

4 Dateneingabe

Die Dateneingabe sollte anhand eines Erhebungsbogens oder einer Tabelle erfolgen, um eine einheitliche Datenstruktur zu gewährleisten. Nach Möglichkeit sollte man die Verwendung von alphanumerischen Variablen vermeiden und die einzelnen Merkmalsausprägungen mit Zahlen verschlüsseln. Dieses Vorgehen vereinfacht die Dateneingabe und erleichtert die Datenauswertung. Erfasst man die Daten in Excel kann man die eingegebenen Zahlen über den Befehl BEARBEITEN ⇨ ERSETZEN gegen den gewünschten Text austauschen oder zur Vermeidung unterschiedlicher Bezeichnungen für die gleiche Merkmalsausprägung mit einer Auswahlliste (Befehl DATEN ⇨ GÜLTIGKEIT) arbeiten oder ganz genau festlegen, welche Ausprägungen in welcher Spalte zugelassen sind.

Die Datenwerte werden in einem Dateneingabeschritt (DATA-Schritt) oder mit der Importfunktion in ein SAS-Datenfile (**dataset**) eingelesen. Dieses selbst entspricht im Aufbau einer Tabelle mit Zeilen und Spalten. Die Zeilen nennt man SAS-Observations (Beobachtungen, Datensätze, cases) und die Spalten Variablen (Merkmale, Felder).

Beispiel für eine Datentabelle:

		Zeile								
Spalte	Lfd.Nr.	Sex	Größe	Gewicht	Alter	Therapie	Gruppe	Aufnahme	Entlassung	
	1	1	weiblich	180	80,0	29	Placebo	A	12.10.93	24.10.93
2	2	männlich	160	45,9	30	Med A	A	10.10.92	11.10.92	
3	3	männlich	175	78,0	26	Placebo	B	01.01.91	30.03.91	
4	4	männlich	203	90,3	24	Med B	A	27.06.94	30.06.94	
5	5	weiblich	173	68,7	18	Placebo	B	16.04.92	24.04.92	
6	6	weiblich	150	42,5	8	Med B	B	31.08.90	01.09.90	
7	7	weiblich	180	112,0	27	Med A	A	27.05.92	28.05.92	
8	8	männlich	170	80,0		Med A	A	16.09.93	17.09.93	
9	9	männlich	120	30,0	11	Med B	B	08.07.94	09.07.94	
10	10	weiblich	192	90,0	23	Med A	B	01.02.92	03.03.92	
11										

Jede Tabellenzeile stellt eine Beobachtung dar, die sich aus Informationen über ein Objekt, eine Zelle, ein Versuchstier oder eine Person zusammensetzt. Die unterschiedlichen Informationen in den Spalten bezeichnet man als Variablen (Merkmale), die numerische Werte (Zahlen) oder alphanumerische Werte (Zeichenketten, character) annehmen können. Die Variablen werden über Variablennamen benannt, vgl. Spaltenbeschriftung in einer Tabelle. Die Tabellenspalte enthält die beobachteten Werte (Merkmalsausprägungen) für eine Variable.

Jede Tabellenzelle stellt für das SAS-System einen Datenwert (DATA VALUE) dar, wobei dieser für SAS die kleinste verarbeitbare Informationseinheit ist. Eine SAS-Datendatei (= DATENTEIL, DATA PART) beinhaltet dann alle Datenwerte und die zusätzlichen Informationen, wie den Variablennamen und den zugehörigen Variablentyp (= DATENBESCHREIBUNG, DESCRIPTOR PART). Die gesamte Tabelle ergibt den Inhalt einer SAS-Datendatei.

Die Daten müssen beim Einlesen maschinenlesbar in Form einer Tabelle vorliegen, wobei dies aus externen Textdateien (Datei ohne Formatierungen), aus Tabellenkalkulationsprogrammen und Datenbankprogrammen möglich ist. Die Datenwerte können aber auch direkt in SAS eingegeben werden.

Syntax des DATA-Schritts:

```
DATA LIBNAME.name ;
    (LIBNAME, siehe "1.4 SAS-Libraries und Libnames")
```

```
DATA name ;
```

```
DATA name ; SET LIBNAME.name ;
```

name steht für den Namen des SAS-Datenfiles. Gibt man einen zweistelligen Namen an, erstellt SAS eine permanente Datei, die auf dem über den LIBNAME festgelegten Datenträger abgelegt wird. Bei einem einstelligen Namen schreibt SAS das angelegte Datenfile in die Library WORK und es entsteht eine temporäre Datei, die beim Beenden von SAS gelöscht wird (siehe "1.5 SAS-Dateien").

Beispiele:

```
DATA meyer.werte ;
DATA daten ;
DATA meyer ; SET meyer.daten ;
```

Im ersten und dritten Beispiel ist **meyer** der LIBNAME. Im ersten Beispiel arbeitet der Benutzer mit dem permanenten Datenfile **werte.SAS7BDAT**, das im durch den LIBNAME festgelegten Ordner abgelegt wurde. Im zweiten Beispiel wird eine temporäre Datendatei mit dem Namen **daten** erstellt. Im dritten Beispiel verwendet der Benutzer das Datenfile **meyer** in der Library WORK, das aus dem permanenten Datenfile **meyer.daten** generiert wird.

4.1 Der INPUT-Befehl

Dieser Befehl liest die Daten eines Datensatzes in der vorgegebenen Form in die angegebenen Variablen ein und gibt die Struktur, die Position und den Namen der einzelnen Variablen an. Man unterscheidet zwischen dem freien und dem spaltengesteuerten Eingabeformat.

4.1.1 Freies Eingabeformat im INPUT-Befehl

Allgemeine Form:

Angaben in Klammern sind optional. [\$] weist darauf hin, dass ein \$-Zeichen angegeben werden kann, wenn eine alphanumerische Variable eingelesen werden soll, bei numerischen Variablen lässt man es einfach weg.

```
INPUT varname1 varname2 varname3 [$] varname4 ... ;
```

Beispiel:

```
INPUT nr sex $ alter groesse ;
```

Voraussetzung für diese Form des Dateneinlesens ist, daß die einzelnen Variablenausprägungen durch mindestens ein Leerzeichen getrennt sind und alphanumerische Variablen kein Leerzeichen enthalten. Fehlende Werte müssen durch die Eingabe eines Punktes gekennzeichnet werden. Das \$ hinter einem Dateinamen gibt an, daß es sich bei der betreffenden Variablen um eine alphanumerische handelt:

nr	sex	alter	groesse
1	m	22	180
2	w	18	160
3	w	28	165

4.1.2 Spaltengesteuertes Eingabeformat im INPUT-Befehl

Allgemeine Form:

```
INPUT variablenname [$] startposition - endeposition ;
```

Beispiel:

```
INPUT nr 1-3 sex $ 4 alter 5-6 groesse 7-9 1 ;
```

Es erfolgt die Angabe der genauen Spaltenposition hinter dem jeweiligen Variablennamen. Soll diese Form des Dateneinlesens funktionieren, müssen die Datenwerte exakt an der festgelegten Spaltenposition stehen und für jede Variable der längste Datenwert bestimmt werden, damit genügend Platz reserviert werden kann. Fehlende Werte benötigen keine besondere Kennzeichnung, da eingegebene Leerzeichen automatisch als fehlender Datenwert interpretiert werden.

Im Beispiel bedeutet **nr 1-3**, daß die **Variable nr in der ersten Spalte beginnt und in der dritten Spalte endet**. Das **\$**-Zeichen hinter der Variablen **sex** zeigt an, daß es sich um eine alphanumerische Variable handelt, die Zahl 4, daß sich Variable **sex** in der 4. Spalte befindet. Die 1 hinter der Spaltenangabe der Variablen **groesse** legt fest, daß diese als dreistellige Variable mit einer Nachkommastelle eingelesen wird und die Angabe 7-9, daß Variable **groesse** an den Positionen 7 bis 9 zu finden ist.

Variablenname	nr			sex	alter		groesse		
Spalte/Position	1	2	3	4	5	6	7	8	9
	0	0	1	m	2	2	1	8	0
	0	0	2	w	1	8	1	6	0

4.1.3 Einlesen mehrerer Eingabezeilen pro Beobachtung

Man benötigt diese Angaben im INPUT-Befehl, wenn eine Eingabezeile nicht ausreicht, weil lange und viele Zeichenketten vorliegen.

