

R-, SAS-, SPSS- und Stata Beispielprogramme zu
Schimpl-Neimanns, Bernhard (2009):
Schätzung des Stichprobenfehlers im Mikrozensus mit Stata – Eine
Einführung mit Beispielen zum Campus File Mikrozensus 2002.
GESIS-Methodenbericht 2009/02.
Mannheim: GESIS.

URL: http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_methodenberichte/2009/gesis_mb_09_02.pdf

Der Bericht zeigt, wie mit Stata (Version 10) Gesamt-, Verhältnis- und Mittelwerte bei freier Hochrechnung (Designgewichtung) sowie gebundener Hochrechnung (Poststratifikation, Redressment) mit dem Campus File Mikrozensus 2002 geschätzt werden können. Daneben wird beschrieben, wie bei statistischen Modellen vorgegangen werden kann, um evtl. durch das Stichprobendesign bedingte Modellverletzungen zu beheben.

Ergänzend zu Stata werden im WWW unter der Adresse <http://www.gesis.org/dienstleistungen/tools-standards/mikrodaten-tools/sp-fehler/> auch Beispielprogramme für die Statistikpakete R, SAS und SPSS angeboten (Programmstatus: April 2009). Diese Programme unterscheiden sich von den Stata-Programmen vorwiegend aufgrund des Leistungsumfangs; z. B. bieten SAS und SPSS keine Optionen zur Umsetzung der gebundenen Hochrechnung an. Für typische Schätzungen finden sich jeweils am Ende der Stata-Programme (Do-Files; nach „exit“) Hinweise auf Umsetzungen mit dem Scientific Use File Mikrozensus 2002 (siehe Beispiele 1, 3, 6 und 8). Vor Verwendung der Programme müssen die Verzeichnisse und ggf. die Dateinamen geändert werden.

In den hier verwendeten Daten des Campus Files Mikrozensus 2002 sind keine fehlenden Werte mehr enthalten und weitere Modifikationen in enger Anlehnung an das Setup für das Scientific Use File Mikrozensus 2002 vorgenommen worden (siehe www.gesis.org/dienstleistungen/daten/amtliche-mikrodaten/mikrozensus/grundfile/mz2002/stata-einleseroutine/). Insofern unterscheiden sie sich vom Campus File, wie es von den Forschungsdatenzentren bereitgestellt wird (siehe www.forschungsdatenzentrum.de/bestand/mikrozensus/cf/2002/-index.asp).

Inhalt der Datei Stata_Beispiele.zip

readme.pdf	Diese Datei
Beisp_01.do	Designbasierte Schätzung des Gesamtwertes Zahl der Erwerbslosen
Beisp_02.do	Designbasierte Schätzung der Zahl der Erwerbstätigen (10 PSUs)
Beisp_03.do	Designeffekte im Campus File
Beisp_04.do	Gruppenvergleiche
Beisp_05.do	Gesamtwerte mit gebundener Hochrechnung (Poststratifikation)
Beisp_06.do	Gesamtwerte mit gebundener Hochrechnung (Regressionsschätzung)
Beisp_07.do	Erwerbslosenquote mit Designgewichtung
Beisp_08.do	Erwerbslosenquote mit gebundener Hochrechnung (Poststratifikation)
Beisp_09.do	Durchschnittliche Arbeitsstunden (Designgewichtung)
Beisp_10.do	Durchschnittliche Arbeitsstunden mit gebundener Hochrechnung (Poststratifikation)
Beisp_11.do	Regression des Nettoeinkommens auf Bildung, Geschlecht und Staatsangehörigkeit
mz02cf_Beisp.dta	Variablenauswahl Campus File Mikrozensus 2002 (s. o.)

Inhalt (ausführliche Programmbeschr. s. o.)	R_Beispiele.zip	SAS_*.zip	SPSS_*.zip
Diese Datei	readme.pdf	readme.pdf	readme.pdf
Gesamtwert Erwerbslose: Designgewichtung	Beisp_01.R	*.sas	*.sps
Gruppenvergleiche	Beisp_04.R	*.sas	*.sps
Gesamtwert: Poststratifikation	Beisp_05.R	*.sas	*.sps
Gesamtwert: Regressionsschätzung	Beisp_06.R	*.sas	*.sps
Erwerbslosenquote: Designgewichtung	Beisp_07.R	*.sas	*.sps
Erwerbslosenquote: Poststratifikation	Beisp_08.R	*.sas	*.sps
Durchschn. Arbeitsstd.: Designgewichtung	Beisp_09.R	*.sas	*.sps
Durchschn. Arbeitsstd.: Poststratifikation	Beisp_10.R	*.sas	*.sps
Regression Nettoeinkommens	Beisp_11.R	*.sas	*.sps
Systemfile: Campus File Mikrozensus 2002 (Variablenauswahl; s. o.)	mz02cf_Beisp.dta	*.sas7bdat formats.sas7bcat	*.sav