

Datendateien umstrukturieren

SPSS-Datendateien folgen in Ihrem Aufbau einer klaren Struktur: Die Werte sind in Zeilen und Spalten angeordnet. Die Zeilen enthalten die Fälle (oder Beobachtungen), die Spalten die Variablen (oder Merkmale zu den Fällen). Für manche Analysen kann es notwendig werden, von dieser Anordnung abzuweichen und die Datenwerte anders strukturiert abzulegen. Dazu ist keine „Handarbeit“ notwendig, da SPSS mit dem Menüpunkt *Daten > Umstrukturieren* einen passenden Befehl zur Verfügung stellt, der im weiteren Verlauf näher erläutert wird.

Das folgende Beispiel greift auf die drei Variablen *Krankheit_Akut*, *Krankheit_Reha* sowie *Krankheit_6 Monate* der SPSS-Beispieldatei *Krankheitsverlauf.sav* zu. Hier der entsprechende Ausschnitt aus der Variablenansicht:

Name	Typ	Spaltenfo...	Dezimals...	Variablenlabel	Wertelabels	Fehlende W...	Spalten
Krankheit_Akut	Numerisch	8	0	Krankheitsverlauf bei Entlassung aus Krankenhaus (Akutversorgung ohne Reha)	Keine	Keine	8
Krankheit_Reha	Numerisch	8	0	Krankheitsverlauf nach Entlassung (Reha)	Keine	Keine	8
Krankheit_6Monate	Numerisch	8	0	Krankheitsverlauf nach 6 Monaten	Keine	Keine	8

Die drei Variablen *Krankheit_Akut*, *Krankheit_Reha* sowie *Krankheit_6 Monate* sollen in eine Variablen überführt werden. Um die drei Ausgangsvariablen später noch rekonstruieren zu können, wird eine zweite neue Variable als Indexvariable benötigt.

Ausgangsdatei

	Krankheit_Akut	Krankheit_Reha	Krankheit_6Monate	v
1	5	5	5	
2	5	5	5	
3	3	3	1	
4	5	5	5	
5	4	3	2	
6	4	5	5	
7	5	5	5	
8	5	5	5	
9	5	5	5	
10	4	4	5	
11	5	5	5	
12	5	5	5	
13	5	5	5	
14	5	5	5	
15	5	5	5	
16	3	3	.	
17	3	2	1	
18	4	4	5	
19	5	5	5	
20	5	5	5	
21	3	3	2	
22	3	2	.	
23	5	5	5	
24	5	5	5	
25	4	4	.	
26	3	2	.	

