

## Auswertungsbeispiele: Staatsangehörigkeit der Haushaltsbezugsperson

Für zahlreiche Auswertungen ist es notwendig, Angaben zu einer bestimmten Person in einer Einheit allen Personen dieser Einheit zuzuspielen. Ein Beispiel ist die Staatsangehörigkeit der Haushaltsbezugsperson.

In Form so genannter Bandsatzerweiterungen sind solche Variablen häufig bereits in den Daten des Mikrozensus enthalten. Sie werden vom Statistischen Bundesamt generiert. Hier wird nun gezeigt, wie solche Variablen gebildet werden, um entweder die im Datensatz bereits enthaltenen Variablen zu rekonstruieren und so eventuell vorhandene unplausible Angaben zu korrigieren, oder um eigene Variablen zu generieren.

Die Staatsangehörigkeit der Haushaltsbezugsperson kann folgendermaßen auf alle Personen des jeweiligen Haushalts übertragen werden (hier am Beispiel des Mikrozensus 2003):

### Statistik-Programm-Routinen

#### SPSS

/\* Es gibt in SPSS mehrere Möglichkeiten die Konstruktion von Bandsatzerweiterungen umzusetzen.\*/

#### Option 1: COMPUTE

`SORT CASES BY EF3 EF4.`

`COMPUTE X52 = EF52 * (EF507=1 & EF506=1).  
AGGREGATE OUTFILE = * MODE=ADDVARIABLES  
/PRESORTED  
/BREAK EF3 EF4  
/V559 "Staatsangeh. HH-Bezugsperson" = sum(X52).`

`DELETE VARIABLES X52.`

#### Option 2: MATCH FILES

`SELECT IF EF506=1.  
SORT CASES BY EF3 EF4.  
SELECT IF EF507=1.`

`SAVE OUTFILE='<Pfadangabe>/V559.sav'  
/ KEEP=EF3 EF4 EF52  
/ RENAME EF52=V559.  
MATCH FILES TABLE='<Pfadangabe> /V559.sav'  
/ FILE=*  
/ BY EF3 EF4.  
EXECUTE.`

#### Stata

/\* Es gibt in STATA mehrere Möglichkeiten die Konstruktion von Bandsatzerweiterungen umzusetzen.\*/

#### Option 1: egen

`egen v559 = max(ef52*(ef506==1 & ef507==1)), by(ef3 ef4)`

#### Option 2: merge

`sort ef3 ef4  
save Laufwerk:\<Padangabe>\filename1.dta, replace  
sort ef3 ef4  
generate v559 = ef52 if ef506==1 & ef507==1  
collapse(sum)v559, by(ef3 ef4)  
sort ef3 ef4  
save Laufwerk:\<Padangabe>\filename2.dta, replace  
use Laufwerk:\<Padangabe>\filename1.dta, clear  
merge ef3 ef4 using ///  
Laufwerk:\<Padangabe>\filename2.dta, uniquising`

### Erläuterungen

#### SPSS

##### Option 1: COMPUTE

Zunächst wird der Datensatz nach den zur Abgrenzung von Haushalten notwendigen Variablen (EF3, EF4) sortiert. Daraufhin wird das interessierende Merkmal (EF52) für die entsprechende Person (EF506=1 & EF507=1) in der Hilfsvariablen X52 definiert und auf Haushaltsebene aggregiert. Die Hilfsvariable (X52) kann nun wieder gelöscht werden.

##### Option 2: MATCH FILES

Zunächst erfolgt eine Einschränkung der interessierenden

#### Stata

##### Option 1: egen

Mit *egen* wird die Variable v559 (Staatsangehörigkeit der Haushaltsbezugsperson) gebildet, welche die bereits vorhandene Bandsatzerweiterung ef559 replizieren soll. Die Funktion *max(exp)* führt nun dazu, dass die Ausprägung der Staatsangehörigkeit (ef52) der Haushaltsbezugsperson (ef507==1) den Personen des Haushaltes in Privathaushalten (ef506==1) zugespielt wird. Die Sortierung des Datensatzes nach den zur Abgrenzung von Haushalten notwendigen Variablen (ef3, ef4) erfolgt in den Optionen nach dem Komma.

Population auf die Bevölkerung in Privathaushalten (EF506=1). Der Datensatz wird nach den zur Abgrenzung von Haushalten notwendigen Variablen (EF3, EF4) sortiert. Dann wird die Staatsangehörigkeit (EF52) der Haushaltsbezugsperson (EF507=1) umbenannt (in V559) und in einen neuen Datensatz übertragen. Dieser Datensatz wird dem Ausgangsdatsatz zugespielt, so dass jede Person eines Haushalts die Angabe der Staatsangehörigkeit der jeweiligen Haushaltsbezugsperson (V559) erhält.

#### Option 2: merge

Zunächst wird der Datensatz nach den zur Abgrenzung von Haushalten notwendigen Variablen (ef3, ef4) sortiert und daraufhin gespeichert. Es folgt die Abgrenzung auf Individualebene, d.h. auf Privathaushalte (ef506==1) und die Haushaltsbezugsperson (ef507==1). Die Aggregation der Beobachtungen auf Haushaltsebene erfolgt mittels *collapse*. Anschließend wird erneut sortiert und gespeichert. Das Individualdatenfile (filename1.dta) wird nun eingelesen und mittels *merge* mit dem Aggregatdatenfile (filename2.dta) zusammengeführt.

#### Weitere Erläuterungen finden sich in:

**John, Kristina (2008):** Rekonstruktion von abgeleiteten Variablen im Mikrozensus 2005 im Vergleich zu den Mikrozensen bis 2004. GESIS Methodenbericht 2008/7.

**John, Kristina (2007):** Rekonstruktion von Bandsatzerweiterungen mittels zeilen- bzw. satzübergreifender Operation in STATA im Mikrozensus. Gesis-Methodenberichte 3/2007.

**John, Kristina (2006):** Rekonstruktion von Bandsatzerweiterungen mittels zeilen- bzw. satzübergreifender Operationen in STATA.

**Lengerer, Andrea; Boehle, Mara (2006):** Rekonstruktion von Bandsatzerweiterungen zu Haushalt, Familie und Lebensformen im Mikrozensus. ZUMA-Methodenbericht 2006/05.