Allgemeine Form:

Beispiel:

```
INPUT #1 var1 var2 var3 var4 INPUT #1 nr sex ort $ famstat $
      var6 var7          groesse gewicht
#2 var8 var9          #2 therapie aufnahme
  ;
#n          ;
```

Im Beispiel müssen für jede Beobachtung zwei Eingabezeilen vorliegen. Ist dies nicht der Fall, gibt es Probleme beim Einlesen der Daten. Fehlt beispielsweise die zweite Zeile einer Beobachtung liest SAS anhand des INPUT-Befehls die erste Zeile der folgenden Beobachtung als zweite Zeile der vorangegangenen Beobachtung ein. In der ersten Zeile erwartet SAS die Variablen **nr bis gewicht** und in der zweiten Zeile die Variablen **therapie und aufnahme**:

Beispiel:

nr	sex	ort \$	famstat \$	groesse	gewicht
therapie		aufnahme			
1	1	ma	ledig	180	75
3		20.05.98			
2	2	hd	verh	163	50
1		31.06.98			

4.1.4 Mehrfaches Einlesen einer Eingabezeile

Im listengesteuerten Eingabeformat (siehe "3.2 Direkte Dateneingabe") kann man SAS mit der Formatsteuerung **@@** anweisen, daß **in einer Eingabezeile die Datenwerte für mehrere Beobachtungen** stehen.

Allgemeine Form:

Beispiel:

```
INPUT var1 var2 ... @@ ; INPUT gewicht groesse @@ ;
DATALINES ; DATALINES ;
Daten 57 154 78 190 65 156 98 204
; 40 130 49 160
;
```

SAS liest die Werte **abwechselnd als gewicht und groesse** ein. So werden die Werte 57, 78, 65, 98, 40 und 49 als *gewicht* und die Werte 154, 190, 156, 204, 130 und 160 als *groesse* interpretiert.

4.2 Direkte Dateneingabe

Bei der **direkten Dateneingabe**, auch als listengesteuertes Einlesen bezeichnet, findet die Erfassung der Daten im SAS-Auswertungsprogramm statt. Der Beginn der Dateneingabe wird durch ein DATALINES-Statement gekennzeichnet und das Ende durch ein Semikolon angezeigt.

Die Erfassung der Daten selbst erfolgt zeilenweise, der Definition der Variablen im INPUT-Statement (siehe "3.1 Der Input-Befehl" und "4.1.1 Programm zur direkten Dateneingabe") entsprechend, wobei die Länge der Eingabezeile 80 Spalten/Stellen nicht überschreiten sollte. Für fehlende Werte schreibt man einen Punkt.

Liegen viele Variablen vor, verwendet man am besten die spaltengebundene Dateneingabe (siehe "3.1.2 Spaltengesteuertes Eingabeformat im INPUT-Befehl") und tippt die Datenwerte rechtsbündig direkt hintereinander ein. Für die Erfassung **weniger Variablen** benutzt man das freie Eingabeformat ("3.1.1 Freies Eingabeformat im INPUT-Befehl") und trennt die einzelnen Variablen durch Leerzeichen voneinander.

Alphanumerische Werte (Zeichenketten) dürfen im freien Eingabeformat (nachfolgendes Beispiel, linke Spalte) nicht länger als acht Zeichen sein und keine Leerzeichen enthalten. Im freien Eingabeformat müssen zwei Datenwerte durch mindestens ein Leerzeichen getrennt werden.

Die eingegebenen Daten werden mit dem Programm abgelegt!!!

Beispiel (im PROGRAM EDITOR):

00001 OPTIONS ... ;	00001 OPTIONS ... ;
00002 INPUT nr sex alter	00002 INPUT nr 1-3 sex 4 alter 5-6
00003 groesse gew ;	00003 groesse 7-9 gew 10-11 ;
00004 DATALINES ;	00004 DATALINES ;
00005 001 2 28 180 70	00005 00122818070
00006 002 1 8 130 30	00006 0021 813030
00007 003 2 . 170 68	00007 0032 .17068
00008 ;	00008 ;
00009 RUN ;	00009 RUN ;
wenige Variablen - freie Eingabe	viele Variablen - spaltengebundene Eingabe

Diese Möglichkeit findet jedoch nur Verwendung, wenn eine geringe Menge von Beobachtungen betrachtet wird. Gibt man unter dem DATALINES-Statement viele Datensätze ein, verliert das Programm an Übersichtlichkeit.

4.3 Dateneingabe über externe Datenfiles

Bei dieser Methode erfolgt die **Dateneingabe in einer separaten Datei**, auf die im zugehörigen Auswertungsprogramm ein Hinweis gemacht werden muß. Dazu verwendet man das INFILE-Statement (siehe "4.1.2 Programm zur Dateneingabe über externe Datenfiles"):

Allgemeine Form (name = Name des SAS-Datenfiles):

```
DATA libname.name ;
INFILE 'laufwerk:\ordner\dateiname' ;
INPUT variablenliste ;
:
```

Beispiel:

```
DATA meier.daten ;
INFILE 'D:\sas\daten.dat' ;
INPUT nr sex alter groesse gew ;
:
```

Die Definition und Erstellung des SAS-Datenfiles findet über ein LIBNAME-Statement statt (siehe "1.4 SAS-Libraries und Libnames").

Auch in einer .DAT-Datei gibt man die Daten zeilenweise ein. Da bei dieser Art der Datenerfassung meist noch kein zugehöriges Auswertungsprogramm existiert, orientiert man sich bei der Dateneingabe am zuvor erstellten Erhebungsbogen bzw. einer Tabelle. Liegen viele Variablen vor, verwendet man am besten die spaltengebundene Dateneingabe (siehe "3.1.2 Spaltengesteuertes Eingabeformat im INPUT-Befehl") und tippt die Datenwerte rechtsbündig direkt hintereinander ein.

Für die Erfassung weniger Variablen benutzt man das freie Eingabeformat (siehe "3.1.1 Freies Eingabeformat im INPUT-Befehl") und trennt die einzelnen Variablen durch Leerzeichen voneinander. Die Beschreibung in der INPUT-Anweisung muß dabei der Anordnung der Datenwerte im externen Datenfile entsprechen.

Alphanumerische Werte (Zeichenketten) dürfen im freien Eingabeformat (linke Spalte) nicht länger als acht Zeichen sein und keine Leerzeichen enthalten. Im freien Eingabeformat müssen zwei Datenwerte durch mindestens ein Leerzeichen getrennt werden. Fehlende Werte kennzeichnet man durch die Eingabe eines Punktes!

Beispiel (im PROGRAM EDITOR):

00001 001 2 28 180 70	00001 00122818070
00002 002 1 8 130 30	00002 0021 813030
00003 003 2 . 170 68	00003 0032 .17068
wenige Variablen - freie Eingabe	viele Variablen - spaltengebundene Eingabe

Speichert man die Daten zum ersten Mal, wählt man den Befehl FILE ⇨ SAVE AS und trägt die gewünschten Angaben an zutreffender Stelle in das erscheinende Dialogfeld ein (Laufwerk, Suchpfad, Dateiname). Um die Daten unter dem ursprünglichen Dateinamen zu sichern, verwendet man den Befehl FILE ⇨ SAVE. Näheres siehe 2.2.2 *Speichern einer Programmdatei oder einer Datendatei mit der Extension .DAT während einer SAS-Sitzung.*

4.4 Dateneingabe in Excel

Nach dem Starten von Excel öffnet sich standardmäßig eine leere Arbeitsmappe mit drei Einzeltabellen. **Man bezeichnet diese Tabellen als Tabellen(blätter) oder auch als Arbeitsblätter.** Die Anzahl der Tabellenblätter pro Arbeitsmappe kann beliebig vorgegeben werden (min. Anzahl: 1, max. Anzahl: 255).

Ein Excel-Tabellenblatt ist in Zeilen (max. $65.536 = 2^{16}$) und Spalten (max. $256 = 2^8$) unterteilt. Eine Zelle ergibt sich aus der Kombination einer Zeile mit einer Spalte. Die aktive Zelle wird von einem dick gezeichneten Rechteck umrahmt und der Spaltenbuchstabe und die Zeilennummer farblich hervorgehoben.

In jede Zelle kann man Zahlen, Texte, Datumsangaben, Funktionen, Formeln oder Rechenvorschriften eintragen. Dazu positioniert man den Mauszeiger in der betreffenden Zelle und tippt die gewünschten Zeichen ein. Eine Eingabe wird abgeschlossen, wenn man die Eingabetaste oder eine der Cursortasten drückt. Je nach dem, ob in der Karteikarte BEARBEITEN des Befehls EXTRAS ⇨ OPTIONEN das Kontrollfeld MARKIERUNG NACH DEM DRÜCKEN DER EINGABETASTE VERSCHIEBEN aktiviert und eine Richtungsangabe für den Cursor gewählt wurde oder nicht, springt Excel nach dem Betätigen der Eingabetaste zur nächsten Zelle oder bleibt in der aktuellen Zelle stehen.

