

# Musterprogramm STATA.do

\*\*\*\*\*

\*\*\* Aufzeichnung in Protokoll starten.

capture log close

log using "Pfad wird im Forschungsdatenzentrum ergänzt\log\_<Dateiname des Programmcodes>.log", replace

\*\*\*\*\*

\*\*\*\*\*

\*\*\* Titel des Projekts: <Frauen und Arbeit in der  
Bundesrepublik>

\*\*\* Datengrundlage: <Mikrozensus 1985>

\*\*\*

\*\*\* Dateiname des Programmcodes: Musterprogramm.do

\*\*\* erstellt: <Datum>

\*\*\* von: <Name>

\*\*\* E-Mail: <E-Mail-Adresse>

\*\*\* Tel.: <Telefonnummer>

\*\*\*

\*\*\* Dateiname des Output-Files: <Musterprogramm.txt>

\*\*\*

\*\*\*

\*\*\* Grundriss des Programms: <Programm zur Untersuchung von  
mehrfachen Personensätzen

\*\*\* (Check der Datenbasis); Zusammenhang

\*\*\* von Alter und Familienstand>

\*\*\*

\*\*\* Verwendete Variablen:

\*\*\* Originalvariablen: <ef1: Land der Bundesrepublik

ef2: Regierungsbezirk

ef3: Auswahlbezirks-Nummer

ef4: Lfd. Nr. d. Haushalts

ef5: Lfd. Nr. d. Person im Haushalt

ef6: Lfd. Nr. d. Familie im Haushalt

ef23: Alter

ef26: Bevölkerungsgruppe

ef27: Bevölkerung in Privathaushalten

ef28: Bevölkerung am Familienwohnsitz

ef38: Familienstand>

\*\*\*

\*\*\* Neu angelegte Variablen: <persnr: Personennummer

famnr: Familiennummer

hhnr: Haushaltsnummer

nrdiff: Differenz Haushaltsnummer - Familiennummer

piddiff: Test einmalig Personennummer>

\*\*\*

```

*** Gewichtungsvariablen: <gew1: Gewichtungsfaktor aus ef253>
***
***
*****
*****

**** Version festlegen

version <8.2>

**** Bildschirm Ausgabe steuern

set more off

**** Arbeitsverzeichnis festlegen

cd '<Pfad>'

**** Datensatz einlesen

use mz_1985.dta, clear

**** Auswahl der Bevölkerung in Privathaushalten (ef27 eq 1) am Ort der Hauptwohnung (ef26 lt 3)
**** und am Familienwohnsitz ef28=1

keep if ef27 == 1 & ef26 < 3 & ef28 == 1

**** Generierung Haushalts- und Familiennummer aus ef1-ef6
**** pers ef1-ef5
**** hh ef1-ef4
**** fam ef1-ef4 + ef6

sort ef1 ef2 ef3 ef4 ef5
generate persnr=(ef1*10000000000)+(ef2*1000000000)+(ef3*100000)+(ef4*1000)+(ef5*10)

sort ef1 ef2 ef3 ef4
generate hhnr=(ef1*10000000000)+(ef2*1000000000)+(ef3*100000)+(ef4*1000)

sort ef1 ef2 ef3 ef4 ef6
generate famnr=(ef1*10000000000)+(ef2*1000000000)+(ef3*100000)+(ef4*1000)+(ef6*1)

**** Beschreibung der Variablen persnr hhnr famnr

```

# Musterprogramm STATA.do

```
lab var hhnr "Haushaltsnummer"  
lab var famnr "Familienummer"  
lab var persnr "Personennummer"
```

```
**** Multiplikation des Gewichtungsfaktors mit 0,1  
generate gew=(ef253 * 0.1)
```

```
**** vergl. 'Haushaltsnummer' (hhnr) - 'Familienummer' (famnr) - 'Personennummer' (personennr)
```

```
generate nrdiff=0
```

```
**** Wenn Familiennummer (famnr) ungleich der Personennummer (personennr) dann Differenz Haushaltsnummer -  
**** Familiennummer (nrdiff) gleich 1
```

```
replace nrdiff=1 if famnr ~= hhnr
```

```
lab var nrdiff "Differenz Haushaltsnummer - Familiennummer"
```

```
lab def nrddiff 1 "Haushaltsnummer und Familiennummer unterschiedlich" 0 "Haushaltsnummer und Familiennummer gleich"  
lab val nrdiff nrddiff
```

```
****Output Nr. 1: Haushaltsnummer ungleich Familiennummer
```

```
tabulate nrdiff
```

```
**** test einmalig personennr  
sort persnr
```

```
**** Personennummer (persnr) gleich der des Vorgängers dann Personennummertest (piddiff) gleich 1
```

```
generate piddiff=0  
replace piddiff=1 if persnr == persnr[_n-1]
```

```
lab var piddiff "Test einmalig Personennummer"
```

```
lab def piddiff 1 "Personennummer mehrfach" 0 "Personennummer einfach"  
lab val piddiff piddiff
```

```
****Output Nr. 2: Mehrfache Personennummern  
tabulate piddiff
```

# Musterprogramm STATA.do

```
**** Erstellen einer benutzerdefinierten Tabelle (mit Anzahl)
**** Mittelwert und gültige Häufigkeit (N) des Alters (ef23) je Ausprägung Familienstand (ef38)
**** Output Nr. 3: Kreuztabelle Alter und Familienstand

tabstat ef23, statistics(mean count) by (ef38)

**** Speichern der Arbeitsdatei
save MZ85arb1.dta

log close

exit
```