Auswertungsbeispiele: Alter des jüngsten Kindes in der Lebensgemeinschaft

Oft benötigt man Angaben zu einer Einheit von Personen, etwa zum Haushalt, zur Familie oder zur Lebensgemeinschaft. Ein Beispiel ist das Alter des jüngsten Kindes in der Lebensgemeinschaft.

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.
Das Alter des jüngsten Kindes in der Lebensgemeinschaft kann folgendermaßen ermittelt und als Variable allen Personen der jeweiligen Lebensgemeinschaft zugespielt werden (hier am Beispiel des Mikrozensus 2000):

Statistik-Programm-Routinen

SPSS	Stata

SELECT IF EF513=1.
IF EF514=3 alt_k=EF30.
SORT CASES BY EF3 EF4 EF643.
AGGREGATE OUTFILE=* MODE=ADDVARIABLES
  / PRESORTED
  / BREAK EF3 EF4 EF643
  / alt_jk=MIN(alt_k).
EXECUTE.
	/* Für die Zuspielung von Individualdaten zu Aggregatdaten gibt es in STATA mehrere Möglichkeiten.*/
	
	Option 1: egen
	generate kind_alt=EF30 if EF513==1 & EF514==3
	egen alt_jk= min(kind_alt*(EF513==1)), by(EF3 EF4 EF643)
	
	Option 2: merge
	sort EF3 EF4 EF643
	save Laufwerk:\<Padangabe>\filename1.dta, replace
	
	sort EF3 EF4 EF643
	  generate alt_jk_merge1 = EF30 if EF513==1 & EF514==3
	  collapse(min) alt_jk_merge1, by(EF3 EF4 EF643)
	  sort EF3 EF4 EF643
	  save Laufwerk:\<Padangabe>\filename2.dta, replace
	use Laufwerk:\<Padangabe>\filename1.dta, clear
	merge EF3 EF4 EF643 using ///
	Laufwerk:\<Padangabe>\filename2.dta, uniqusing

Erläuterungen

SPSS	Stata
Zunächst erfolgt eine Einschränkung der interessierenden Population auf die Bevölkerung am Wohnsitz der Lebensgemeinschaft (EF513=1). Eine Hilfsvariable alt_k wird gebildet, die für ledige Kinder in der Lebensgemeinschaft (EF514=3) das Alter (EF30) enthält. Dann wird diese Hilfsvariable alt_k auf Ebene der Lebensgemeinschaft (EF3, EF4, EF643) aggregiert und ihr Minimum wird als neue Variable (alt_jk) allen Personen der jeweiligen Lebensgemeinschaft zugewiesen. Gibt es kein Kind in der Lebensgemeinschaft, hat die Variable alt_jk einen ungültigen Wert.	**Option 1: egen** Zunächst wird eine Hilfsvariable kind_alt (Alter des Kindes) gebildet, die für ledige Kinder (EF514==3) in der Lebensgemeinschaft (EF513==1) das Alter (EF30) enthält. Mit *egen* wird die Variable alt_jk (Alter des jüngsten Kindes) gebildet. Die Funktion min(exp) bewirkt, dass das Minimum des Alters des Kindes (kind_alt) für die Bevölkerung am Wohnsitz der Lebensgemeinschaft (EF513==1) gebildet wird. Die Sortierung des Datensatzes nach den zur Abgrenzung von Lebensgemeinschaften notwendigen Variablen (EF3, EF4, EF643) erfolgt in den Optionen nach dem Komma. Gibt es kein Kind in der Lebensgemeinschaft, weist die Variable alt_jk einen ungültigen Wert auf. **Option 2: merge** Zunächst wird der Datensatz nach den zur Abgrenzung von Lebensgemeinschaften notwendigen Variablen (EF3, EF4, EF643) sortiert und daraufhin gespeichert. Nach erneuter Sortierung wird die Variable alt_jk_merge1

	(Alter des jüngsten Kindes in der Lebensgemeinschaft) generiert. Über diese Variable erfolgt die Abgrenzung auf Individualebene, d.h. auf die Bevölkerung am Wohnsitz der Lebensgemeinschaft (EF513==1) und auf ledige Kinder in der Lebensgemeinschaft (EF514==3). 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:

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

Stata
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.