Will man während der Eingabe einen Tippfehler verbessern, betätigt man die Rücklöschtaaste. Zum Abbrechen einer Eingabe, drückt man die Esc-Taste oder den Schalter X (= Abbrechen) in der Bearbeitungsleiste. Ist die Eingabe bereits abgeschlossen, muß man die zu bearbeitende Zelle zuerst per Doppelklick markieren. Excel setzt den Cursor hinter den letzten Buchstaben bzw. vor die erste Zahl der selektierten Zelle. Das/die zu bearbeitende(n) Zeichen kann(können) dann über die Cursortasten (←, →) oder per Mausclick angewählt und bearbeitet werden.

4.4.1 Besonderheiten bei der Eingabe von Zahlen

- Excel erkennt **Zahlen** automatisch und ordnet sie standardmäßig **rechtsbündig** in der Zelle an.
- Währungs- oder Maßeinheiten werden **nicht** mit der Zahl eingegeben, sondern über **Zahlenformate** zugewiesen.
- Einer **negativen Zahl** muß ein **Minuszeichen** vorangestellt werden.
- Excel rundet standardmäßig bei mehr als 9 eingegebenen Nachkommastellen kaufmännisch auf 9 Stellen nach dem Komma, rechnet jedoch, unabhängig von der Darstellung, mit dem ungerundeten Inhalt der Zelle.
- Um **Zahlen ohne Dezimalstellen** anzuzeigen, deaktiviert man im Befehl EXTRAS ⇨ OPTIONEN unter der Karteikarte BEARBEITEN das Kontrollfeld FESTE DEZIMALSTELLE SETZEN.
- **Sehr große Zahlen** werden in der **Exponentialschreibweise** dargestellt, z.B. $2,407E+10$. 2,407 ist dabei die Mantisse, +10 der Exponent zur Basis 10. $2,407E+10$ bedeutet daher: $2,407 * 10^{10}$ (=24070000000).
Zeigt Excel anstelle der eingegebenen Zahl einfach **Nummernzeichen (#)** an, muß man die Breite der Spalte vergrößern, weil diese trotz der Kurzschreibweise zu schmal eingestellt ist.
- Eine Zahl darf auch mit einem **Prozentzeichen** am Ende eingegeben werden, z.B. 5%. Excel weiß dann, daß es die Zahl bei Berechnungen durch 100 teilen muß, aus 5% wird dann bei Rechenvorgängen 0,05.

4.4.2 Besonderheiten bei der Eingabe von Text

- **Text** wird in der Standardeinstellung **linksbündig** in der Zelle ausgegeben.
- Excel erkennt an den eingegebenen Zeichen automatisch, ob es sich um Text, Zahlen oder Formeln handelt.
- Reicht die eingestellte Breite einer Spalte nicht aus, um den Texteintrag vollständig anzuzeigen, werden die Nachbarzellen rechts davon, falls sie **leer** sind, überdeckt. Enthalten diese jedoch Einträge, schneidet Excel den Text an der rechten Zellbegrenzung der aktuellen Zelle scheinbar ab. Der Inhalt der aktiven Zelle bleibt jedoch erhalten und wird beim Aktivieren der Zelle in der Bearbeitungsleiste komplett angezeigt.
- Um eine Zahl als Text zu definieren, weist man **über den Befehl FORMAT** ↗ **ZELLEN** das **Format TEXT** zu. **Auf keinen Fall** sollte man, wie in einigen Handbüchern oder Hilfen beschrieben, vor der Zahl ein Hochkomma (⇧ + #-Taste) eingeben, z.B. '1999. Excel entfernt das Hochkomma nach dem Betätigen der Eingabetaste und es gibt keine Möglichkeit mehr, die Zahl wieder als solche zu verwenden.
- Eine Formel oder Funktion **ohne** vorangestelltes Gleichheitszeichen wird als Text aufgefaßt.

4.4.3 Besonderheiten bei der Eingabe von Datumswerten oder Zeitangaben

- **Datumswerte** werden in der Standardeinstellung wie Zahlen **rechtsbündig** in der Zelle ausgegeben.
- Der **Excel-Kalender** beginnt am **01.01.1900** und endet am **31.12.9999**. Alle Datumsangaben dazwischen werden von Excel auch als Datum gewertet, wobei die Angaben für Tag, Monat und Jahr stimmen müssen. Gibt man den 29. Februar für ein Jahr ein, das kein Schaltjahr ist, interpretiert Excel dieses als Text und richtet es in der Zelle linksbündig aus.
- Excel speichert eine **Datumsangabe als serielle Zahl**. Für ein **Datum** wird als **Ausgangsbasis der Wert 1** verwendet und in ganzzahligen Schritten weiter gezählt wird. Das Datum 01.01.1900 entspricht der Ausgangsbasis 1, d.h. Excel speichert dann das Datum 02.01.1900 als 2 und das Datum 03.01.1900 als 3 und so weiter. Ein Datum wird nur als solches angezeigt, wenn die betreffende Zelle mit einem Datumsformat formatiert ist. Weist man ein Zahlenformat zu, gibt Excel das Datum als serielle Zahl aus.
- Excel legt die **Uhrzeit als serielle Zahl zwischen 0 und 0,999** ab, was einer Zeitangabe von 00:00:00 bis 23:59:59 entspricht. Die serielle Uhrzeit berechnet sich nach der Formel: $x/24$, wobei x nach der Formel $STUNDE + MINUTE/60 + SEKUNDE/60 \cdot 60$ berechnet wird.

- Gibt man bei einer Zahl als Dezimaltrennzeichen statt einem Komma einen Punkt, einen Bindestrich oder einen Schrägstrich ein, interpretiert Excel diese Zahl als Datum und weist der Zelle jetzt als Standardformat das Datumsformat zu und alle Zahlen, die man in diese Zelle einträgt (auch bei der Verwendung des Kommas als Dezimaltrennzeichen), werden dann als Datum interpretiert. Um diese Einstellung zu löschen, markiert man die betreffende Zelle und führt den Befehl **BEARBEITEN** ↗ **LÖSCHEN** ↗ **FORMATE** aus.

4.4.4 Eingabe von Sonderzeichen

- Zu den Sonderzeichen gehören Zeichen, die nicht direkt über die Tastatur eingegeben werden können, Zeichen aus anderen Sprachen oder Symbole.
- Jede Schriftart hat andere Zeichen. Die Schriften Wingdings und Webdings bietet viele Zeichen, die in Formulare eingesetzt werden können.

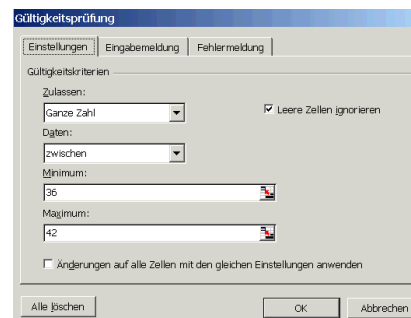
- Verwendung über den Befehl **EINFÜGEN** ↗ **SYMBOL**, Schriftart wählen, Zeichen anklicken und zum Beenden dieses Befehls den Schalter **OK** betätigen.

4.4.5 Gültigkeitsregeln für die Dateneingabe

Standardmäßig können in eine Tabellenzelle beliebige Zeichen eingegeben werden. Excel erkennt dabei, ob es sich um Zahlen, Buchstaben, ein Datum oder eine Uhrzeit handelt. Will man für eine Zelle nur bestimmte Zahlenwerte zulassen, legt man Eingabebeschränkungen fest oder definiert für Texte eine Liste mit den zu verwendenden Einträgen.

Eingabebeschränkungen für Zahlenwerte:

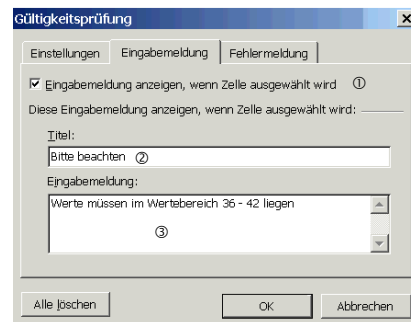
- Zelle(n), für die die Beschränkung gelten soll(en), markieren. Handelt es sich um eine komplette Spalte oder Zeile, setzt man die Markierung über den Spalten- bzw. Zeilenkopf:
- Befehl **DATEN** ↗ **GÜLTIGKEIT** ausführen und Karteikarte **EINSTELLUNGEN** öffnen: In den Feldern **ZULASSEN** und **DATEN** die gewünschte Option aus der Liste wählen und bei den zusätzlich erscheinenden Feldern die gewünschten Werte eintippen.



