

gesis

Leibniz-Institut
für Sozialwissenschaften

TECHNICAL *Reports*

2013|10

Mikrozensus Scientific Use File 2010

Dokumentation und Datenaufbereitung

Andreas Herwig, Bernhard Schimpl-Neimanns

GESIS-Technical Reports 2013|10

Mikrozensus Scientific Use File 2010

Dokumentation und Datenaufbereitung

Andreas Herwig, Bernhard Schimpl-Neimanns

GESIS-Technical Reports

GESIS – Leibniz-Institut für Sozialwissenschaften
Postfach 12 21 55
68072 Mannheim
Telefon: (0621) 1246 - 263
Telefax: (0621) 1246 - 100
E-Mail: bernhard.schimpl-neimanns@gesis.org

ISSN: 1868-9043 (Print)
ISSN: 1868-9051 (Online)
Herausgeber,
Druck und Vertrieb: GESIS – Leibniz-Institut für Sozialwissenschaften
Unter Sachsenhausen 6-8, 50667 Köln

Zusammenfassung

Die Aufbereitung und Dokumentation des faktisch anonymisierten Einzelmaterials des Mikrozensus als Scientific Use File erfolgt durch das Statistische Bundesamt, Gruppe F2 Bevölkerung, Mikrozensus, Wohnen und Migration, in Kooperation mit dem German Microdata Lab der GESIS. In dem vorliegenden Bericht werden das Erhebungsprogramm und der Stichprobenplan des Mikrozensus im Allgemeinen sowie die Besonderheiten des Mikrozensus 2010 skizziert. Diese Daten umfassen das Zusatzprogramm zur Wohnsituation und das Ad-hoc-Modul zum Thema Vereinbarkeit von Familie und Beruf der Arbeitskräfteerhebung der EU. Außerdem werden die Datenaufbereitung und Substichprobenziehung sowie die Qualitätsprüfung des Scientific Use Files beschrieben.

Schlüsselwörter: Mikrozensus, Scientific Use File

JEL-Klassifikation: J00

German Microcensus Scientific Use File 2010: Documentation and Data Preparation

Abstract

The processing of the factually anonymized microdata of the German Microcensus into a scientific use file and its documentation is carried out by the German Federal Statistical Office, Group F 2 Microcensus, in cooperation with the German Microdata Lab of GESIS. This technical report gives a general outline of the Microcensus' survey program and sampling plan. Furthermore, the unique features of the Microcensus 2010 will be discussed. The survey 2010 includes the supplementary program on housing situations, and the ad hoc module reconciliation between work and family life made available by the EU Labour Force Survey. In addition, a description of the processing of data, sub-sampling and the quality testing of the scientific use file is provided.

Keywords: German Microcensus, Scientific Use File

JEL classification: J00

1 Einleitung

Seit dem Bundesstatistikgesetz von 1987 (§ 16, Abs. 6, BStatG) können anonymisierte Individualdaten der amtlichen Statistik durch die statistischen Ämter an unabhängige inländische Forschungseinrichtungen weitergegeben werden. Das Scientific Use File (SUF) ist eine faktisch anonymisierte 70 %-Substichprobe der Haushalte des Mikrozensus, wobei alle Personen eines ausgewählten Haushalts enthalten sind. Statt einer Haushaltssubstichprobe wird in den Erhebungsjahren, in denen das Zusatzprogramm zur Wohnsituation (z. B. 2002, 2006 und 2010) durchgeführt wird, eine Wohnungssubstichprobe gezogen. Das SUF ist eine systematische Zufallsauswahl aus den Originaldaten des Statistischen Bundesamtes. Das Ziehungsverfahren orientiert sich eng am Stichprobendesign des Mikrozensus. Das Konzept der faktischen Anonymität bezieht sich auf das sogenannte Grundfile, das neben der Ziehung einer Substichprobe die eingeschränkte Weitergabe von Regionalinformationen vorsieht. Im SUF werden nur die Länder und vergrößerte Angaben zur Gemeindegrößenklasse ausgewiesen. Des Weiteren werden Merkmalsausprägungen vergrößert, weil Staatsangehörigkeiten oder Gruppen von Nationalitäten in der Grundgesamtheit mindestens 50.000 Personen umfassen müssen. Für andere Merkmale liegt dieses Anonymisierungskriterium bei 5.000 Personen.

