

Universität Münster

Institut für Biometrie und Klinische Forschung

Querschnittsfach Epidemiologie, Biometrie und
Medizinische Informatik

Anleitung zur Bedienung des Statistik-Programms SPSS
(Version 27.0)

Inhalt

1. Benutzeroberfläche von SPSS

1.1	Daten-Editor.....	2
1.2	Variablenansicht	2
1.3	Der Ausgabe-Viewer	3

2. Datenaufbereitung

2.1	Öffnen und Speichern einer Datendatei.....	3
2.2	Öffnen und Speichern einer Ausgabedatei	5
2.3	Zusammenfügen von Datendateien (fall- und variablenweise).....	6

3. Datenselektion

3.1	Auswahl von Fällen	9
3.2	Fälle sortieren	10
3.3	Aufteilung der Fälle in Gruppen	10

4. Datenmodifikation

4.1	Berechnung von neuen Variablen	11
4.2	Rechnen mit Datumsangaben	12
4.3	Umkodieren und Klassieren von Werten.....	15
4.4	Aggregieren von Daten	16

1. Benutzeroberfläche von SPSS

SPSS bezeichnet eine Software für statistische Datenanalysen. Es gibt zwei Auswahlansichten: den Daten-Editor und die Variablenansicht.

1.1 Daten-Editor

Der Dateneditor, ein in Zeilen und Spalten aufgeteiltes Arbeitsblatt, ermöglicht das Öffnen einer Datei, das Eingeben, Bearbeiten und das Auswerten von Daten. Die einzelnen Zeilen entsprechen den Fällen, die einzelnen Spalten den Variablen. Die Zellen enthalten den Wert der jeweiligen Variable innerhalb eines entsprechenden Falles. Die Auswertung der Daten erfolgt in der Regel über Menübefehle (ggf. Programmsyntax) und diesen zugeordnete Dialogfelder. Die Menüleiste im oberen Bildrand wird durch eine darunter befindliche Symbolleiste ergänzt.

AML99.sav [DataSet1] - IBM SPSS Statistics Dateneditor

DateiBearbeitenAnsichtDatenTransformierenAnalysierenGrafikExtrasErweiterungenFensterHilfe

</

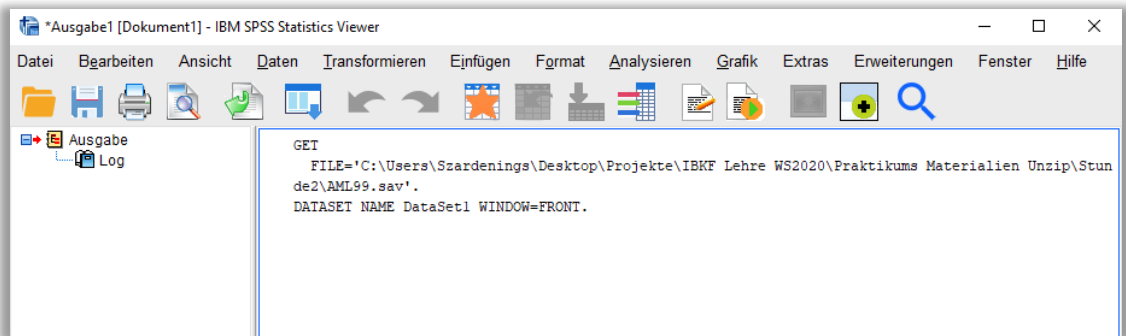
1.2 Variablenansicht

Die Variablenansicht dient zur Vergabe von Variablennamen. Dabei unterliegt die Vergabe von Variablennamen bestimmten Festlegungen (keine Leerzeichen, kein Beginn mit einer Ziffer, keine Sonderzeichen wie !, ?, *, & und .). Weiterhin werden Variablentyp, Spaltenformat, Variablenlabel, Wertelabels und das Messniveau (Skalenniveau) der Variablen definiert.

	Name	Typ	Breite	Dezimals...	Beschriftung	Werte	Fehlend	Spalten	Ausrichtung	Messniveau	Rolle
1	PID	Zeichenfolge	18	0		Ohne	Ohne	8	Links	Nominal	Eingabe
2	Therapie	Numerisch	8	0	Behandlungsgr...	{1, TAD-HA...	Ohne	8	Rechts	Nominal	Eingabe
3	Geschlecht	Numerisch	8	0	Geschlecht	{1, Männlich...	Ohne	8	Rechts	Nominal	Eingabe
4	alter	Numerisch	2	0	Alter bei Diagn...	Ohne	Ohne	10	Rechts	Metrisch	Eingabe
5	KGW	Numerisch	3	0	Körpergewicht i...	Ohne	Ohne	8	Rechts	Metrisch	Eingabe
6	KGR	Numerisch	3	0	Körpergröße in ...	Ohne	Ohne	8	Rechts	Metrisch	Eingabe
7	Zustand	Numerisch	8	0	Allgemeinzusta...	{0, Normal}...	Ohne	8	Rechts	Ordinal	Eingabe
8	Zustand_bin	Numerisch	8	0	Zustand, binär ...	{0, Normal o...	Ohne	8	Rechts	Ordinal	Eingabe
9	Fieber	Numerisch	5	2	Körpertemperat...	Ohne	Ohne	8	Rechts	Metrisch	Eingabe

1.3 Der Ausgabe-Viewer

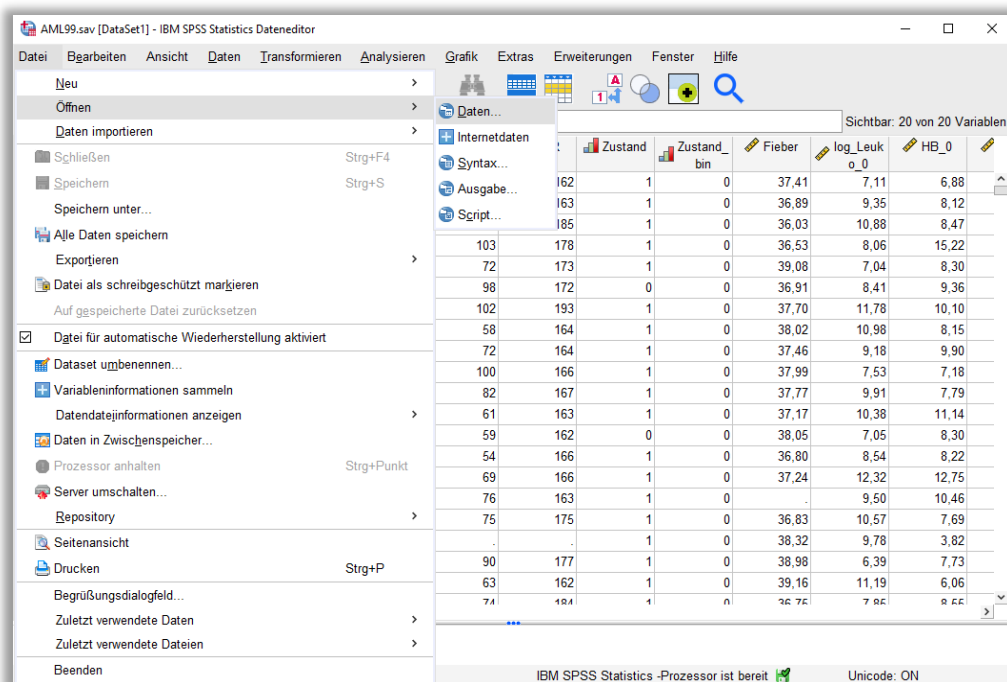
Der Ausgabe-Viewer zeigt die Auswertungsergebnisse an (Tabellen, Statistiken, Diagramme). Der Darstellungsbereich besteht aus zwei Teilen. Im linken Teil wird eine Gliederung aller bisher ausgeführten Analysen als Inhaltsübersicht aufgelistet; der rechte Teil enthält die Ergebnisse mit den Tabellen bzw. Graphiken. Der Ausgabe-Viewer sieht wie folgt aus:



