

Satzübergreifendes Arbeiten und der Umgang mit Jahresüberhängen

Oft benötigt man Angaben zu einer Einheit von Personen, etwa zum Haushalt, zur Familie oder zur Lebensgemeinschaft. Ein Beispiel ist die Zahl der Personen im Haushalt. In Form sogenannter Bandsatzerweiterungen sind solche Variablen häufig bereits in den Daten des Mikrozensus enthalten. Sie werden vom Statistischen Bundesamt generiert. Im Folgenden wird nun gezeigt, was seit der Erhebung 2006 beachtet werden muss, wenn solche Variablen selbst gebildet werden. Das Beispiel geht dabei auf die Besonderheiten von Jahresüberhängen ein, die sich seit 2006 in den Daten befinden. Jahresüberhänge sind Haushalte bzw. Personen, die im Erhebungszeitraum i. d. R. aufgrund schwerer Erreichbarkeit nicht befragt werden konnten, sondern erst im Folgejahr.

Beim satzübergreifenden Arbeiten sollten die Jahresüberhänge immer berücksichtigt werden, da im SUF im Fall von Jahresüberhängen die Haushaltsnummer im Auswahlbezirk (EF4) mit der Haushaltsnummer für den jeweils aktuellen Jahresbestand identisch. Bei eigenen Abgrenzungen auf der Ebene von Haushalten, Lebensformen oder Familien muss deshalb für eine eindeutige Identifikation der Einheiten die folgende Sortierfolge verwendet werden: Auswahlbezirk (EF3), Berichtsquartal (EF12), Jahresüberhang (EF5b) und Haushaltsnummer (EF4) sowie nach Bedarf Nummer der Lebensform im Haushalt (EF27) oder Nummer der Familie im Haushalt (EF25). Anderenfalls kann es zu Doppelungen kommen, wie das folgende Beispiel zeigen wird.

Im Folgenden wird anhand des Beispiels „EF663 Personen im Haushalt: Anzahl (Haushaltstypisierung)“ mit dem Mikrozensus 2008 gezeigt, wie sich die Nichtberücksichtigung von Jahresüberhängen auswirkt.

Hier wird die Variable EF663 Personen im Haushalt: Anzahl (Haushaltstypisierung) beispielhaft ohne Berücksichtigung von Jahresüberhängen generiert:

```
egen v663_1 = total(ef31==1), by(ef1 ef3 ef4)
```

Mit egen wird die Variable v663_1 gebildet, welche die bereits vorhandene Bandsatzerweiterung ef663 replizieren soll. Die Funktion total bewirkt, dass die Summe der Bevölkerung in Privathaushalten (EF31==1) gebildet wird. Die Sortierung des Datensatzes erfolgt hier nur nach e1, ef3 und ef4 in den Optionen nach dem Komma. Die in diesem Jahr zur Abgrenzung von Haushalten eigentlich benötigten Variable EF5b wird also hier nicht berücksichtigt.

```
recode v663_1 (9/90=9)
```

Zur Angleichung an die bereits vorhandene Bandsatzerweiterung EF633 werden Haushalte mit über 9 Personen in einer Kategorie zusammengefasst.

Die gleiche Prozedur wird nun wiederholt, dieses Mal werden jedoch die Jahresüberhänge wie empfohlen berücksichtigt.

```
egen v663_2 = total(ef31==1), by(ef1 ef3 ef4 ef5b ef12)  
recode v663_2 (9/90=9)
```

Nun werden beide Variablen innerhalb eines Beispielhaushalts verglichen um zu verstehen, welche Auswirkung eine Nichtberücksichtigung von Jahresüberhängen haben kann:

```
list ef1 ef3 ef4 ef5a ef5b ef12 ef31 ef661 ef662 ef663 v663_1 v663_2 if ef1==1 & ef3==34 & ef4==1, nol noobs  
sepby(ef5b ef12)
```

EF1	EF3	EF4	EF5a	EF5b	EF12	EF31	EF661	EF662	EF663	EF663_1	EF663_2
1	34	1	1	0	4	1	1	2	2	4	2
1	34	1	2	0	4	1	9	7	2	4	2
1	34	1	1	1	1	1	1	2	2	4	2
1	34	1	2	1	1	1	9	7	2	4	2

Im Fall von EF663_2 wurden beim Erstellen der Variable die Jahresüberhänge berücksichtigt, daher werden wie bei der vom Statistischen Bundesamt erstellten Variable EF663 zwei Personen im Haushalt gezählt. Wenn wie im Fall von EF663_1 EF5b nicht berücksichtigt wird, werden zwei Bezugspersonen im Haushalt (EF661) und zwei Haupteinkommensbezieher im Haushalt (EF662) gezählt, und insgesamt vier Personen im Haushalt (EF662_1). Der Haushalt wurde zweimal befragt (Jahresbestand und Jahresüberhang) und wird ohne Berücksichtigung von EF5b doppelt gezählt, da die Haushaltsnummer identisch ist. Daher muss beim Erstellen von sogenannten Bandsatzerweiterungen die Information zum Jahresüberhang (EF5b) ab dem Mikrozensus 2006 immer berücksichtigt werden.