrufbar.

1. Informationen über eine nicht geöffnete SPSS-Datendatei

Auf diese Weise erhält man eine vollständige Auflistung aller Datei- und Variableninformationen.

a) Befehl: *SYSDATA LIST*

Befehlssyntax für Sysfile Info

```
SYSFILE INFO [FILE=] datei
```

*FILE=datei* SPSS-Datendateibezeichnung

## b) Daten-Editor-Menü

►Datei ►Info über Datendatei ►Dateityp: SPSS(\*.sav) ►Dateiname:

## 2. Informationen über eine geöffnete SPSS-Datendatei (Arbeitsdatei)

Es können Informationen über einzelne Variablen und über die gesamte Arbeitsdatei abgerufen werden.

### a) Befehl: *DISPLAY*

Befehlssyntax für Display (Auswahl!)

```
DISPLAY [{DICTIONARY }] [/VARIABLES=varlist]
        {VARIABLES }
```

- Voreinstellung            Auflistung der Variablennamen
- DICTIONARY*                Auflistung des gesamten Datenlexikons
- VARIABLES*                 Tabellarische Variablenliste
- VARIABLES=varlist*        Variablenauswahl

## b) Daten-Editor-Menü

Informationen über einzelne Variablen:

►Extras ►Variablen

Informationen über die gesamte Arbeitsdatei:

►Extras ►Datei-Info

## 2.9.4 Daten auflisten (LIST)

Die Daten einer Arbeitsdatei werden im Daten-Editor-Fenster angezeigt. Möchte man nur die Werte einzelner Variablen und Fälle auflisten, so empfiehlt sich der SPSS-Befehl *LIST*, der die ausgewählten Daten im Ausgabefenster auflistet.

Befehlssyntax für List (Auswahl!)

```
LIST [[VARIABLES=]{ALL** } ]
        {varlist}
[/FORMAT=[{UNNUMBERED**}]
        {NUMBERED } ]
[/CASES=FROM n TO n ]
```

(Nur über die Syntax verfügbar.)

- Voreinstellung            Auflistung der Werte aller Variablen und Fälle ohne Fallnummer
- VARIABLES=varlist*        Variablenauswahl
- FORMAT=NUMBERED*        interne Fallnummer (*\$CASENUM*) wird ausgegeben
- CASES=FROM n TO n*        Fallauswahl

## 2.9.5 SPSS-Datendateien manipulieren

SPSS bietet mehrere Möglichkeiten zur Manipulation einer Arbeitsdatei. Die Daten können z.B. sortiert und zusammengefaßt werden, die Arbeitsdatei kann in Unterdateien aufgeteilt und um Variablen und Fälle erweitert werden. Auch das Transponieren von Zeilen und Spalten der Arbeitsdatei, also das Vertauschen von Fällen und Variablen, ist möglich. Wir beschreiben hier das Sortieren und Aufteilen einer Arbeitsdatei.

### 2.9.5.1 SPSS-Datendateien sortieren

Einige Prozeduren wie z.B. *AGGREGATE* und *SPLIT FILES* setzen voraus, daß die Daten sortiert nach den Werten einer oder mehrerer Variablen vorliegen. *SORT CASES* sortiert eine Arbeitsdatei.

Befehlssyntax für Sort Cases

```
SORT CASES [BY] varlist[({A})]
                {D}
```

*varlist* eine oder mehrere Variablen nach deren Werten die Arbeitsdatei aufsteigend (*A*) oder absteigend (*D*) sortiert werden soll

Menü-Aufruf:

►Daten ►Fälle sortieren

### 2.9.5.2 SPSS-Datendateien aufteilen

Für einige Analysen ist die Aufteilung der Daten in mehrere Gruppen erforderlich. *SPLIT FILE* kann eine Arbeitsdatei so aufteilen, daß die Ergebnisse für die erzeugten Gruppen (z.B. Häufigkeitstabellen) separat oder für Vergleichszwecke zusammen angezeigt werden. Die Arbeitsdatei muß vorher nach der oder den Variablen, nach deren Werten die Aufteilung vorgenommen wird, sortiert werden.