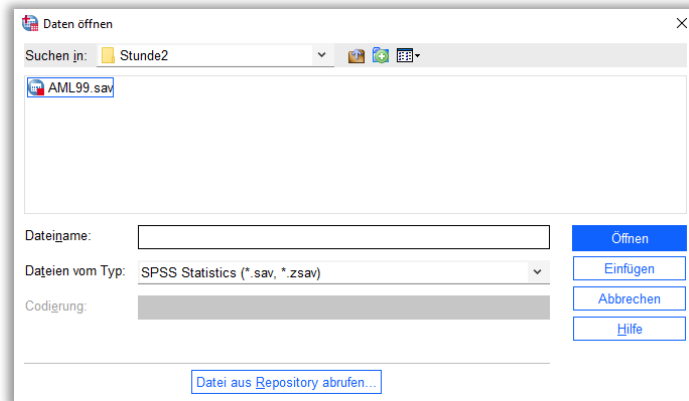
2. Datenaufbereitung

2.1 Öffnen und Speichern einer Datendatei

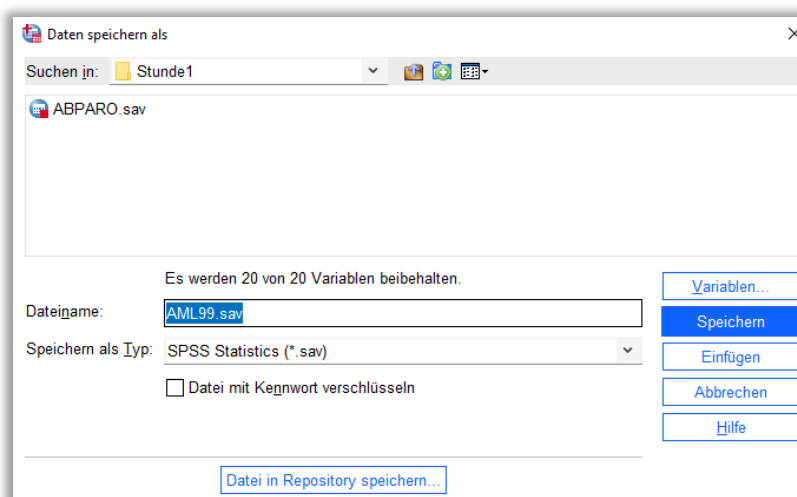
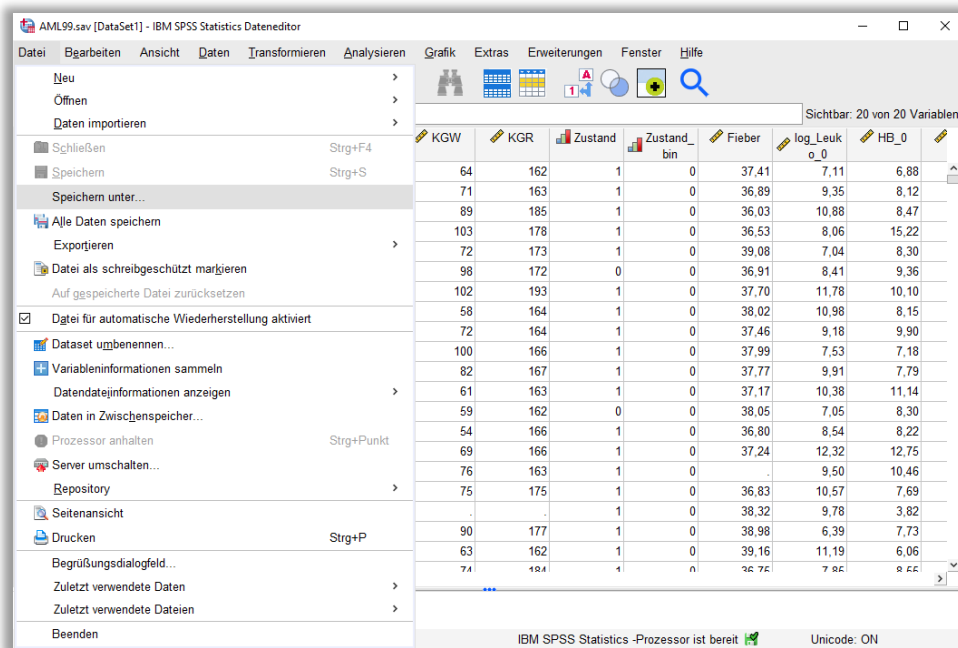
Starten Sie SPSS für Windows, indem Sie doppelt auf das SPSS-Symbol klicken. Es öffnet sich der SPSS-Daten-Editor. Um eine SPSS-Datendatei zu öffnen, betätigen Sie die Befehlsfolge **Datei > Öffnen > Daten**. Sie gelangen in das Dialogfeld *Datei öffnen*. Der voreingestellte Dateityp ist SPSS Statistics mit der Erweiterung **.sav**.



Durch Auswahl eines anderen Dateityps können auch Daten eingelesen werden, die in einem anderen Format vorliegen (z.B. Excel-Dateien).

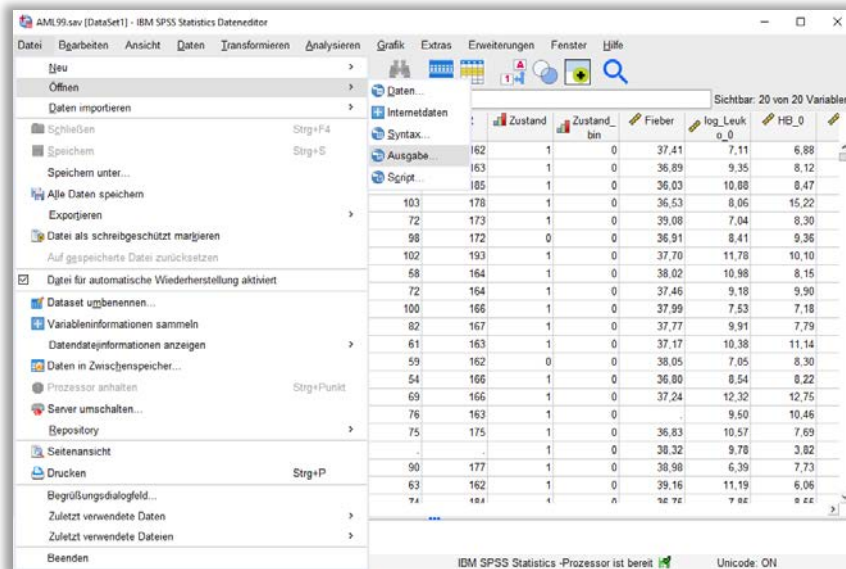


Um eine SPSS-Datei abzuspeichern betätigen Sie die Befehlsfolge **Datei > Speichern unter**. Im Dialogfeld *Daten speichern als Typ* können Sie die Datei mit der Endung **.sav** in Ihrem gewünschten Verzeichnis ablegen.