Kontrollfeld **LEERE ZELLEN IGNORIEREN** unterdrückt die Ausgabe einer Fehlermeldung für leere Zellen.

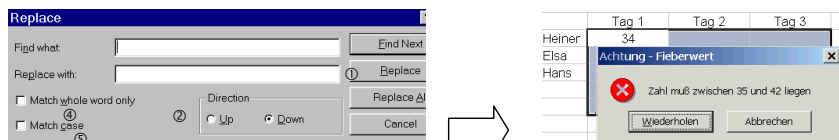
Kontrollfeld **ÄNDERUNGEN AUF ALLE ZELLEN MIT DEN GLEICHEN EINSTELLUNGEN ANWENDEN** steht dann zur Verfügung, wenn in der Karteikarte **EINSTELLUNGEN** Änderungen vorgenommen werden. Aktiviert man das Kontrollfeld übernimmt Excel die Änderungen für alle Zellen mit gleichen Gültigkeitsregeln.

- ◆ Im Beispiel oben wurden die Optionen **GANZE ZAHL** ① und **zwischen** ② eingestellt.
- ◆ Die in den Feldern **MINIMUM** ③ und **MAXIMUM** ④ vorgegebenen Werte sollen die untere bzw. obere Grenze für einen Fieberwert festlegen, d.h. für die markierten Zellen sind nur Werte zwischen 35 und 42 zugelassen.
- Karteikarte **EINGABEMELDUNG** öffnen. Sie dient zum Festlegen einer Hinweismeldung, die beim Eingeben der Zellwerte angezeigt wird:



- ◆ Deaktiviert man das Kontrollfeld **EINGABEMELDUNG ANZEIGEN WENN ZELLE AUSGEWÄHLT WIRD** ①, wird keine Eingabemeldung angezeigt.
- ◆ Option **TITEL** ② legt die Überschrift im Infocfeld fest. Sie wird fett formatiert ausgegeben.

- ◆ Der unter der Option EINGABEMELDUNG ③ eingegebene Text erscheint als Meldetext. Er darf maximal 255 Zeichen lang sein. Um innerhalb des Meldetextes eine neue Zeile zu beginnen, drückt man die Eingabetaste.
- In der Karteikarte FEHLERMELDUNG kann man eine Meldung eingeben, die im Falle einer falschen Angabe auf dem Bildschirm angezeigt wird:



- ◆ Deaktiviert man das Kontrollfeld FEHLERMELDUNG ANZEIGEN, WENN UNGÜLTIGE DATEN EINGEGEBEN WURDEN ①, wird keine Fehlermeldung angezeigt.
- ◆ Option TITEL ② legt die Überschrift für den Hinweis an den Benutzer fest. Sie erscheint in der Titelleiste des erscheinenden Dialogfensters.
- ◆ Der unter der Option FEHLERMELDUNG ③ eingegebene Text erscheint als Meldetext. Er darf maximal 255 Zeichen lang sein. Um innerhalb des Meldetextes eine neue Zeile zu beginnen, drückt man die Eingabetaste.
- ◆ Im Listenfeld TYP ④ kann man festlegen als was die Fehlermeldung angezeigt wird und das weitere Vorgehen bei einer Falscheingabe bestimmen.
STOPP ⑤ (Standardvorgabe, siehe Beispiel): Zeigt die Schaltflächen WIEDERHOLEN und ABBRECHEN. WIEDERHOLEN bedeutet, daß Excel den eingegebenen Wert löscht und auf eine neue Eingabe wartet. ABBRECHEN stellt den vorherigen Zelleintrag wieder her.
INFORMATION: Schlägt die Schaltflächen OK und ABBRECHEN vor. Schaltfläche OK übernimmt den falschen Wert, ABBRECHEN stellt den vorherigen Zelleintrag wieder her.
WARNUNG: Stellt die Frage FORTFAHREN? und bietet die Schaltflächen JA, NEIN und ABBRECHEN an. Schaltfläche JA übernimmt den falschen Wert, obwohl er gegen die eingestellten Gültigkeitsregeln verstößt. NEIN bricht die Eingabe ab, markiert die Zelle und wartet auf eine neue Eingabe. ABBRECHEN stellt den vorherigen Zelleintrag wieder her.

- Schalter OK betätigen, um die Eingabe der Gültigkeitsregeln zu bestätigen.

Anmerkung: Die Gültigkeitsprüfung gehört zu den Zellformatierungen und wird allen markierten Zellen zugewiesen, deshalb sollte man sie für ganze Spalten definieren.

Hinweis: Man kann auch für bereits eingegebene Daten Gültigkeitsregeln festlegen. Die Überprüfung der Daten erfolgt dann über die **Formelüberwachung**. Dazu weist man den betreffenden Zellen, wie oben beschrieben, die gewünschten Regeln zu, **eine Eingabe- und Fehlermeldung muß nicht festgelegt werden**, und führt den Befehl EXTRAS ⇨ FORMELÜBERWACHUNG ⇨ DETEKTIVSYMBOLLEISTE ANZEIGEN aus und betätigt den Schalter (= UNGÜLTIGE DATEN MARKIEREN). Excel kennzeichnet die Zellen mit fehlerhaftem Inhalt durch eine rote Ellipse:

	A
1	35
2	36
3	40
4	43

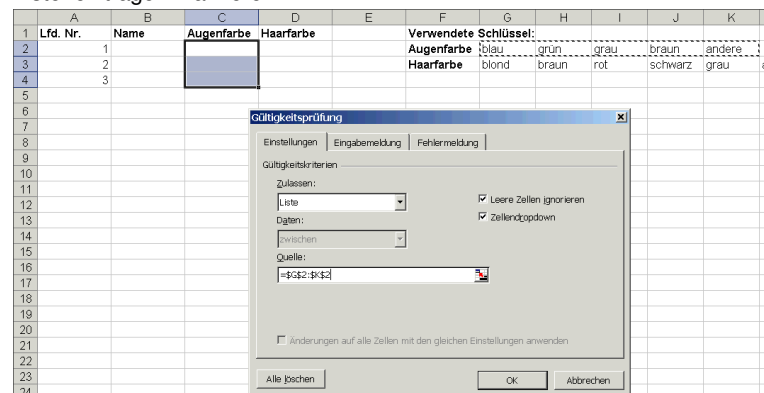
Will man die Markierung der fehlerhaften Zellen aufheben, betätigt man den Schalter (= GÜLTIGKEITSKREISE LÖSCHEN) in der FORMELÜBERWACHUNGS-Leiste.

Eingabeliste für Texte:

Bei dieser Form der Dateneingabe definiert man eine Liste mit gültigen Texteinträgen. Die Texte für die Listeneinträge gibt man im gleichen Tabellenblatt, wie die Daten ein, am besten unter oder rechts neben der Datentabelle.

Die folgenden Schritte müssen für alle Zellen, bei denen eine Eingabeliste verwendet werden soll, durchgeführt werden.

- Zelle(n), für die die Eingabeliste gelten soll(en), markieren:
- Befehl DATEN ⇨ GÜLTIGKEIT ausführen, Karteikarte EINSTELLUNGEN öffnen und im Listenfeld ZULASSEN den Listeneintrag LISTE wählen. Eine Eingabe- und Fehlermeldung muss nicht definiert werden.
- Das Eingabefeld QUELLE aktivieren, mit der Maus die Zellen mit den zu verwendenden Listeneinträgen markieren



und das Dialogfenster über den Schalter OK schließen.

⇒ Excel zeigt neben der selektierten Zelle den Schalter an, um zu zeigen, dass eine Liste vorhanden ist.

	A	B	C	D	E	F	G	H	I	J	K
1	Lfd. Nr.	Name	Augenfarbe	Haarfarbe		Verwendete Augenfarbe	Schlüssel:				
2	1		blau	grün		blau	grün	grau	braun	andere	
3	2		grün	braun		grün	braun	rot	schwarz	grau	an
4	3		grau	braun		braun	rot	schwarz	grau	an	
5			braun	andere							
6			andere								

- Daten eingeben und dabei für die entsprechenden Variablen bei der Eingabe das zugehörige Listenfeld öffnen und den gewünschten Listeneintrag auswählen.

4.5 EXCEL-Daten in SAS einlesen und als SAS-Datenfile abspeichern

SAS ist in der Lage Excel-Daten und auch Daten aus anderen Programmen zu verarbeiten. Dazu ist es allerdings notwendig diese nach SAS zu importieren. Für den Import-Vorgang gibt es zwei Möglichkeiten: den IMPORT-Assistent (was für mehrere Arbeitsmappen bzw. Tabellenblätter ziemlich aufwendig werden kann) oder die IMPORT Prozedur.