Die Aufbereitung und Dokumentation des SUF erfolgt durch das Statistische Bundesamt, Gruppe F2 Bevölkerung, Mikrozensus, Wohnen und Migration, in Kooperation mit dem German Microdata Lab (GML) der GESIS. Durch die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder werden Rohdaten im ASCII-Format sowie Daten im dokumentierten Format (Systemfiles) der am häufigsten verwendeten Statistikprogramme SPSS, SAS und STATA bereitgestellt.

Um die Datendokumentation und Auswertungsmöglichkeiten für die Nutzer des SUF weiter zu verbessern, sind ab der Erhebung 2007 zwischen der Mikrozensusgruppe des Statistischen Bundesamtes und dem GML die Schritte zur Datenaufbereitung und der -dokumentation neu gestaltet worden. Ein Ergebnis ist die Bereitstellung von Systemfiles. Außerdem wird das bisherige Schlüsselverzeichnis durch ein Datenhandbuch mit umfassenden Informationen zu den Variablen ersetzt. Diese und weitere Informationen stehen auf den Web-Seiten von GESIS zur Verfügung.¹

In diesem Bericht werden nach einem Überblick zum Mikrozensus und zum SUF 2010 zunächst die Datenaufbereitung und die Schritte zur Prüfung der Daten beschrieben bevor die wichtigsten Veränderungen im Vergleich zum SUF 2009 benannt und die Verwendung der amtlichen Bevölkerungskonzepte sowie die Hochrechnung beispielhaft erläutert werden.

¹ Weitere Informationen sind auf den Web-Seiten des Statistischen Bundesamtes (2013a; [Statistik von A bis Z](#)) und der Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder (FDZ 2013; siehe [Datenangebot, Mikrozensus](#)) abrufbar.

2 Kurzbeschreibung des Mikrozensus

Der Mikrozensus ist die amtliche Repräsentativitätsstatistik über die Bevölkerung und den Arbeitsmarkt in Deutschland mit einem Auswahlsatz von einem Prozent der Haushalte und der Bevölkerung. Er wird seit 1957 in Westdeutschland und seit 1991 auch in den neuen Bundesländern durchgeführt. Der Mikrozensus dient der Bereitstellung statistischer Informationen über die wirtschaftliche und soziale Lage der Bevölkerung sowie über die Erwerbstätigkeit, den Arbeitsmarkt und die Ausbildung. Er schreibt die Ergebnisse der Volkszählung fort. Darüber hinaus dient er der Rationalisierung anderer amtlicher Statistiken. Die Arbeitskräfteerhebung (AKE) der Europäischen Union (engl.: Labour Force Survey (LFS)), einschließlich der in diesem Rahmen durchgeführten Ad-hoc-Module, ist in den Mikrozensus integriert. Im Folgenden werden nur die Haupteigenschaften skizziert (für Details siehe den Qualitätsbericht zum Mikrozensus 2010 des Statistischen Bundesamtes (2011a), der auch in den Fachserien zum Mikrozensus enthalten ist.

Grundgesamtheit Zur Grundgesamtheit zählt die gesamte wohnberechtigte Bevölkerung in Privathaushalten und Gemeinschaftsunterkünften am Haupt- und Nebenwohnsitz. Nicht zur Erhebungsgesamtheit gehören Angehörige ausländischer Streitkräfte sowie ausländischer diplomatischer Vertretungen mit ihren Familienangehörigen. Personen ohne Wohnung (Obdachlose) haben im Mikrozensus aufgrund der Anlage als Flächenstichprobe keine Erfassungschance.

Berichtszeitraum Seit 2005 wird das Befragungsvolumen gleichmäßig über alle Kalenderwochen des Jahres verteilt (kontinuierliche Erhebung). Der Berichtszeitraum ist die letzte Kalenderwoche vor der Befragung (gleitende Berichtswoche).