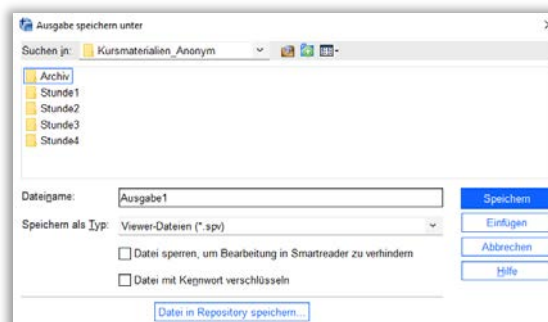
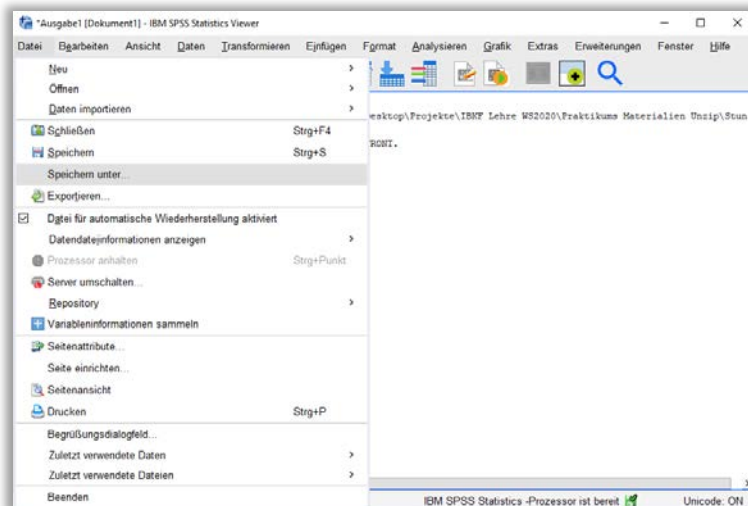


2.2 Öffnen und Speichern einer Ausgabedatei

Über die Befehlsfolge **Datei > Öffnen > Ausgabe** gelangen Sie in das Dialogfeld *Ausgabe öffnen*. Der Dateityp ist die Viewer-Datei mit der Erweiterung **.spv**.



Um den gesamten Inhalt des Ausgabefensters abzuspeichern, muss man sich im Ausgabefenster befinden. Betätigen Sie die Befehlsfolge **Datei > Speichern unter**. Sie gelangen in das Dialogfeld *Ausgabe speichern unter*. Der Dateityp ist die Viewer-Datei mit der Erweiterung **.spv**.



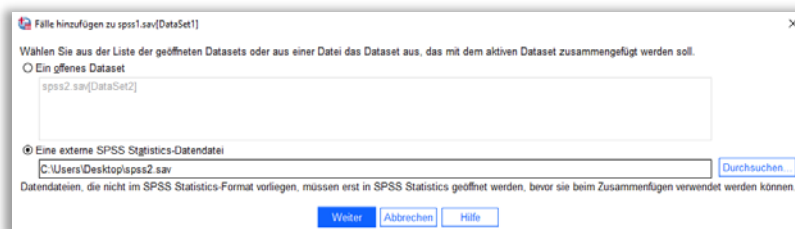
2.3 Zusammenfügen von Datendateien (fall- und variablenweise)

In SPSS besteht die Möglichkeit, zwei Datendateien (z.B. *spss1.sav* und *spss2.sav*) zusammenzuführen. Dabei gibt es zwei Möglichkeiten: fallweises Zusammenfügen und variablenweises Zusammenfügen.

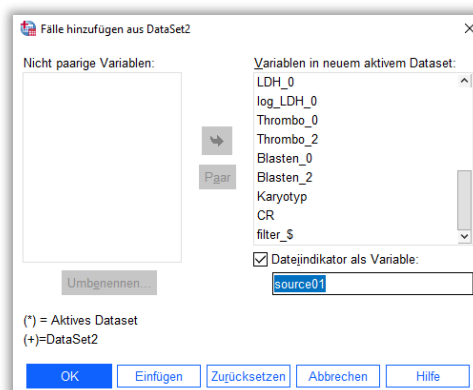
a) Fallweises Zusammenfügen

Zusammenfügen zweier Datendateien mit unterschiedlichen Fällen, deren Variablen identisch oder unterschiedlich sind.

Hierfür gehen Sie wie folgt vor: Öffnen Sie zunächst die Datei *spss1.sav*. Wählen Sie nun aus dem Menü die Befehlsfolge **Daten > Dateien zusammenfügen > Fälle hinzufügen**. Es öffnet sich die Dialogbox *Fälle hinzufügen zu spss1.sav*. Laden Sie die entsprechende zweite SPSS-Datei, aus der Fälle hinzugefügt werden sollen und gehen Sie auf **Weiter**.



Wenn die Variablen in beiden Dateien übereinstimmen, sind alle Variablen im Feld *Variablen in neuer Arbeitsdatei* aufgeführt und das Feld *Nicht gepaarte Variablen* bleibt leer. Um in der späteren zusammengefassten Datei zu kennzeichnen, welche Daten aus welcher Datei stammen, aktivieren Sie zusätzlich die Option *Datei-Indikator als Variable*. Die Variable *source01* als Datei-Indikator hat den Wert 0 für die zuerst geladene Datei *spss1.sav* und den Wert 1 für die hinzugefügte Datei *spss2.sav*. Bestätigen Sie mit **OK**.



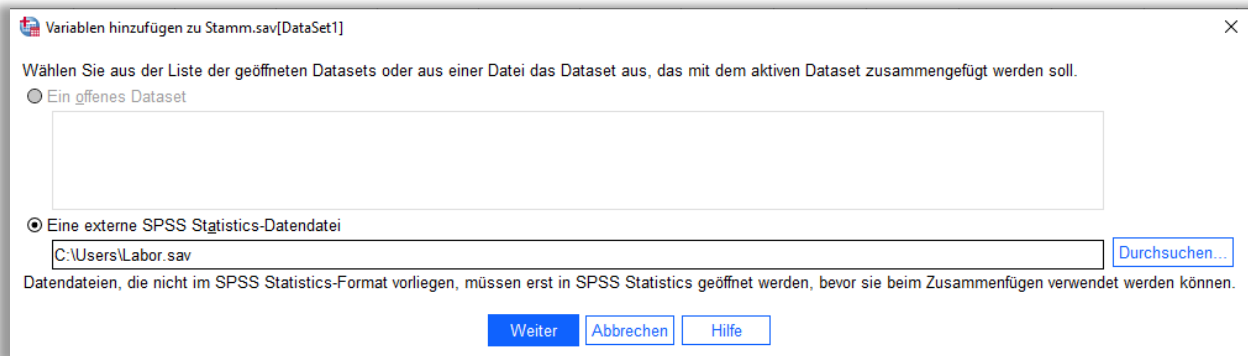
Die fertige zusammengefügte Datei hat folgende Struktur:

	PID	Therapie	Geschlecht	alter	KGW	KGR	Zustand	Zustand_bin	source01
1		1	2	60	64	162	1	0	0
2	P2	...	1	39	71	163	1	0	0
3	P3	...	1	71	89	185	1	0	0
4	P4	...	1	61	103	178	1	0	0
5	P6	...	1	62	98	172	0	0	0
6	P7	...	1	44	102	193	1	0	0
7	P8	...	1	60	58	164	1	0	0
8	P9	...	1	66	72	164	1	0	0
9	P10	...	1	61	100	166	1	0	0
10	P13	...	1	62	59	162	0	0	0
11	P16	...	1	66	76	163	1	0	0
12	P18	...	1	33	.	.	1	0	0
13	P19	...	1	43	90	177	1	0	0
14	P21	...	1	55	74	184	1	0	0
15	P24	...	1	30	72	176	0	0	0
16	P25	...	1	55	74	175	0	0	0
17	P26	...	1	64	100	176	1	0	0
18	P27	...	1	63	87	172	1	0	0
19	P28	...	1	41	98	184	0	0	0
20	P29	...	1	62	76	173	0	0	0

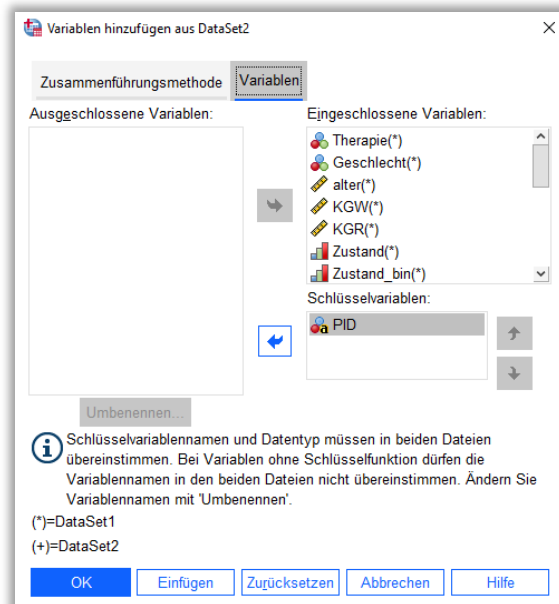
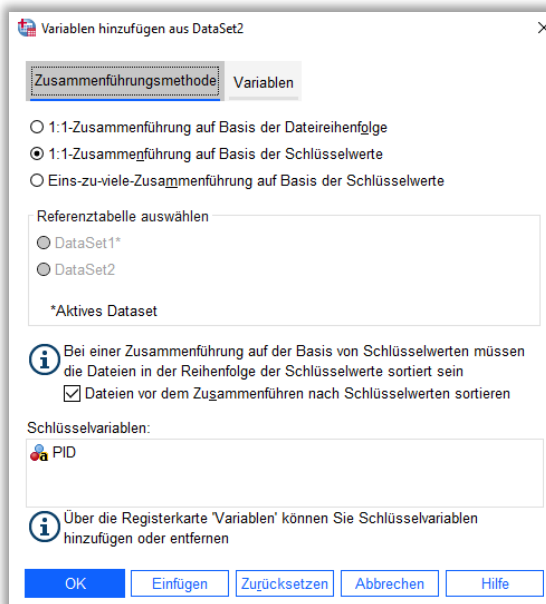
b) Variablenweises Zusammenfügen

Zusammenfügen zweier Datendateien mit gleichen Fällen, deren Variablen unterschiedlich sind, z.B. eine Datei enthält die Stammdaten der Patienten (*Stamm.sav*), die andere Datei die Laborparameter (*Labor.sav*) dieser Patienten. Empfehlenswert ist die Benutzung einer Schlüsselvariablen in beiden Dateien, da hier die Richtigkeit des Zusammenfügens beider Dateien sichergestellt ist.

Um beide Dateien zu einer Datei zusammenzufügen, gehen Sie wie folgt vor: Öffnen Sie zunächst die Datei *Stamm.sav* und wählen aus dem Menü die Befehlsfolge: **Daten > Dateien zusammenfügen > Variablen hinzufügen**. Es öffnet sich die Dialogbox *Variablen hinzufügen zu Stamm.sav*. Klicken Sie auf den Button **Durchsuchen** und wählen die Datei *Labor.sav* aus.



Klicken Sie auf den Button **Öffnen** und bestätigen Sie anschließend mit **Weiter**. Es öffnet sich die Dialogbox *Variablen hinzufügen aus....* Unter *Variablen* wählen Sie als Schlüsselvariablen Variablen, die in beiden Dateien enthalten sind (z.B. die ID des Patienten *PID*). Variablen, die in beiden Dateien auftreten und nicht zur Identifikation von Fällen dienen, können im Feld *Ausgeschlossene Variablen* vom Hinzufügen ausgeschlossen werden. Aktivieren Sie unter *Zusammenführungsmethode* die Option *1:1-Zusammenführung auf Basis der Schlüsselwerte*. Bestätigen Sie mit **OK**. (Vorsicht: Stellen Sie vorher sicher, dass das Häkchen bei *Dateien vor dem Zusammenführen nach Schlüsselwerten sortieren* gesetzt ist.

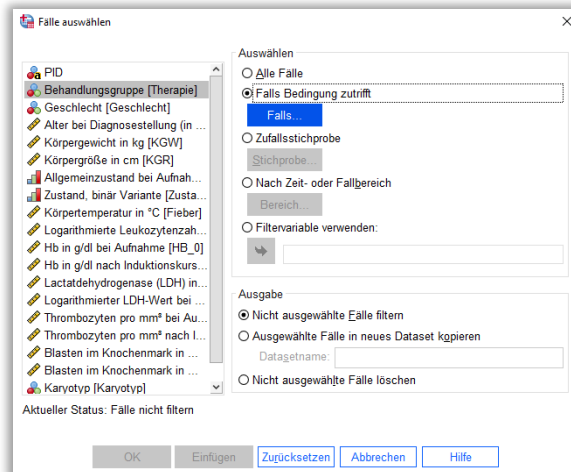


Die fertige, zusammengeführte Datei hat folgende Struktur:

3. Datenselektion

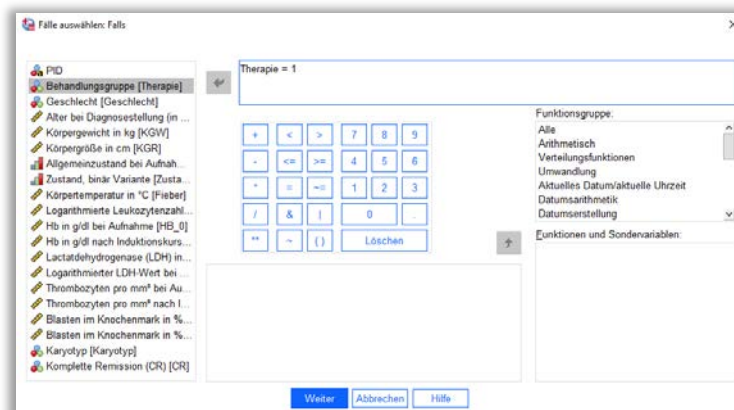
3.1 Auswahl von Fällen

Unter Datenselektion versteht man die Auswahl von Fällen nach bestimmten Kriterien. So sollen z.B. bei der Auswertung des AML-Datensatzes *AML99.sav* nur Patienten berücksichtigt werden, die der Behandlungsgruppe TAD-HAM (*Therapie* = 1) zugeteilt wurden. Dazu wählen Sie aus dem Menü die Befehlsfolge **Daten > Fälle auswählen > Falls Bedingung zutrifft > Falls**.