Voraussetzungen, die in der EXCEL-Datei erfüllt sein müssen:

- Die **Spalten** enthalten die **Merkmale**, die untersucht werden sollen, die **Zeilen** stellen die **Beobachtungen** dar.

	Zeile					
Spalte	lfdnr	sex	groesse	gewicht	diagnose	alter
	1	maennlich	180	82	Kopfschmerzen	38
	2	weiblich	165	54	Ohrensausen	67
	3	männlich	175	78	Halsschmerzen	24
	...					

- Variablenamen sollten möglichst kurz sein und **nur** aus Buchstaben und Zahlen bestehen, wobei das erste Zeichen immer ein Buchstabe sein muß. SAS erlaubt für Variablenamen höchstens 32 Zeichen; **besser sind aber max. 8**, denn je länger die Namen desto mehr muss man im Auswertungsteil schreiben! Leerzeichen dürfen innerhalb eines Variablenamens nicht vorkommen.

Die Buchstaben ä, ö, ü und ß dürfen **nicht** verwendet werden! Sonderzeichen, wie z.B. / oder . oder - oder %, sind **keine** Buchstaben, Ausnahme ist der Unterstrich: _!!!

Beispiel:

erlaubte Namen:

nr, lfd_nr, sex, geschl, alter, kompl1

unzulässige Namen:

lfd.nr, anteil%, rh-fakt, kompl/1

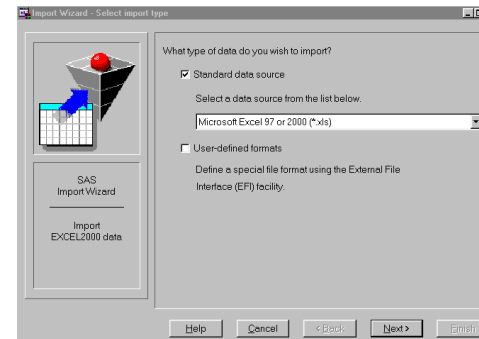
- In der Excel-Datei dürfen **nur** die Werte (Rohdaten) enthalten sein, **keine** Berechnungen oder Graphiken.
- Die Werte qualitativer Variablen sollten, der einfacheren Verarbeitung wegen verschlüsselt (mit Zahlen oder Text) und die Tabellenspalte mit optimaler Breite versehen werden. **Keine Romane in die Zellen schreiben, sondern kurze, aussagekräftige und prägnante Texte verwenden!!!**
- Sind viele Variablen vorhanden, diese auf mehrere Tabellenblätter aufteilen, dabei aber in jedem Tabellenblatt zur späteren Verknüpfung eine Identifikationsvariable angeben, z.B. eine laufende Nummer, die zeigt, welche Daten zusammengehören.
- Numerische Variablen** enthalten nur Zahlen, **alphanumerische** eine Kombination aus Buchstaben und Zahlen. SAS kennzeichnet fehlende Zahlenwerte mit einem Punkt und fehlende Texte mit einem Leerzeichen. **Die Merkmalsausprägungen einer Variablen müssen vom gleichen Typ sein, also entweder numerisch oder alphanumerisch!**
- Kommen in der ersten Datenzeile fehlende Werte vor, einen fiktiven Datensatz eingeben, der z. B. für fehlende Zahlen den Wert 0 enthält, für fehlende Buchstaben einen beliebigen Buchstaben und für eine nicht vorhandene Datumsangabe z.B. 01.01.1900 (Excel übergibt sonst einen falschen Variablentyp an SAS). Dieser Datensatz wird dann in SAS über ein Statement wieder aus den Daten gelöscht.

- Keine** personenbezogenen Daten eingeben. **Datenschutz!**

4.5.1 EXCEL-Daten mit dem IMPORT-Assisten in SAS einlesen und als SAS-Datenfile abspeichern

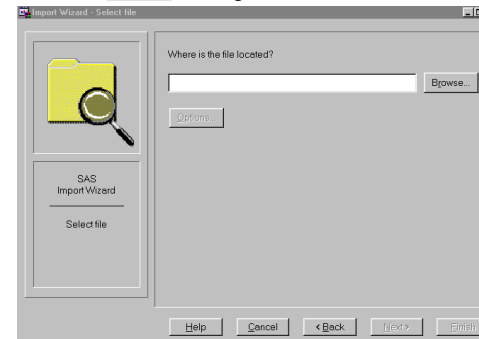
- Falls noch nicht geschehen: SAS starten und den zugehörigen LIBNAME einstellen (siehe "1.4 SAS - Libraries und Libnames").

- Den Befehl FILE ↗ IMPORT DATA... ausführen. SAS öffnet daraufhin das Dialogfenster des IMPORT WIZARDS:



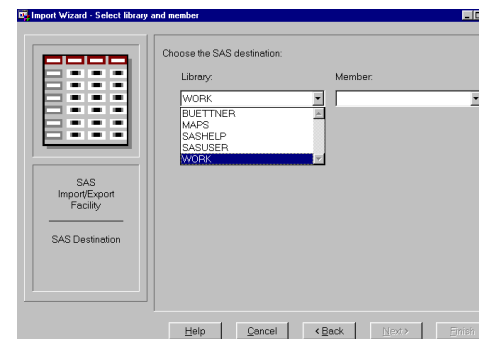
- Im Listenfeld SELECT A DATA SOURCE FROM THE LIST BELOW des Fensters IMPORT WIZARD - SELECT IMPORT TYPE die vorgeschlagene Excelversion MICROSOFT EXCEL 97 OR 2000 (*.XLS) für die einzulesende Datei übernehmen.

- Schalter NEXT> betätigen. SAS öffnet das Dialogfenster IMPORT WIZARDS - SELECT FILE:



- Im Textfeld WHERE IS THE FILE LOCATED? des Dialogfensters IMPORT WIZARD - SELECT FILE den Namen der zu importierenden einschließlich Suchpfad eingeben oder den Schalter BROWSE zum Einstellen des Pfades benutzen und den Dateinamen wählen.

- Schalter NEXT> betätigen. SAS öffnet dann das Dialogfenster IMPORT WIZARDS - SELECT LIBRARY AND MEMBER



07. Das Listenfeld LIBRARY öffnen, in der Liste den gewünschten LIBNAME anklicken und unter MEMBER den Namen für das zu erstellende SAS-Datenfile eingeben.

08. Schalter FINISH betätigen.

⇒ SAS zeigt die Meldung **Note: LIBNAME.member was successfully created** im LOG-Fenster und legt eine SAS-Datei mit dem Namen **member.SAS7BDAT** im durch den LIBNAME festgelegten Verzeichnis an. LIBNAME steht dabei für die gewählte Library und **member** für das angelegte Datenfile.

4.5.2 EXCEL-Daten über die Prozedur IMPORT in SAS einlesen und als SAS-Datenfile abspeichern

01. SAS starten und die folgenden Programmzeilen in den PROGRAM EDITOR eingeben:

```

OPTIONS LINESIZE=130 PAGESIZE=69 NOLABEL ;
LIBNAME name 'Laufwerk:\Ordner' ;

PROC IMPORT OUT=name.daten
  DATAFILE="Laufwerk:\Ordner\dateiname.xls"
  DBMS=EXCEL2000 REPLACE ;
  SHEET="Tabellenblattname" ;
  GETNAMES=YES;
RUN ;

DATA tname ; SET name.daten ;
    
```

Beschreibung der Programmzeilen:

Statement **OPTIONS** legt die Anzahl der Zeichen pro gedruckter Zeile (**LINESIZE=130**) und die Anzahl der Zeilen pro gedruckter Seite (**PAGESIZE=69**) fest. Option **NOLABEL** verhindert, dass der Variablenname als zusätzliche Erläuterung bei den Auswertungen auftaucht.

Statement **LIBNAME** definiert den LIBNAME für die einzulesenden Daten, siehe "1.4 SAS - Libraries und Libnames".

Die Prozedur **IMPORT** startet das Einlesen der Excel-Daten in ein SAS-Datenfile und benötigt dafür

den **SAS-Datenfile-Namen (OUT=name.daten)**; **name** ist dabei der Name des Libnames und **daten** der Name des zu erstellenden SAS-Datenfiles.

den **Speicherort** der Excel-Arbeitsmappe mit den einzulesenden Daten (**DATAFILE="Laufwerk:\Ordner\dateiname.xls"**); **dateiname.xls** steht für den Namen der Excel-Arbeitsmappe.

das **für die Dateneingabe verwendete Programm (DBMS=EXCEL2000 REPLACE)**; die Angabe **DBMS=EXCEL2000** gibt an, dass für die Dateneingabe mindestens Excel 2000 verwendet wurde und die **Option REPLACE** bewirkt, dass bei mehrfachem Einlesen das SAS-Datenfile immer wieder überschrieben wird.

den **Namen des einzulesenden Tabellenblattes (SHEET="Tabellenblattname")**.