Fragenprogramm Das Fragenprogramm des Mikrozensus besteht aus einem gleichbleibenden Grundprogramm, das sozioökonomische Grundinformationen liefert und die laufende Beobachtung des Arbeitsmarktes ermöglicht.

Neben den jährlich zu erhebenden Merkmalen des Grundprogramms werden in vierjährigen Zusatzprogrammen weitere Fragen gestellt. Das Zusatzprogramm 2010 umfasst Angaben zur Wohnsituation. Darüber hinaus wurde im Rahmen der EU-Arbeitskräfteerhebung (AKE) 2010 das Ad-hoc-Modul zum Thema „Vereinbarkeit von Familie und Beruf“ durchgeführt. Im Labour Force Survey werden seit 2005 einige Fragen zur Erwerbstätigkeit als sogenannte Strukturvariablen nur als Jahresdurchschnittswerte erhoben. Der Auswahlsatz für diese AKE-Strukturvariablen und das Ad-hoc-Modul beträgt 0,1 %.

Stichprobenplan Der Mikrozensus ist eine einstufig geschichtete Flächen- oder Klumpenstichprobe, das heißt, aus dem Bundesgebiet werden Auswahlbezirke ausgewählt, in denen alle Haushalte und Personen befragt werden. Sowohl die Auswahlbezirke als Primäreinheiten als auch die Sekundäreinheiten (Haushalte und Personen) haben die gleiche Auswahlwahrscheinlichkeit von einem Prozent. Die Auswahlbezirke der Grundausswahl wurden aus Ergebnissen der Volkszählung 1987 gebildet; für die neuen Bundesländer wurde auf der Basis des Bevölkerungsregisters Statistik eine vergleichbare Auswahlgrundlage erstellt.

Die Stichprobe der Grundausswahl wird durch Berücksichtigung der Neubautätigkeit über die Meldungen zur Bautätigkeitsstatistik aktualisiert (Neubauauswahl). Jährlich werden ein Viertel aller in der Stichprobe enthaltenen Auswahlbezirke ausgetauscht.

Die aus einem Auswahlbezirk wegziehenden Haushalte und Personen werden nicht weiter befragt, sondern durch zuziehende Haushalte und Personen ersetzt. Der Mikrozensus ist somit eine Wiederholungsbefragung mit teilweiser Überlappung der Erhebungseinheiten (partielle Rotation).

Erhebungsverfahren Die Interviews werden i. d. R. als mündliche Befragung mit Hilfe eines Laptops (CAPI) durchgeführt. Die Haushalte können auch selbst einen schriftlichen Fragebogen ausfüllen. Zulässig sind Proxy-Interviews, d. h., ein erwachsenes Haushaltsmitglied darf stellvertretend für andere Haushaltsmitglieder antworten (Fremdauskünfte für Personen ab 15 Jahren: ca. 24 %). Für die ausgewählten Haushalte und die darin lebenden Personen besteht gemäß der gesetzlichen Regelungen Auskunftspflicht. Einige Fragen sind freiwillig zu beantworten.

Insbesondere wegen verzögerter Beantwortung der Haushalte liegen die Berichtswochen nicht immer im Monat oder Quartal des Auswahlplanes, so dass es mitunter zu Verschiebungen von Haushalten in die Stichprobe des Folgejahres, sogenannten Jahresüberhängen, kommt (außer im Jahreswechsel 2006/2007).

Hochrechnung Bei der Hochrechnung wird ein zweistufiges Verfahren verwendet. Im ersten Schritt der Ausfallkorrektur wird von der Nettostichprobe der erfolgreich befragten Haushalte auf die Bruttostichprobe aller zu befragenden Haushalte hochgerechnet. Der Unit-Nonresponse beträgt im Mikrozensus 2010 2,5 Prozent der Haushalte (ohne Jahresüberhänge; s. u.). Im zweiten Schritt wird diese Bruttostichprobe an Eckzahlen aus der laufenden Bevölkerungsfortschreibung angepasst. Ab 2005 werden dabei auch Ergebnisse des Ausländerzentralregisters verwendet. Die Anpassung erfolgt quartalsweise auf unterschiedlichen regionalen Ebenen. Mit Ausnahme der AKE-Strukturvariablen, des Ad-hoc-Moduls des EU Labour Force Surveys und den Variablen zur Wohnsituation liegen für alle Merkmalsbereiche gleiche Hochrechnungsfaktoren vor.