Es öffnet sich das Dialogfeld *Fälle auswählen: Falls*.

Übertragen Sie die Variable *Therapie* über die Pfeiltaste ⇌ in das rechte Dialogfeld. Nun kann der angeforderte logische Ausdruck über die Buttons der Rechenastatur in das dafür vorgesehene Dialogfeld geschrieben werden. Die Bedingung lautet *Therapie* = 1, d.h. es sollen alle Fälle ausgewählt werden, für welche die Variable *Therapie* den Wert »1« (TAD-HAM) hat. Die Dialogbox stellt sich nun wie folgt dar:

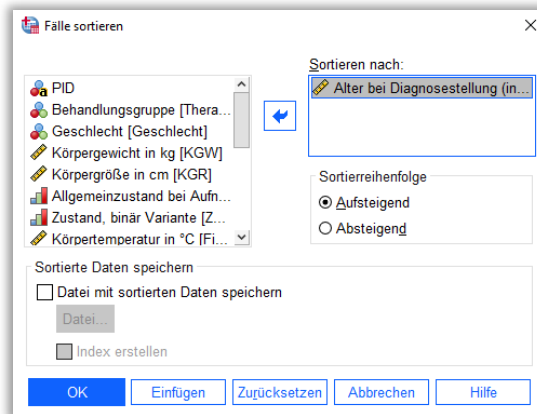


Bestätigen Sie mit **Weiter**. Belassen Sie die Voreinstellung *Nicht ausgewählte Fälle filtern*. Bestätigen Sie mit **OK**.

Sie gelangen zurück zum Dateneditor. Der Filter ist nun eingeschaltet, d.h. alle Patienten, die nicht die Therapie TAD-HAM bekommen haben, sind gefiltert und werden aus allen weiteren Analysen ausgeschlossen. Die Selektion ist erkennbar an den durchgestrichenen Fallnummern, sowie am unteren Rand des Daten-Editors: Meldung »Filter an«.

3.2 Fälle sortieren

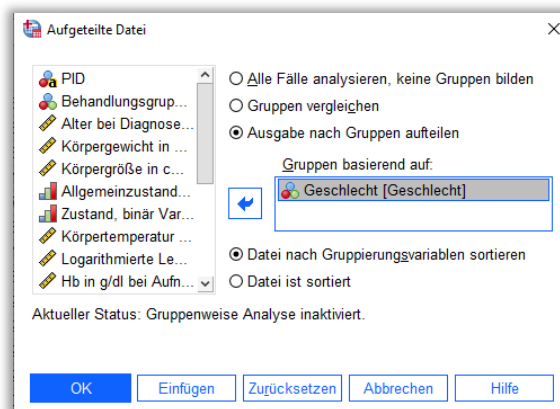
In SPSS können die Daten nach den Werten von einer oder mehreren Variablen in auf- oder absteigender Reihenfolge sortiert werden. Um z.B. die Daten der Datei *AML99.sav* nach dem Alter in aufsteigender Reihenfolge zu sortieren, gehen Sie wie folgt vor: **Daten > Fälle sortieren**. Es öffnet sich das Dialogfeld *Fälle sortieren*. Übertragen Sie die Variable *Alter* über die Pfeiltaste ⇨ in das rechte Dialogfeld *Sortieren nach*. Im Auswahlkasten *Sortierreihenfolge* ist die Sortierreihenfolge *Aufsteigend* vorgegeben. Bestätigen Sie mit **OK**.



Im Daten-Editor sehen Sie nun die Datendatei *AML99.sav* aufsteigend sortiert nach den Werten der Variable *Alter*.

3.3 Aufteilung der Fälle in Gruppen

Die Datenanalyse kann in SPSS getrennt nach Gruppen erfolgen. Eine solche Gruppenvariable kann z.B. die Variable *Geschlecht* sein. So können z.B. Häufigkeitstabellen nach den Merkmalen männlich und weiblich getrennt durchgeführt werden. Wählen Sie aus dem Menü **Daten > Datei aufteilen**. Es öffnet sich die Dialogbox *Datei aufteilen*. Aktivieren Sie die Option *Ausgabe nach Gruppen aufteilen* und bestätigen Sie mit **OK**. So erhalten Sie die Ausgabe nachfolgender Analysen getrennt nach Gruppen.



Geschlecht = Männlich**Statistiken^a**

Behandlungsgruppe

N	Gültig	546
	Fehlend	0

Behandlungsgruppe^a

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	TAD-HAM	284	52,0	52,0	52,0
	HAM-HAM	262	48,0	48,0	100,0
	Gesamt	546	100,0	100,0	

Geschlecht = Weiblich**Statistiken^a**

Behandlungsgruppe

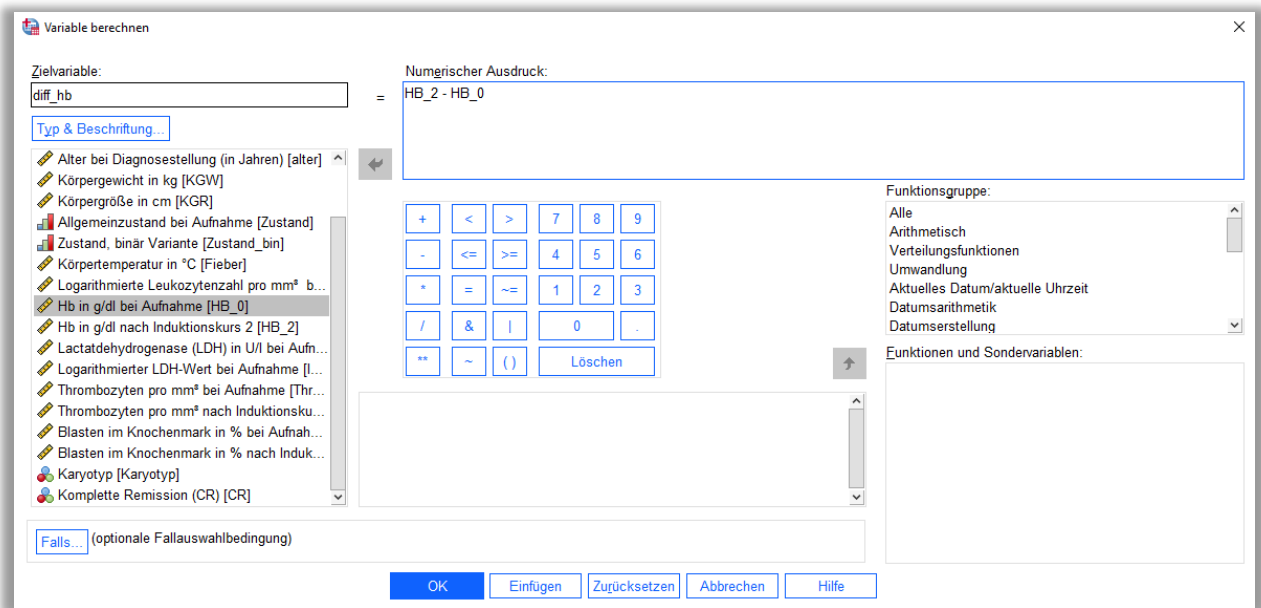
N	Gültig	461
	Fehlend	0

Behandlungsgruppe^a

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	TAD-HAM	243	52,7	52,7	52,7
	HAM-HAM	218	47,3	47,3	100,0
	Gesamt	461	100,0	100,0	