Statement **GETNAMES=YES** ; übergibt SAS die Spaltenbeschriftung als Variablenamen. Statement **RUN** beendet die Prozedur.

Die Zeile **DATA tname ; SET name.daten (tname=Name des temporären Datenfiles, name=Name des Libnames, daten=Name des SAS-Datenfiles)**; macht aus dem permanenten Datenfile **name.daten** das temporäre Datenfile **tname**, das in die Library **WORK** geschrieben wird.

4.5.3 Bearbeiten von Excel-Daten

In die zu bearbeitende Zelle klicken und die gewünschten Änderungen am Zellinhalt in der Bearbeitungsleiste bzw. direkt in der Zelle durchführen. Das/die zu bearbeitende(n) Zeichen kann(können) dann über die Cursortasten (←, →) oder per Mausclick angewählt und bearbeitet werden.

Excel zeigt den Mitteilungstext BEARBEITEN in der Statuszeile an. Man befindet sich dann im Bearbeitungsmodus und kann die entsprechenden Zeichen ändern, überschreiben, löschen oder neue einfügen.

Wichtige Tasten:

←-Taste Korrekturtaste/Rücklöschtaete: löscht das Zeichen links vom Cursor

ENTF-Taste Löschtaste: entfernt das Zeichen rechts vom Cursor

Mit der Einfügetaste schaltet man den Überschreibenmodus ein (Kürzel ÜB erscheint in der Statuszeile), d.h. Excel überschreibt alle Zeichen, die rechts vom Cursor stehen mit den neu eingegebenen Zeichen.

EINFÜG-Taste Drückt man die Taste erneut, schaltet Excel zum standardmäßig eingestellten Einfügemodus um und das Kürzel ÜB in der Statuszeile verschwindet. Im Einfügemodus werden die getippten Zeichen an der Cursorposition eingefügt. Rechts vom Cursor stehende Zeichen werden dabei weiter nach rechts geschoben.

Änderungen müssen durch das Drücken der **↵**-Taste (= Eingabetaste, ENTER-Taste, RETURN-Taste) bestätigt werden. Sollen Änderungen nicht übernommen werden, drückt man die **ESC**-Taste.

Die geänderten Excel-Daten müssen selbstverständlich wieder in SAS eingelesen werden, denn nur dann kann SAS mit den neuen Daten arbeiten!!!

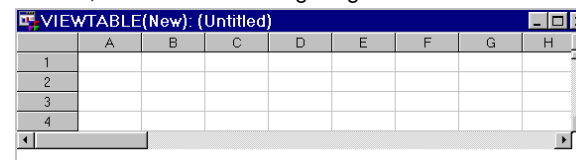
4.6 Dateneingabe und -bearbeitung über den VIEWTABLE-Editor

Die Dateneingabe mit Hilfe dieses SAS-Editors ist ziemlich umständlich, das Ändern von Daten dagegen recht einfach. Erzeugt werden temporäre oder permanente SAS-Datenfiles. Die Auswertung erfolgt mit SAS-Prozeduren.

4.6.1 Datenfile erstellen

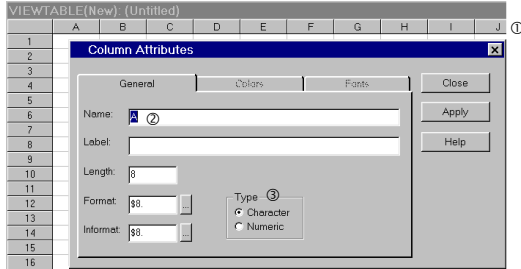
01. Falls noch nicht geschehen: SAS starten und den zugehörigen LIBNAME einstellen (siehe "1.4 SAS - LIBRARIES und LIBNAMES").

02. Der Befehl **TOOLS ⇨ TABLE EDITOR** startet den Editor. SAS zeigt daraufhin die folgende Tabelle, in die die Daten eingetragen werden können:



Eine Zeile stellt eine Beobachtung (Patient) und eine Spalte eine einzelne Variable mit ihren Ausprägungen dar.

03. Um die standardmäßige Bezeichnung der Tabellenspalten zu ändern, wählt man den Befehl **DATA COLUMN** ⇨ **ATTRIBUTES...**. Die Variablenamen können dann im erscheinenden Dialogfenster eingetragen werden:



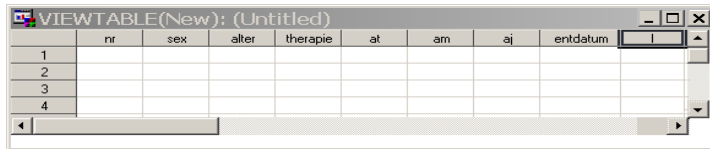
Wie geht man vor?

In den Spaltenkopf der zu benennenden Spalte klicken ① und den Namen der Variablen eingeben ② (Vorgaben für Variablenamen siehe "1.3 Regeln für die Definition von Dateinamen und LIBNAMES, Variablenamen und -typen"). Verdeckt das Dialogfenster COLUMN ATTRIBUTES... die Spaltenbeschriftung, klickt man mit der Maus in die Titelleiste und zieht es bei gedrückter Maustaste an eine andere Stelle auf dem Bildschirm.

Unter TYPE ③ wird für jede Variable der Merkmalstyp festgelegt, d.h. Variablen mit numerischen Ausprägungen erhalten den Typ NUMERIC und Variablen, deren Ausprägungen Kombinationen aus Buchstaben und Zahlen enthalten den Typ CHARACTER.

Für ein Datum wird zusätzlich auf den Schalter ... hinter dem Listenfeld FORMAT (Ausgabeformat für das Datum, d.h. das Aussehen des Datums in der Tabelle) geklickt, das Format ddmmyy aktiviert und hinter WIDTH die Zahl 10 angegeben, dann wird der Schalter ... hinter dem Listenfeld INFORMAT (Eingabeformat für das Datum, d.h. wie wird das Datum in die Tabelle eingegebenen) gedrückt, auch hier das Format ddmmyy aktiviert und hinter WIDTH wieder die Zahl 10 angegeben. Das Format erstellt ein Datum in der Form Tag.Tag.MonatMonat.JahrJahrJahr, z.B. 12.05.1996, 23.12.2005.

04. Das Betätigen des APPLY-Schalters, veranlasst SAS den eingegebenen Variablenamen und den gewählten Merkmalstyp zu übernehmen.
05. **Die Schritte 04) und 05) sind für alle Variablen durchzuführen.** Wurde allen Tabellenspalten ein Name und ein Merkmalstyp zugewiesen, beendet man die Vergabe der Variablenamen durch Drücken des CLOSE-Schalters:



Im Beispiel wurden die Variablen nr, sex, alter, therapie, at, am, aj und entdatum in der Datentabelle definiert.

06. Zum Eingeben der Datenwerte positioniert man den Cursor per Mausklick in die erste Tabellenzelle und tippt den zur Spaltenbeschriftung gehörenden Datenwert ein.

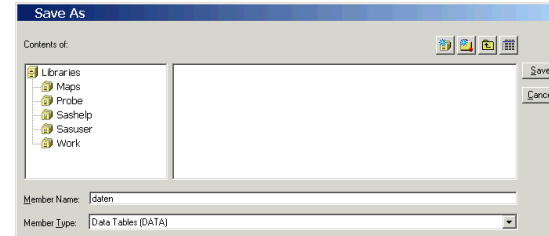
Anmerkungen:

Die TAB-Taste läßt den Cursor zur nächsten Zelle (nächste Variable) springen. Will man eine neue Datenreihe eingeben, zieht man die horizontale Bildschirmleiste nach links und klickt in die erste Tabellenzelle der nächsten Zeile.

Führende Nullen müssen nur beim Datum eingegeben werden. Sollen die Daten über den Zahlenblock rechts auf der Tastatur eingegeben, muß dieser, falls noch nicht geschehen, über die Taste NUM aktiviert werden.

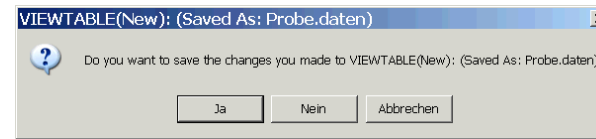
Achtung!!! Die Punkte zwischen den Datumskomponenten nicht eingeben.