3 Das Scientific Use File Mikrozensus 2010

Wie bereits einleitend erwähnt, ist das SUF eine faktisch anonymisierte 70 %-Substichprobe; i. d. R. eine Haushaltssubstichprobe. Wegen des Zusatzprogramms zur Wohnsituation wurde 2010 eine 70 %-Substichprobe der Wohnungen gezogen. Das Ziehungsverfahren orientiert sich eng am Stichprobendesign des Mikrozensus.

Die Substichprobe ist eine systematische Zufallsauswahl aus den Originaldaten. Die Daten werden dazu nach Berichtsquartal, Bundesland, Unterstichprobenkennung, Regierungsbezirk, Gemeindegrößenklasse, Anzahl der Personen in der Wohnung, Gebäudegrößenklasse, Auswahlbezirksnummer und Wohnungsnummer sortiert. Das Merkmal Gebäudegrößenklasse wurde im SUF 2006 erstmals verwendet. Die Unterstichprobenkennung wird ab 2008 eingesetzt, um die Ergebnisqualität des Ad-hoc-Moduls und der AKE-Stukturvariablen zu verbessern. Durch diese Anordnung ist gewährleistet, dass die Substichprobe hinsichtlich dieser Merkmale nur geringe zufallsbedingte Abweichungen zu den Originaldaten aufweist.

Bei Gemeinschaftsunterkünften erhält jede Person eine eigene fortlaufende „Wohnungs“-Nummer. Zur Stichprobenziehung dient die letzte (Einer-) Stelle der Wohnungsnummer. Alle Wohnungen, deren Einerstelle nicht mit drei zufällig gezogenen einstelligen Zahlen übereinstimmen, gelangen nach diesem sogenannten Schlussziffernverfahren in die Stichprobe (70 % der Wohnungen). Nicht enthalten sind sogenannte Nullbezirke, das sind Auswahlbezirke ohne Befragte (z. B. leer stehende Häuser). Diese werden vor der Substichprobenziehung gelöscht.

Weiterhin unterscheidet sich das SUF von den Originaldaten des Mikrozensus dadurch, dass bestimmte Variablen im SUF, bedingt durch die Anonymisierung, in klassierter und vergrößerter Form verfügbar sind.² So sind im SUF die Regionalangaben Bundesland und eine vergrößerte Gemeindegrößenklasse, die nicht für die kleinen Länder Bremen und Saarland ausgewiesen ist, enthalten. Mithilfe eines speziellen Codes der Variablen Gemeindegrößenklasse (EF563 = 9) kann zwischen West- und Ost-Berlin unterschieden werden. Auch bei den weiteren Variablen des SUF werden je nach Bedarf Merkmalsvergrößerungen vorgenommen, sodass jede Ausprägung in den univariaten Verteilungen mindestens 5.000 Personen in der Grundgesamtheit aufweist. Die Ausprägungen der Variable Nationalität werden so aggregiert, dass jede Nationalitätengruppe in der Grundgesamtheit wenigstens 50.000 Einwohner umfasst. Im Fall von Vergrößerungen wird im SUF i. d. R. die am stärksten besetzte Kategorie ausgewiesen.

Für Schätzungen des Stichprobenfehlers enthält das SUF Informationen zum Stichprobendesign (Schimpl-Neimanns 2011). Mit den Variablen Bundesland (EF1) und Gebäudegrößenklasse (EF570) sowie den Angaben zum Berichtsquartal (EF12) stehen Schichtungsmerkmale zur Verfügung. Die Information zur Klumpung liegt in Form der Auswahlbezirksnummer (EF3) vor. Des Weiteren können die Hochrechnungsfaktoren für Quartals- und Jahresdurchschnittsergebnisse (EF951 und EF952) sowie für die Merkmale der 0,1 %-Substichprobe mit den Strukturvariablen des EU Labour Force Survey (EF955) und des Ad-hoc-Moduls (EF956) genutzt werden. Für die Auswertung der Angaben zur Wohnsituation