4. Datenmodifikation**4.1 Berechnung von neuen Variablen**

Es gibt in SPSS die Möglichkeit, neue Variablen zu berechnen und der Datendatei anzufügen. So sind etwa in der AML-Studie zu zwei Zeiten (vor und nach Therapie) Messungen des Hämoglobins erfolgt (Variablen *HB_0* und *HB_2*). Interessieren Sie sich für die Änderung der Hämoglobinkonzentration zwischen den beiden Zeitpunkten, so kann die Berechnung einer neuen Variable *diff_hb* von SPSS ausgeführt werden. Wählen Sie dazu im Menü **Transformieren > Berechnen**. Es öffnet sich die Dialogbox *Variable berechnen*. Im Feld *Zielvariable* wird der Name der Variablen eingegeben, die den neu berechneten Wert zugewiesen bekommt. Im Feld *Numerischer Ausdruck* wird der numerische Ausdruck eingetragen, der zur Berechnung des Wertes der Zielvariablen verwendet wird.



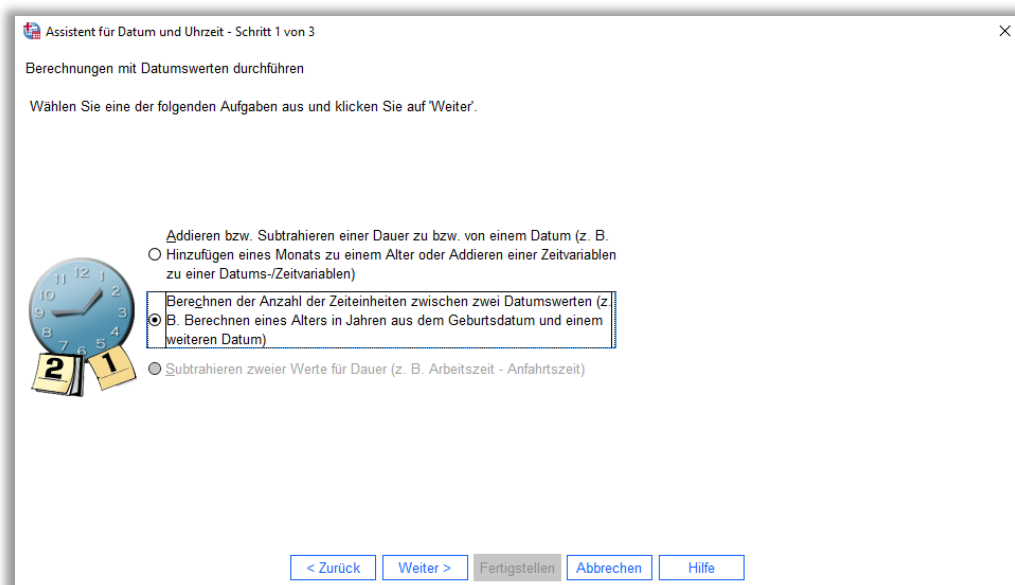
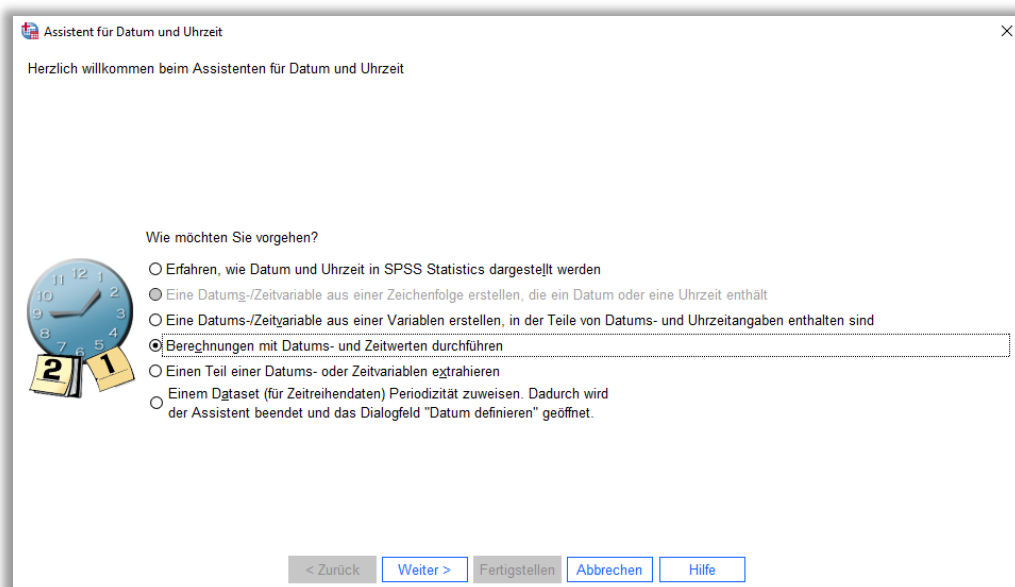
Tragen Sie im Feld *Zielvariable* den Namen *diff_hb* ein, im rechten Dialogfeld *Numerischer Ausdruck* die Formel »*HB_2* – *HB_0*«. Die Formel kann entweder per Hand eingegeben werden oder durch Benutzung der Variablenliste über die Pfeiltaste ⇨ und die Buttons im Feld mit der Rechentastatur. Bestätigen Sie mit **OK**. Sie finden die neue Variable *diff_hb* im Datenblatt hinten angefügt.

4.2 Rechnen mit Datumsangaben

In SPSS ist es möglich, die Differenz aus zwei Datumsvariablen, die bereits in der Datei vorkommen, zu berechnen. Z.B. möchten Sie bestimmen, wie alt jeder Patient zum Zeitpunkt der Untersuchung (in Jahren) ist. Es liegen dafür das Geburtsdatum (*birth_dt*) und der Untersuchungszeitpunkt (*accept_dt*) vor (siehe Abbildung unten):

patnr	birth_dt	accept_dt
1	10-Jan-1937	15-Nov-1967
2	02-Mar-1916	02-Jan-1968
3	19-Sep-1913	06-Jan-1968
4	23-Dec-1927	28-Mar-1968
5	28-Jul-1947	10-May-1968
6	08-Nov-1913	13-Jun-1968
7	29-Aug-1917	12-Jul-1968
8	27-Mar-1923	01-Aug-1968
9	11-Jun-1921	09-Aug-1968
10	09-Feb-1926	11-Aug-1968
11	22-Aug-1920	15-Aug-1968
12	09-Jul-1915	17-Sep-1968
13	22-Feb-1914	19-Sep-1968