07. Zum Speichern des Datenfiles wählt man den Befehl **FILE** ⇨ **SAVE AS**, klickt im Listenfeld CONTENTS OF auf den Namen der zu verwendenden LIBRARY (im Beispiel Probe), gibt in das Eingabefeld MEMBER NAME den gewünschten Dateinamen (im Beispiel daten) ein, (Vorgaben für Variablenamen siehe "1.3 Regeln für die Definition von Dateinamen und LIBNAMES, Variablenamen und -typen") und betätigt den Schalter **SAVE**.



Das Datenfile wird dann im durch den LIBNAME (im Beispiel **Probe**) festgelegten Ordner unter den angegebenen Namen (im Beispiel **daten**) abgelegt.

08. Der Befehl **FILE** ⇨ **CLOSE** schließt den VIEWTABLE-Editor. Erscheint die Frage "**Do you want to save the changes you made to VIEWTABLE...**", den Schalter **JA** betätigen, um das Speichern des Datenfiles zu bestätigen:



4.6.2 Datenfile laden und bearbeiten

Das Bearbeiten von Daten im VIEWTABLE-Editor ist nur bei kleinen Änderungen sinnvoll. Das Hinzufügen neuer Zeilen oder Spalten beispielsweise ist sehr umständlich. Liegen die benötigten Daten auch als Excel-Tabelle vor, empfiehlt es sich die Daten in der Excel-Arbeitsmappe zu bearbeiten und die korrigierten Daten erneut aus Excel in SAS einzulesen.

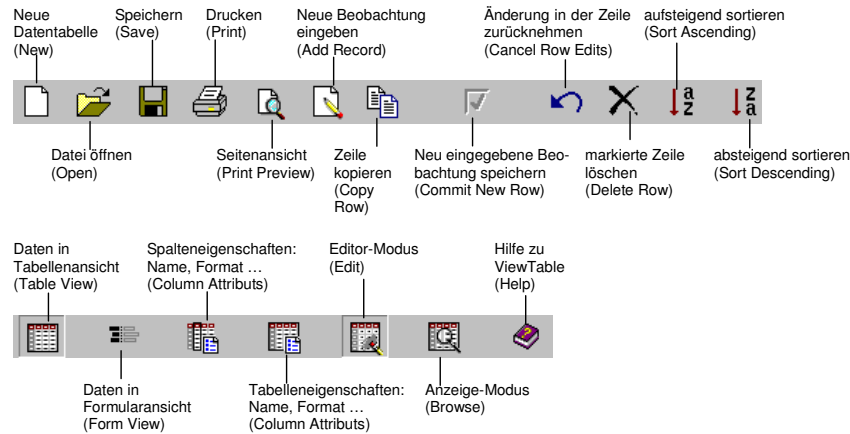
01. Falls noch nicht geschehen: SAS starten und den zugehörigen Libname einstellen (siehe "1.4 SAS - LIBRARIES und LIBNAMES").
02. Im SAS-Explorer-Fenster zuerst die benötigte LIBRARY und dann den Namen des gewünschten Datenfiles per Doppelklick öffnen.
03. Da SAS die Datendatei standardmäßig im BROWSE-Modus öffnet, den Befehl **EDIT** ⇨ **EDIT MODE** ausführen, um Änderungen an den Daten durchführen zu können.

Das **Menü EDIT** liefert Befehle zum Einfügen von Zeilen (ADD ROW), zum Kopieren und Löschen (COPY ROW und DELETE ROW) der aktuellen Zeile, zum Speichern neu eingegebener Zeilen (COMMIT NEW ROW), zum Durchführen von Suchabfrage (FIND, REPEAT FIND), zum Umschalten vom Anzeige- (BROWSE MODE) in den Editiermodus (EDIT MODE), zum Regeln des Zugriffs mehrerer Benutzer auf die gleich Datendatei (ROW LEVEL EDIT ACCESS zum Ermöglichen des gleichzeitigen Zugriffs mehrerer Benutzer auf eine

Tabelle bzw. Table Level EDIT ACCESS zum Verhindern des gleichzeitigen Zugriffs mehrerer Benutzer auf eine Tabelle).

Mit Hilfe des **Menüs DATA** können Bedingungen gestellt (WHERE) und wieder zurückgenommen (WHERE CLEAR), Spalten ein- und ausgeblendet (HIDE/UNHIDE), sortiert (SORT), beim Blättern fixiert (HOLD...) und Spalten- (COLUMN ATTRIBUTES) und Tabelleneigenschaften (TABLE ATTRIBUTES) festgelegt werden.

Viele der in den Menüs EDIT und DATA aufgeführten Befehle sind auch per Mausklick in der Symbolleiste abrufbar:



04. Zum Bearbeiten der Datenwerte positioniert man den Cursor per Mausklick in der gewünschten Tabellenzelle und führt die gewünschte Änderung durch.

05. Der Befehl FILE ⇨ SAVE speichert das Datenfile mit den durchgeführten Änderungen im durch den LIBNAME festgelegten Ordner unter den angegebenen Namen.

09. Der Befehl FILE ⇨ CLOSE schließt den VIEWTABLE-Editor.

4.7 Dateneingabe über SAS/INSIGHT

Die Dateneingabe mit Hilfe von SAS/INSIGHT ist recht einfach zu handhaben und erzeugt temporäre oder permanente SAS-Datenfiles, die direkt oder durch Verwendung von SAS-Prozeduren ausgewertet werden können.

Die Eingabe von Datums- und Zeitangaben ist in SAS/INSIGHT sehr umständlich. Die einfachste Lösung ist die Datumskomponenten Tag, Monat und Jahr bzw. die Zeitkomponenten Stunde, Minute und Sekunde jeweils als eigene Variable einzugeben und die Komponenten dann später im Programm mit Hilfe der Funktion MDY(monat;tag;jahr) bzw. HMS(stunde,minute,sekunde) in ein Datum bzw. Uhrzeit umzuwandeln.

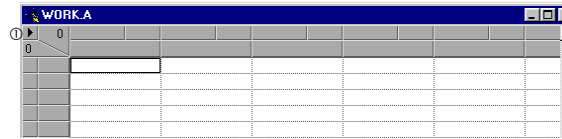
Im nachfolgenden Beispiel wurde das Aufnahme datum mit den drei Variablen *at*, *am* und *aj* eingegeben.

4.7.1 Datenfile erstellen

01. Falls noch nicht geschehen: SAS starten und den zugehörigen Libname einstellen (siehe "1.4 SAS - LIBRARIES und LIBNAMES").

02. Den Befehl SOLUTIONS ⇨ ANALYSIS ⇨ INTERACTIVE DATA ANALYSIS wählen.

03. Betätigt man den Schalter **NEW** erscheint eine Tabelle, vgl. Excel-Tabelle, in die die Daten eingetragen werden können:

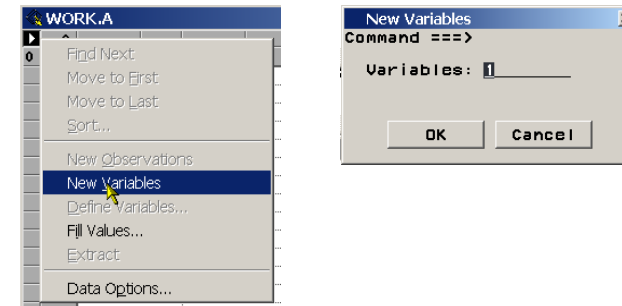


Die Zeilen stellen dabei die Beobachtung (Patient) und die Spalten die einzelnen Variablen dar. SAS nennt diese Tabelle provisorisch **WORK.A**, erlaubt aber beim Speichern einen anderen Namen.

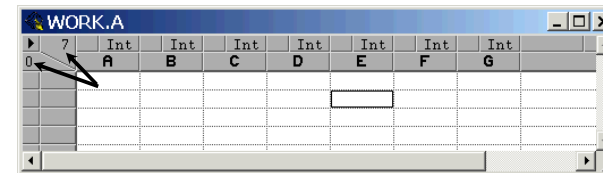
04. Als nächstes muß die Anzahl der aufzunehmenden Variablen festgelegt werden. Dazu drückt man den ►- Zeichen oben links in der Tabelle ①,

wählt den Befehl **NEW VARIABLES** aus der angezeigten Liste (linkes Bild) und betätigt die **EINFG**-Taste um den Überschreibenmodus einzuschalten.

Danach trägt man die Zahl für die zu bearbeitenden Variablen hinter dem Feld **VARIABLES** in das erscheinende Dialogfenster ein (rechtes Bild) und schaltet den Überschreibenmodus durch erneutes Drücken der **EINFG**-Taste wieder aus.