² Zusätzlich zu den im SUF enthaltenen Variablen können im Rahmen eines Aufenthalts als Gastwissenschaftler in den Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder weitere Variablen genutzt werden. Die Auflistung dieser Variablen findet sich am Ende der thematischen Übersicht (Statistisches Bundesamt / GESIS 2013b).

liegt ein spezieller Hochrechnungsfaktor (EF960) vor, mit dem die Ergebnisse des Mikrozensus an ausgewählte Eckzahlen der laufenden Wohnungsfortschreibung angepasst werden. Außerdem werden damit Haushalte kompensiert, die 2010 nicht befragt werden konnten (für Details siehe Statistisches Bundesamt 2012a: 4-5). Diese Hochrechnungsfaktoren sind bereits mit dem Kehrwert der Auswahlwahrscheinlichkeit der Substichprobe (100/70) multipliziert und rechnen auf 1.000 der Population hoch.

3.1 Die wichtigsten Veränderungen gegenüber 2009

Um Nutzern, die bereits mit dem SUF früherer Erhebungszeitpunkte gearbeitet haben, die Auswertung zu erleichtern, werden im Folgenden die wichtigsten Veränderungen gegenüber dem MZ 2009 sowie zum Zusatzprogramm im MZ 2006 skizziert. Sofern nicht gesondert auf Änderungen hingewiesen wird, kann von einer weitgehenden Vergleichbarkeit der Daten ausgegangen werden. Ausführliche Informationen finden sich in den „Wichtige[n] Informationen zur Nutzung des Mikrozensus Scientific Use Files 2010“ (Statistisches Bundesamt / GESIS 2013c) sowie im [Datenhandbuch](#) (Statistisches Bundesamt / GESIS 2013a), das in der Variablenbeschreibung ebenfalls erläuternde Kommentare mit Hinweisen zur zeitlichen Vergleichbarkeit sowie den Fragetext des Selbstausfüllerbogens bzw. des CAPI-Interviews enthält. Zur Vergleichbarkeit mit weiteren Erhebungen des Mikrozensus ab 1973 enthält die [Variablen-Zeitpunkte-Matrix](#) in MISSY zusätzliche Hinweise.

- Im Rahmen der in 4-jährigen Abständen wiederholten Zusatzprogramme sind im Mikrozensus 2010 folgende Variablen zur Wohnsituation neu:

EF489-EF504 *Wohnsituation des Haushalts, Teil 1*

EF588-EF643 *Wohnungstypisierungen*

EF1000u1-EF1011 *Wohnsituation des Haushalts, Teil 2*

Zu beachten ist, dass 2010 im Vergleich zu 2006 teilweise neue Fragen gestellt und neue Antwortkategorien eingesetzt wurden. Des Weiteren enthält 2010 die Variable EF502 die Grundmiete, 2006 dagegen die Bruttokaltmiete (siehe für detaillierte Hinweise Statistisches Bundesamt / GESIS 2013c).

Für Hochrechnungen zum Bestand und zur Struktur der Wohnungen soll der Hochrechnungsfaktor EF960 verwendet werden. Für Hochrechnungen zur Wohnsituation der Haushalte sind die Standardhochrechnungsfaktoren (EF951, EF952) einsetzbar.

Aufgrund eines in den Originaldaten noch vorhandenen Fehlers (siehe Statistisches Bundesamt, 2012a) sind im momentan verfügbaren SUF eigenständige satzübergreifende Arbeiten auf Wohnungsebene nicht möglich und Typisierungen zur Wohnsituation (z. B. EF637 „Personen in der Wohnung: Anzahl“) nicht replizierbar.

- Im Mikrozensus 2010 wurde die Erfassung der Staaten und Staatsangehörigkeiten hinsichtlich Codierung und Differenziertheit geändert. Die Staaten werden nun bei der Befragung mit einem Buchstabencode erhoben und anschließend numerisch codiert. Im Unterschied zur schriftlichen Befragung werden in der computerassistierten Erhebung (CAPI) alle Staaten erfasst (siehe Statistisches Bundesamt 2012b). Die Änderungen betreffen alle Variablen zu Ländern und Staatsangehörigkeiten (z. B. EF190 „Arbeitsstätte: Staat“, EF369 „1. Staatsangehörig.“ oder EF2008 „Staatsangehörig.: erweiterter Migrationsstatus“).