Um das Alter jedes Patienten in Jahren zu berechnen, gehen Sie wie folgt vor: Wählen Sie im Menü die Befehlsfolge **Transformieren > Assistent für Datum und Uhrzeit**. Sie gelangen in das Dialogfeld *Assistent für Datum und Uhrzeit*. Klicken Sie hier auf **Berechnungen mit Datums- und Zeitwerten durchführen** und gehen auf **Weiter**. Sie gelangen ins nächste Dialogfeld *Schritt 1 von 3*. Aktivieren Sie hier den Befehl *Berechnen der Anzahl der Zeiteinheiten zwischen zwei Datumswerten*. Gehen Sie auf **Weiter**. Sie gelangen ins nächste Dialogfeld *Schritt 2 von 3*. Bringen Sie nun die Variable *accept_dt* über die Pfeiltaste ⇨ in das Feld *Datum1* und die Variable *birth_dt* in das Feld *minus Datum2*. Belassen Sie unter *Einheit* die Voreinstellung *Jahre*. Unter *Ergebnisbehandlung* belassen Sie die Voreinstellung *Auf ganze Zahl kürzen*. Gehen Sie auf **Weiter**. Sie gelangen ins nächste Dialogfeld *Schritt 3 von 3*. Unter *Ergebnisvariable* geben Sie den Namen der neuen Variable ein (z.B. *alter*). Optional können Sie noch ein Variablenlabel vergeben (z.B. *Alter in Jahren zum Zeitpunkt der Untersuchung*). Bestätigen Sie Ihre Eingaben durch **Fertigstellen**.



Assistent für Datum und Uhrzeit - Schritt 2 von 3

Berechnung der Anzahl der Zeiteinheiten zwischen zwei Datums- oder Datums-/Zeitvariablen.

Das Ergebnis ist eine ganzzahlige Variable. Alle Bruchteile einer Einheit werden verworfen. Das Ergebnis ist eine Variable für die Dauer. In der unten stehenden Variablenliste werden nur Variablen für die Dauer angezeigt.



Variablen:

Aktuelles Datum und aktuelle Uhrzeit [...]

Datum1:

acceptance into program [accept_dt]

minus Datum2:

birth date [birth_dt]

Einheit:

Jahre

Ergebnisbehandlung

☒ Auf ganze Zahl kürzen

☐ Auf ganze Zahl runden

☐ Bruchteil beibehalten


Für Einheiten von Monat und Jahr basiert das Ergebnis auf durchschnittlichen Einheitenlängen, es sei denn, es wird abgeschnitten

\$TIME ist das aktuelle Datum und die aktuelle Uhrzeit.

< Zurück Weiter > Fertigstellen Abbrechen Hilfe

Assistent für Datum und Uhrzeit - Schritt 3 von 3

Berechnung: accept_dt - birth_dt



Ergebnisvariable:

alter

Einheiten:

Jahre

Variablenbeschriftung:

Alter in Jahren zum Zeitpunkt der Untersuchung

Ausführung

☒ Variable jetzt erstellen ☐ Syntax in Syntaxfenster einfügen

< Zurück Weiter > Fertigstellen Abbrechen Hilfe

Als Ergebnis erscheint folgende Datendatei mit der neu berechneten Variable *alter*, die an den bisherigen Datensatz hinten angefügt wurde:

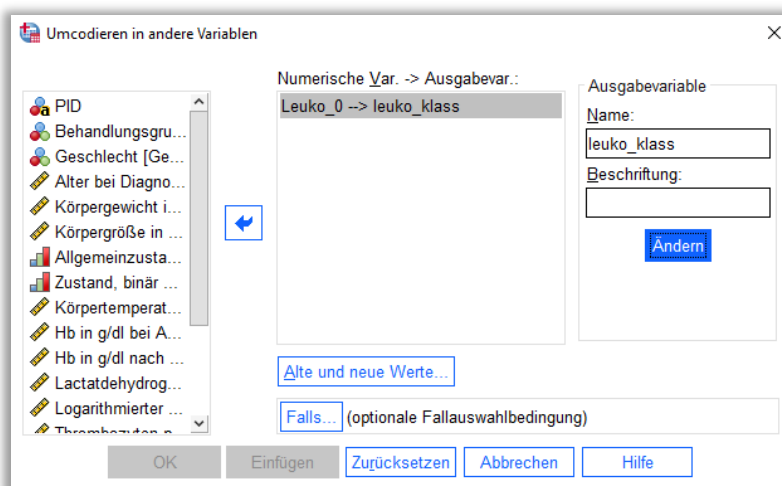
patnr	birth_dt	accept_dt	alter
1	10-Jan-1937	15-Nov-1967	30
2	02-Mar-1916	02-Jan-1968	51
3	19-Sep-1913	06-Jan-1968	54
4	23-Dec-1927	28-Mar-1968	40
5	28-Jul-1947	10-May-1968	20
6	08-Nov-1913	13-Jun-1968	54
7	29-Aug-1917	12-Jul-1968	50
8	27-Mar-1923	01-Aug-1968	45
9	11-Jun-1921	09-Aug-1968	47
10	09-Feb-1926	11-Aug-1968	42
11	22-Aug-1920	15-Aug-1968	47
12	09-Jul-1915	17-Sep-1968	53
13	22-Feb-1914	19-Sep-1968	54

4.3 Umkodieren und Klassieren von Werten

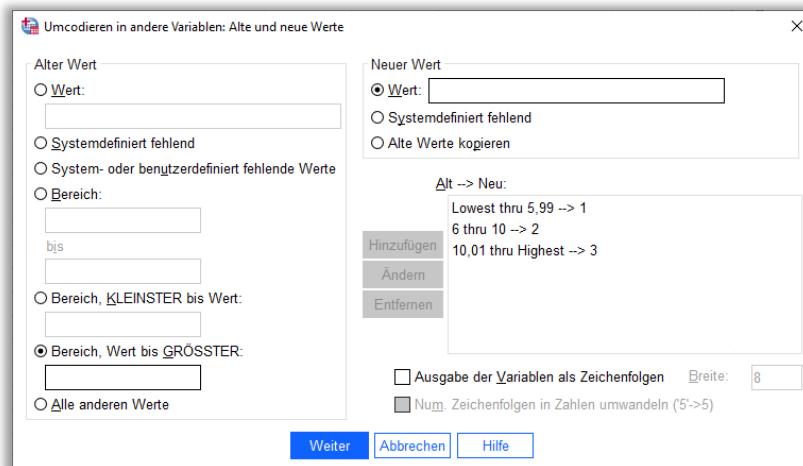
In SPSS können ursprünglich erfasste Daten umkodiert werden. Eine Umkodierung numerischer Daten ist z.B. dann sinnvoll, wenn die original erfassten Rohdaten für die weitere Analyse im folgenden Sinn vereinfacht werden sollen. Man möchte den Wert der Leukozyten bei Aufnahme in 3 Klassen einteilen und so die Leukozytenklasse jedes Patienten bestimmen. Die Klassierung soll folgendermaßen aussehen:

Leukozytenwert	Klasse	Wertelabel
< 6	1	Leukopenie
6-10	2	Normbereich
> 10	3	Leukozytose