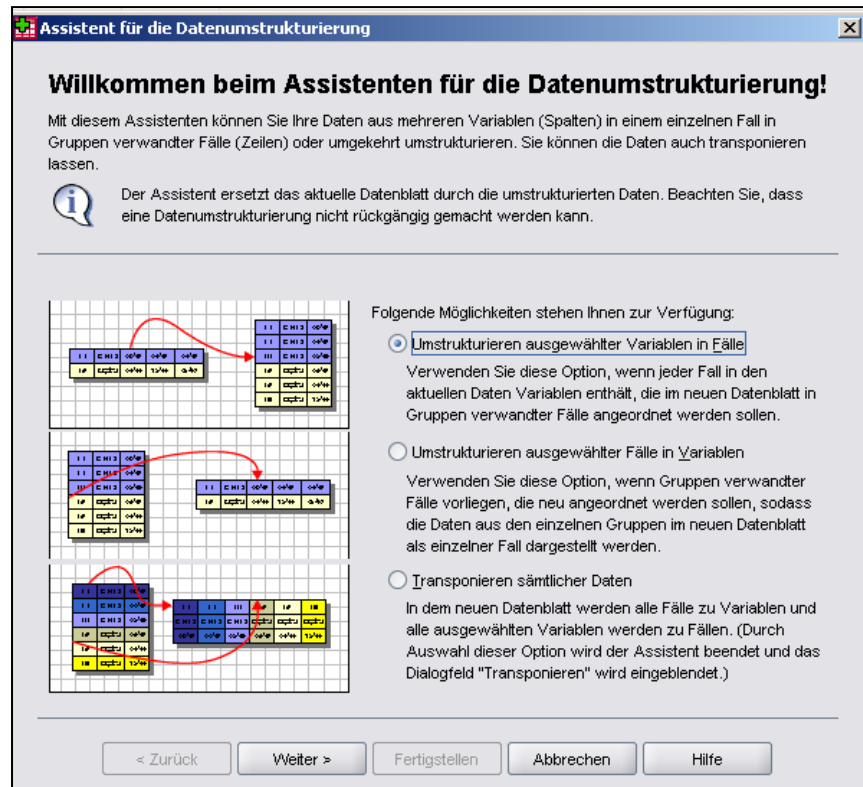
Zieldatei

	id	Geschlecht	Index1
1	1	weiblich	Krankheit_Akut
2	1	weiblich	Krankheit_Reha
3	1	weiblich	Krankheit_6Monate
4	2	männlich	Krankheit_Akut
5	2	männlich	Krankheit_Reha
6	2	männlich	Krankheit_6Monate
7	3	männlich	Krankheit_Akut
8	3	männlich	Krankheit_Reha
9	3	männlich	Krankheit_6Monate
10	4	weiblich	Krankheit_Akut
11	4	weiblich	Krankheit_Reha
12	4	weiblich	Krankheit_6Monate
13	5	weiblich	Krankheit_Akut
14	5	weiblich	Krankheit_Reha
15	5	weiblich	Krankheit_6Monate
16	6	männlich	Krankheit_Akut
17	6	männlich	Krankheit_Reha
18	6	männlich	Krankheit_6Monate
19	7	männlich	Krankheit_Akut
20	7	männlich	Krankheit_Reha
21	7	männlich	Krankheit_6Monate
22	8	weiblich	Krankheit_Akut
23	8	weiblich	Krankheit_Reha
24	8	weiblich	Krankheit_6Monate
25	9	weiblich	Krankheit_Akut
26	9	weiblich	Krankheit_Reha
27	9	weiblich	Krankheit_6Monate

Da der Befehl Ihre aktuelle Arbeitsdatei mit den umstrukturierten Daten überschreibt, sollten Sie vor dem Menüaufruf Ihre Arbeitsdatei abspeichern.

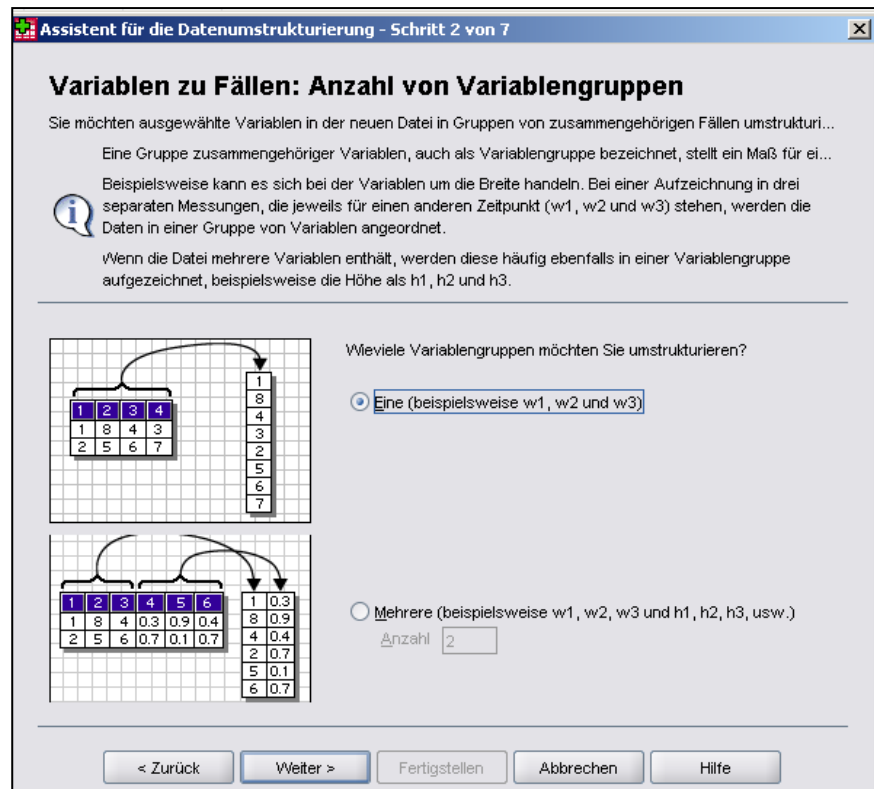
Danach rufen Sie in SPSS 17.0 über das Menü *Daten* > *Umstrukturieren* auf. Ein Assistent fragt nun in sieben Schritten ab, in welcher Art und Weise die geöffnete Datendatei umstrukturiert werden soll.