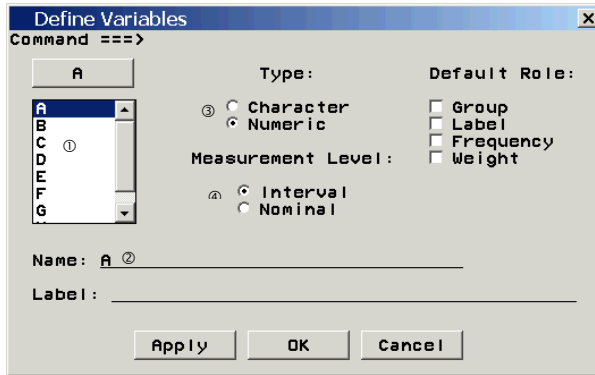
⇒ SAS benennt darauf so viele Tabellenspalten mit den Buchstaben des Alphabetes, wie Variablen definiert wurden (im Beispiel 7):



Die Zahl 7 im Beispiel gibt die Zahl der eingegebenen Variablen an und die 0 ganz links in der zweiten Tabellenzeile die Anzahl der bisher aufgenommenen Datensätze.

Die Abkürzung **INT** über den Spaltenbuchstaben besagt, daß es sich bei der betreffenden Variablen eine stetige Variable handelt. **NOM** steht für nominale Variablen.

05. Um die standardmäßige Bezeichnung der Tabellenspalten zu ändern, klickt man mit der Maus das ►- Zeichen oben links in der Tabelle ein zweites Mal an und wählt den Befehl **DEFINE VARIABLES** in der sich öffnenden Liste. Im erscheinenden Dialogfenster kann man dann den Variablennamen eintragen:



Im Listenfeld ① auf der linken Seite des Dialogfensters den Buchstaben der betreffenden Tabellenspalte markieren.

Das Feld NAME ② durch Anklicken des angezeigten Buchstabens aktivieren und den Namen der Variablen eingeben (Vorgaben für Variablenamen siehe "1.3 Regeln für die Definition von Dateinamen und LIBNAMES, Variablenamen und -typen").

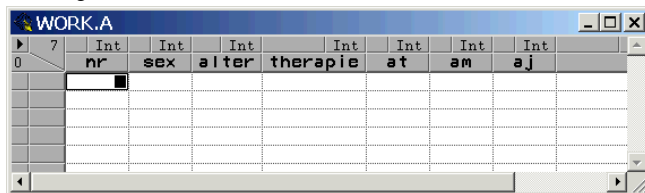
Unter TYPE ③ wird der Variablentyp vorgegeben: für Variablen, deren Ausprägungen als Text vorliegen die Option CHARACTER und für Variablen, deren Ausprägungen Zahlen enthalten die Option NUMERIC einstellen.

Unter MEASUREMENT LEVEL ④ wird für jede Variable der zutreffende Merkmalstyp festgelegt, d.h. quantitative/stetige Variablen erhalten den Typ INTERVAL und nominale/qualitative Variablen den Typ NOMINAL.

06. Zum Übernehmen des eingegebenen Variablenamens und des gewählten Merkmalstyps den APPLY-Schalter betätigen.

Nach dem **ersten** Drücken des Schalters APPLY muß der Schalter DEFINE VARIABLES in der Fensterleiste angeklickt werden.

07. Die Schritte 05) und 06) sind für alle Variablen durchzuführen. Der OK-Schalter beendet die Eingabe der Variablenamen und das Zuweisen des Variablentyps:

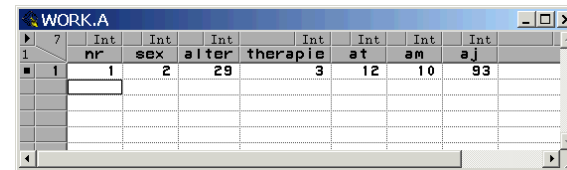


08. Zum Einstellen der Cursorrichtung klickt man das ►- Zeichen oben links in der Tabelle ein drittes Mal, wählt den Befehl **DATA OPTIONS** und aktiviert unter DIRECTION OF "ENTER": die Option **DOWN AND LEFT**.

09. Für die Dateneingabe positioniert man den Cursor per Mausclick in die erste Tabellenzelle und tippt den zur Spaltenbeschriftung gehörenden Datenwert ein. **Fehlende Werte werden durch die Eingabe eines Punktes gekennzeichnet.**

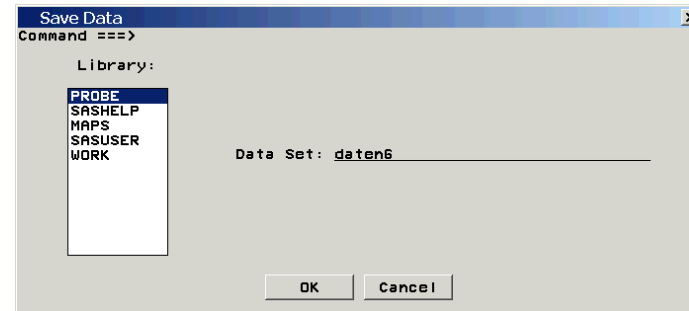
Die TAB-Taste läßt den Cursor zur nächsten Zelle (nächste Variable) springen. Am Ende der Tabellenzeile betätigt man die ↵-Taste um die Datenwerte in die Felder einzutragen.

Vergißt man die ↵-Taste zu drücken, trägt SAS statt der eingegebenen Wert fehlende Werte an.



SAS wartet in der nächsten Zeile auf die Erfassung eines weiteren Datensatzes.

10. Zum Speichern des Datenfiles wählt man den Befehl FILE ⇨ SAVE ⇨ DATA....



Im angezeigten Dialogfenster markiert man im Listenfeld LIBRARY den Namen des zu verwendenden Libnames (im Beispiel **PROBE**), klickt mit der Maus den Buchstaben hinter Feld **DATA SET** an, tippt den gewünschten Namen für das Datenfile ein (im Beispiel daten6) und betätigt den Schalter **OK**.

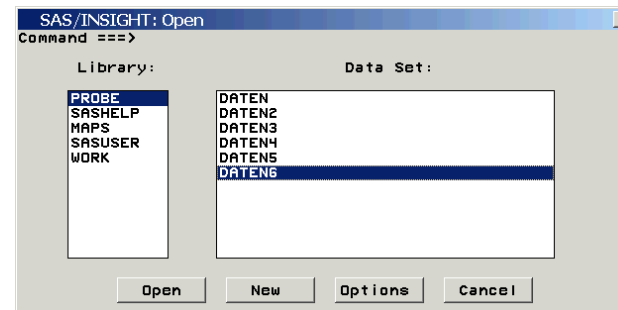
Vorgaben für Variablenamen siehe *1.3 Regeln für die Definition von Dateinamen und Libnames, Variablenamen und -typen*. Dieses wird dann im durch den Libname festgelegten Ordner unter den angegebenen Namen abgelegt.

11. Der Befehl FILE ⇨ END schließt das SAS/INSIGHT- Programm.

4.7.2 Datenfile laden und bearbeiten

01. Falls noch nicht geschehen: SAS starten und den zugehörigen Libname einstellen (siehe *1.4 SAS - LIBRARIES und LIBNAMES*).

02. Den Befehl SOLUTIONS ⇨ ANALYSIS ⇨ INTERACTIVE DATA ANALYSIS wählen.



03. Im angezeigten Dialogfenster markiert man im Feld LIBRARY den Namen des zu verwendenden Libnames (im Beispiel PROBE) und im Feld DATA SET den Namen der zu öffnenden Datendatei (im Beispiel DATEN6). Nach dem Drücken des Schalters OPEN zeigt SAS die Daten des gewählten Datenfiles in einer Tabelle an.
04. Zum Bearbeiten der Datenwerte positioniert man den Cursor per Mausklick in der gewünschten Tabellenzelle und führt die gewünschte Änderung durch:
Will man eine Zeile oder eine Spalte löschen, markiert man diese über den Zeilen- bzw. Spaltenkopf und wählt den Befehl Delete im Menü Edit.
Um Datensätze anzuhängen, positioniert man den Cursor in der erste Zelle der Zeile nach der letzten Datenzeile und gibt die Daten ein.
05. Im angezeigten Dialogfenster markiert man im Listenfeld LIBRARY den Namen des zu verwendenden Libnames (im Beispiel **PROBE**), klickt mit der Maus den Buchstaben hinter Feld **DATA SET** an, wählt den gewünschten Namen für das Datenfile (im Beispiel DATEN6) und betätigt den Schalter **OK**.



06. Der Befehl FILE ⇨ END schließt das SAS/INSIGHT- Programm.