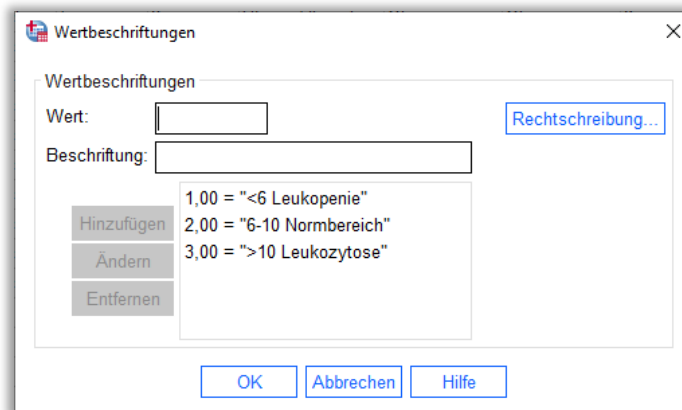
Um obiges Verfahren in SPSS umzusetzen, gehen Sie wie folgt vor: Über die Befehlsfolge **Transformieren > Umkodieren in andere Variablen** gelangen Sie in das Dialogfeld *Umkodieren in andere Variablen*. Übertragen Sie die ursprüngliche Variable *Leuko_0* über die Pfeiltaste ⇨ in das Feld *Numerische Var. -> Ausgabevar.* Anschließend geben Sie der neu zu erstellenden Ausgabevariable einen Namen im Feld *Ausgabevariable* (z.B. *leuko_klass*). Danach klicken Sie auf den Button **Ändern**. So wird der Name der neuen Ausgabevariable festgelegt (*leuko_klass*). Betätigen Sie nun den Button **Alte und neue Werte**.



Sie gelangen in das Dialogfeld *Umkodieren in andere Variablen: Alte und neue Werte*. Im linken Fensterbereich unter *Alter Wert* geben Sie bei *Bereich, KLEINSTER bis Wert* den Wert »5,99« ein. Im rechten Fensterbereich geben Sie unter *Neuer Wert* den Wert »1« ein und bestätigen Ihre Eingaben durch den Button **Hinzufügen**. Dem *Bereich* »6« bis »10« wird der neue Wert »2« zugeordnet. Und im Bereich *Wert bis GRÖSSTER* tragen Sie »10,01« ein und weisen diesem Bereich den Wert »3« zu.



Bestätigen Sie die Eingaben mit **Weiter** und **OK**. In der Datendatei wurde nun die neue umkodierte Variable *leuko_klass* hinten angefügt. Zusätzlich können in der Variablenansicht **Wertelabels** der neuen Variable vergeben werden.

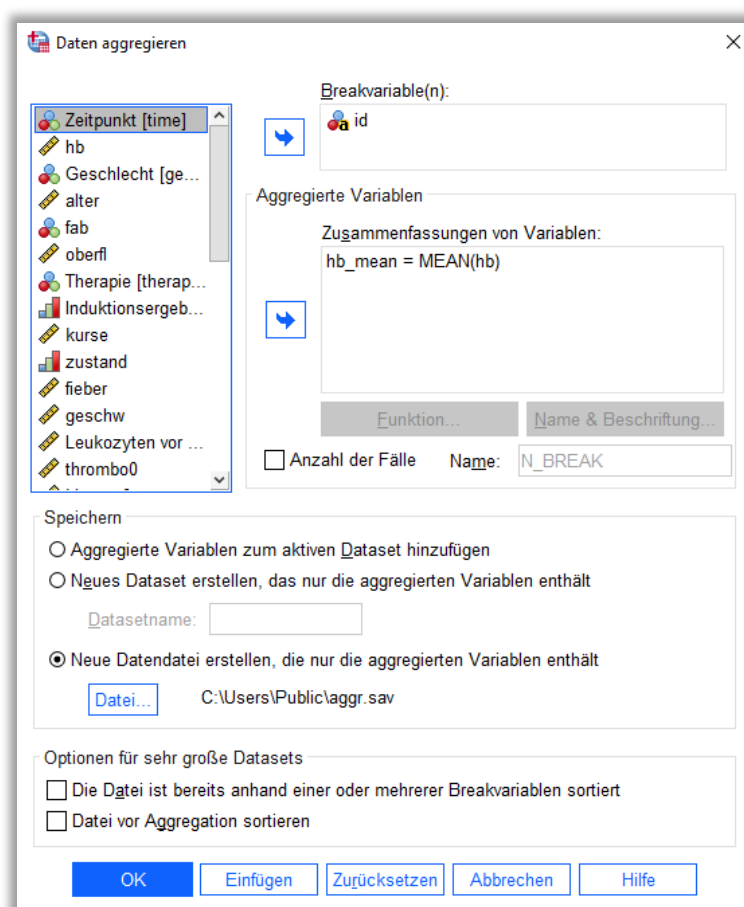


4.4 Aggregieren von Daten

Basierend auf den Werten von einer oder mehreren Gruppenvariablen (Break-Variablen) können Fälle zusammengefasst (aggregiert) und eine neue Datendatei erstellt werden, die für jede Break-Gruppe einen Fall enthält. Z.B. wurde in der AML99-Studie der Hb-Wert (in g/dl) jedes Patienten insgesamt drei Mal erfasst: bei Aufnahme sowie nach dem 1. und 2. Induktionskurs. Die drei Werte sind in der Datendatei als drei Fälle untereinander aufgeführt:

	id	time	hb
1	P1202	1	3,40
2	P1202	2	.
3	P1202	3	.
4	P1204	1	8,60
5	P1204	2	8,40
6	P1204	3	.
7	P1205	1	6,60
8	P1205	2	.
9	P1205	3	9,00
10	P1207	1	8,50
11	P1207	2	10,00
12	P1207	3	.
13	P1208	1	10,80
14	P1208	2	12,20
15	P1208	3	8,80

Es soll für jeden Patienten der mittlere Hb-Wert über alle drei Messzeitpunkte bestimmt und eine neue Datendatei erstellt werden, die nur die aggregierten Variablen enthält. Wählen Sie dazu **Daten > Aggregieren**. Es öffnet sich das Dialogfeld *Daten aggregieren*. Übertragen Sie die Variable *id* in das Feld *Breakvariable(n)*. Übertragen Sie die Variable *hb* in das Feld *Zusammenfassung von Variablen*. Als Aggregierungsfunktion ist der Mittelwert bereits voreingestellt. Alternativ können Sie über den Button **Funktion** andere Kenngrößen anfordern. Wenn die aggregierten Daten in einer neuen Datei gespeichert werden sollen, markieren Sie die Option *Neue Datendatei erstellen, die nur die aggregierten Variablen enthält*. Über den Button **Datei** können Sie einen neuen Dateinamen eingeben und die Datei in einem Verzeichnis Ihrer Wahl ablegen. Bestätigen Sie Ihre Wahl mit **Speichern**. Anschließend sollte die Dialogbox wie folgt aussehen:



Bestätigen Sie Ihre Einstellungen mit **OK**. Die neu erzeugte Datei enthält nun die neu gebildete aggregierte Variable *hb_mean* mit dem mittleren Hb-Wert pro Patient über alle 3 Messzeitpunkte:

	id	hb_mean
1	P1202	3,40
2	P1204	8,50
3	P1205	7,80
4	P1207	9,25
5	P1208	10,60