In unserem Beispiel erfolgt das *Umstrukturieren ausgewählter Variablen in Fälle*.



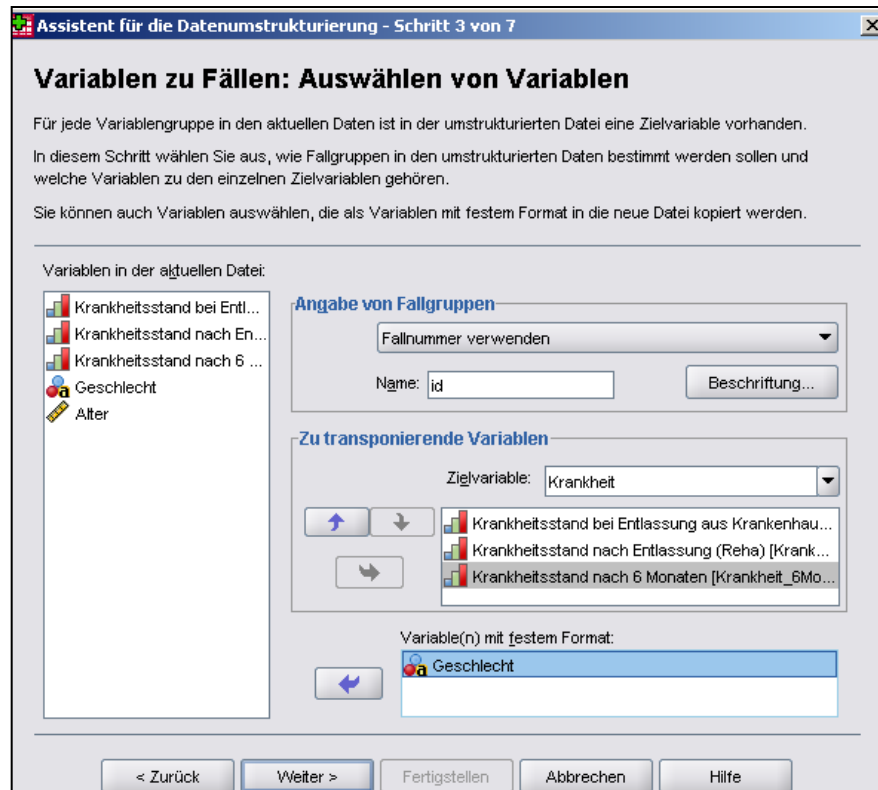
Im nächsten Schritt, Schritt 2 von 7, wird die Anzahl der Variablengruppen bestimmt. Damit legen Sie fest, wie viele neue Variablen gebildet werden.

In unserem Fall ist eine Variablengruppe auszuwählen, da drei Variablen in eine neue Variable umstrukturiert werden sollen.



Im dritten Schritt wird die Variablen­gruppe festgelegt, d.h. die Variablen ausgewählt, die umstrukturiert werden sollen. SPSS bezeichnet das Umstrukturieren mit Transponieren.