Befehlssyntax für Split File

```
SPLIT FILE {OFF
            {[{LAYERED}}[BY varlist]}
            {SEPARAT}
```

*varlist* eine oder mehrere Variablen

*LAYERED* Gruppenvergleich: Ergebnisse werden zusammen angezeigt

*SEPARAT* getrennte Anzeige der Ergebnisse

*OFF* Aufteilung zurücknehmen

Menü-Aufruf:

►Daten ►Datei aufteilen

Der Menü-Aufruf enthält den nötigen Sortierbefehl.

## 2.9.6 Portable SPSS-Datendateien

Die Export-/Import-Funktion ermöglicht den Austausch von portablen SPSS-Datendateien zwischen verschiedenen Rechnerplattformen, z.B. zwischen Windows- und UNIX-Rechnern. Portable SPSS-Datendateien enthalten im Gegensatz zu Rohdatendateien alle Informationen einer SPSS-Datendatei, also sowohl Datenwerte als auch Variablennamen und Variableninformationen. Portable SPSS-Datendateien sind Textdateien, die problemlos z.B. auf Disketten, CD-ROM und per Electronic Mail weitergegeben werden können.

### 2.9.6.1 Portable SPSS-Datendateien schreiben (EXPORT)

*EXPORT* erzeugt aus einer Arbeitsdatei eine portable SPSS-Datendatei.

Befehlssyntax für Export (Auswahl!)

```
EXPORT OUTFILE=datei
[/KEEP={ALL    }] [/DROP=varlist]
    {varlist}
[/RENAME=(alte varlist=neue varlist)...]
[/MAP]
```

*OUTFILE=datei*    Bezeichnung der portablen Datei

vollständige Windows-Dateibezeichnung in Apostroph oder Anführungszeichen: Laufwerksbuchstabe, Verzeichnisname, Dateiname, Dateityp (\* .por)

- Voreinstellung: keine Datei
- Beispiel: "c:\allbus\allbus96.por"

*KEEP=varlist*

Nur die hier aufgeführten Variablen werden in die portable SPSS-Datendatei übernommen.

- Voreinstellung: *ALL*, d.h. alle Variablen

*DROP=varlist*

Die hier aufgeführten Variablen werden nicht übernommen.

*RENAME(alte varlist=neue varlist)*

Die Variablen erhalten in der portablen SPSS-Datendatei einen neuen Namen.

*MAP*

Die Namen der übernommenen Variablen werden im Ausgabefenster aufgelistet.

Statt mit dem *EXPORT*-Befehl kann über das Datei-Menü des Daten-Editors eine portable SPSS-Datendatei erzeugt werden:

►Datei ►Speichern unter ►Dateityp: SPSS portabel (\*.por)

### 2.9.6.2 Portable SPSS-Datendateien lesen (IMPORT)

*IMPORT* liest eine portable SPSS-Datendatei und stellt eine Arbeitsdatei bereit.

Befehlssyntax für Import (Auswahl!)

```
IMPORT FILE=datei
[/KEEP={ALL    }] [/DROP=varlist]
    {varlist}
[/RENAME=(alte varlist=neue varlist)...]
[/MAP]
```

*FILE=datei*    Bezeichnung der portablen Datei

vollständige Windows-Dateibezeichnung in Apostroph oder Anführungszeichen: Laufwerksbuchstabe, Verzeichnisname, Dateiname, Dateityp (\* .por)

- Voreinstellung: keine Datei
- Beispiel: "c:\allbus\allbus96.por"

*KEEP=varlist*

Nur die hier aufgeführten Variablen werden in die Arbeitsdatei übernommen.

- Voreinstellung: *ALL*, d.h. alle Variablen

*DROP=varlist*

Die hier aufgeführten Variablen werden nicht übernommen.

*RENAME(alte varlist=neue varlist)*

Die Variablen erhalten in der Arbeitsdatei einen neuen Namen.

*MAP*

Die Namen der übernommenen Variablen werden im Ausgabefenster aufgelistet.

Eine portable SPSS-Datendatei wird über das Datei-Menü des Daten-Editors zu einer Arbeitsdatei mit:

►Datei ►Öffnen ►Dateityp: SPSS portabel (\*.por)