Die Dropdown-Liste enthält im vorliegenden Beispiel einzig die drei zu transponierenden Variablen *Krankheit_Akut*, *Krankheit_Reha* sowie *Krankheit_6 Monate*, die ausgewählt werden. Die Zielvariable, also die neu zu bildende Variable, erhält standardmäßig den Namen *trans1*. Sie können dort aber auch einen Namen frei wählen, etwa *Krankheit*, wie in unserem Beispiel.



Die Option *Angabe von Fallgruppen* liefert Informationen über Werte, die in der Ursprungstabelle demselben Fall angehören. Im vorliegenden Beispiel wird die Angabe *nach Fallnummer* verwendet, da sich in der Ursprungstabelle keine weitere Variable befindet, die zur Kennzeichnung der Fälle geeignet ist. Somit wird die Zieltabelle mit der Variable *id* ergänzt, die für jeden Fall die Nummer des Ursprungsfalles aus der Ausgangstabelle angibt.

Desweiteren können in diesem Dialogfenster auch Variablen mit festem Format, d.h. Variablen, die nicht transponiert, aber in die Zieltabelle übernommen werden sollen, ausgewählt werden. Im vorliegenden Beispiel enthält die Ausgangstabelle die Variable *Geschlecht*, die mit keiner anderen Variable zusammengefasst wird und keiner Variablen­gruppe angehört. In der Ausgangstabelle wird durch jeden Fall das Geschlecht der Person in der Variable *Geschlecht* angegeben. Damit auch in der Zieltabelle das Geschlecht der Person in drei allen Fällen ausgewiesen wird, sollen deren Werte in die Zieldatei übernommen werden.

Im vierten Schritt erfolgt die Auswahl der Indexvariablen. Die Indexvariable zeigt in der Ziel-tabelle an, ob sich die Werte der Fälle auf den Wert der ersten, zweiten oder dritten Ursprungsvariable beziehen. In unserem Beispiel wird eine Indexvariable erstellt.

Assistent für die Datenumstrukturierung - Schritt 4 von 7

Variablen zu Fällen: Erstellen von Indexvariablen

Bei den aktuellen Daten werden die Werte für eine Variablen-Gruppe in einem einzigen Fall in mehreren Variablen angezeigt. So enthält beispielsweise ein einzelner Fall die Werte für w1, w2 und w3.

Bei den neuen Daten werden die Werte für eine Variablen-Gruppe in mehreren Fällen in einer einzelnen Variablen angezeigt. So gibt es beispielsweise drei Fälle, jeweils einen für w1, w2 und w3.

Ein Index ist eine neue Variable, die eine Gruppe neuer Fälle angibt, die aus dem ursprünglichen Fall erstellt wurde. So hätte beispielsweise ein Index "w" die Werte 1, 2 und 3.

1	1	1	0.07
1	1	2	0.11
1	1	3	0.05
2	1	1	0.08
2	1	2	0.04
2	1	3	0.06

1	1	1	1	0.07
1	1	1	2	0.11
1	1	1	3	0.05
1	1	2	1	0.08
1	1	2	2	0.04
1	1	2	3	0.06

1	1	0.08	2	0.07
2	1	0.11	2	0.11
3	1	0.07	2	0.05
4	1	0.06	2	0.08
5	1	0.09	2	0.04
6	1	0.02	2	0.06

Wieviele Indexvariablen möchten Sie erstellen?

Eine
Verwenden Sie diese Option, wenn eine Variablen-Gruppe die Effekte eines einzelnen Faktors, einer einzelnen Behandlung oder einer einzelnen Bedingung aufzeichnet.

Mehrere

Verwenden Sie diese Option, wenn eine Variablen-Gruppe die Effekte mehrerer Faktoren, Behandlungen oder Bedingungen aufzeichnet.

Keine
Verwenden Sie diese Option, wenn die Indexinformationen in einem dieser zu transponierenden Variablen-Sets gespeichert sind.

< Zurück
Weiter >
Fertigstellen
Abbrechen
Hilfe

Im fünften Schritt wird diese Indexvariable näher beschrieben. Die Indexvariable kann entweder fortlaufende Zahlen oder die Namen der Ursprungsvariablen enthalten. Zusätzlich kann der Name und das Label festgelegt werden.

Für unser Beispiel haben wir die Variablennamen gewählt, Index und Label aber unverändert belassen.

Assistent für die Datenumstrukturierung - Schritt 5 von 7

Variablen zu Fällen: Erstellen einer Indexvariablen

Sie haben ausgewählt, dass genau eine Indexvariable erstellt werden soll. Bei den Werten der Variablen kann es sich um fortlaufende Zahlen oder die Namen der Variablen in einer Gruppe handeln.

In der Tabelle können Sie Namen und Label für die Indexvariable angeben.

Art des Indexwerts:

Fortlaufende Zahlen
Indexwerte: 1, 2, 3

Variablennamen
Indexwerte: Krankheit_Akut, Krankheit_Reha, Krankheit_6Monate

Name und Label der Indexvariablen bearbeiten:

	Name	Variablenlabel	Stufen	Indexwerte
1	Index1			3 Krankheit_Akut, Krankh...

< Zurück
Weiter >
Fertigstellen
Abbrechen
Hilfe

Im folgenden Dialogfenster (Schritt 6) werden weitere Optionen festgelegt, z.B. wie mit weiteren Variablen in der Datendatei umzugehen ist, wie fehlende Werte zu behandeln sind. Für unser Beispiel genügen die Voreinstellungen.

Assistent für die Datenumstrukturierung - Schritt 6 von 7

Variablen zu Fällen: Optionen

In diesem Schritt können Sie Optionen festlegen, die auf die umstrukturierte Datendatei angewendet werden.

Verarbeitung nicht ausgewählter Variablen

Variable(n) aus neuer Datendatei entfernen

Beibehalten und als Variable(n) mit festem Format behandeln

System Missing (fehlender Wert) oder leere Werte in allen transponierten Variablen

Einen Fall in der neuen Datei erstellen

Daten verwerfen

Variable zum Zählen von Fällen

Anzahl neuer Fälle zählen, die vom Fall in den aktuellen Daten erstellt wurden

Name:

Beschriftung:

< Zurück Weiter > Fertigstellen Abbrechen Hilfe

Im letzten und siebten Schritt kann ausgewählt werden, ob die Umstrukturierung der Daten unmittelbar durchgeführt oder die Syntax mit den zuvor vorgenommen Einstellungen in eine Syntaxdatei geschrieben werden soll.

Assistent für die Datenumstrukturierung - Fertigstellen

Fertigstellen

Folgende Möglichkeiten stehen Ihnen zur Verfügung:

Daten jetzt umstrukturieren

Empfiehlt sich, wenn die aktuelle Datei sofort ersetzt werden soll.

Vom Assistenten erzeugte Syntax in ein Syntaxfenster einfügen

Verwenden Sie diese Option, wenn Sie die Syntax vor dem Umstrukturieren der Daten speichern oder ändern möchten.

< Zurück Weiter > Fertigstellen Abbrechen Hilfe

Sobald Fertigstellen angeklickt wird, erfolgt die Umstrukturierung und die bereits oben gezeigte Zielfeld mit folgender Variablenansicht wird erzeugt:

	Name	Typ	Spaltenfo...	Dezimal...	Variablenlabel	Wertelabels	Fehlende W...	Spalten	Ausrichtung	Messniveau
1	id	Numerisch	8	0		Keine	Keine	8	Rechtsbündig	Nominal
2	Geschlecht	String	1	0		{1, männlich...	Keine	8	Linksbündig	Nominal
3	Index1	String	17	0		Keine	Keine	17	Linksbündig	Nominal
4	Krankheit	Numerisch	8	0	Krankheitsstan...	Keine	Keine	8	Rechtsbündig	Ordinal