- Um insbesondere die Erfassung des ILO-Erwerbsstatus zu verbessern, sind seit der Erhebung 2005 Fragen zur Arbeitsmarktbeteiligung und die Feldarbeit verändert worden. In der Erhebung 2010 gibt es neue Fragen (z. B. zu bezahlten und unbezahlten Überstunden (EF162 und EF163) oder EF268 „Arbeitsuche: Warten auf Antwort der Agentur für Arbeit/ sonst. Behörde d. Arbeitsvermittlung, Arbeitnehmertätig.“). Einige Fragen weisen im Vergleich zu 2009 einen anderen Wortlaut auf bzw. enthalten andere Antwortkategorien (z. B. EF132 „Wunsch nach mehr Arbeitsstunden“). Verteilungsunterschiede können deshalb sowohl mit Veränderungen des Arbeitsmarktes als auch methodisch mit geänderten Fragen zusammenhängen (siehe Gauckler und Körner 2011)
- Weitere Änderungen betreffen die Bildungsangaben:
 Beim höchsten allgemeinen Schulabschluss (EF310) werden 2010 die Abschlüsse „Polytechnische Oberschule der DDR mit Abschluss der 8. oder 9. Klasse“ (EF310 = 2) und „Polytechnische Oberschule der DDR mit Abschluss der 10. Klasse“ (EF310 = 7) getrennt ausgewiesen.
 Beim beruflichen Abschluss ist der bis 2009 in der Sammelkategorie „Meister-/Technikerausbildung oder gleichwertiger Fachschulabschluss ...“ (EF312 = 5) enthaltene „Abschluss einer Berufsakademie“ 2010 extra (EF312 = 12) ausgewiesen.
 In den neuen Variablen EF319 „Höchster Abschluss Berufsakademie/(Fach-)Hochschule: Bezeichnung“ und EF320 „Höchster berufl. Abschluss (einschl. Abschlussgrad)“ sind die verschiedenen Abschlussarten differenzierter als zuvor erfasst. Variable EF319 unterscheidet u. a.: Bachelor (1), Master (2), Diplom, Lehramtsprüfung, Staatsprüfung, Magister, künstlerischer Abschluss und vergleichbare Abschlüsse (3). In EF320 ist der höchste berufliche Abschluss (EF312) mit dem Abschlussgrad (EF319) kombiniert.
 Die Ergebnisse zu Bildungsangaben in den Fachserien wurden auf Basis korrigierter Variablen ermittelt, die nicht in den Originaldaten vorlagen, als das SUF erstellt wurde. Beispielsweise wird der Abschluss „Haupt-(Volks-)schulabschluss“ zugeordnet, wenn als höchster allgemeinbildender Schulabschluss „Abschluss der polytechnischen Oberschule mit der 8. oder 9. Klasse“ angegeben war und die Person 1975 oder später geboren wurde. Um die in den Fachserien veröffentlichten Ergebnisse mit Auswertungen des SUF vergleichen zu können, wurden vom Statistischen Bundesamt nachträglich die entsprechenden Änderungen vorgenommen und in den folgenden Variablen mit der Extension "k" (nach Korrektur) abgelegt: EF310k, EF312k, EF320k, EF743k, EF744k, EF831k, EF832k, EF855k, EF856k.
- 2010 ist die AKE-Strukturvariable EF179 „Nichtarbeitsuche oder Teilzeitbeschäftigung: Grund (FB) (0,1 %)“ nicht belegt. Die entsprechenden Angaben wurden innerhalb des Ad-hoc-Moduls erfragt und sind in den Variablen EF1112-EF1115 enthalten.
- Das Thema des Ad-hoc-Moduls 2010 war „Vereinbarkeit von Familie und Beruf“. Die dazu gehörigen Variablen sind EF1105-EF1128 (vgl. Eurostat (2013) zum Design, zur Vergleichbarkeit mit dem gleichnamigen Ad-hoc-Modul 2005 sowie zu methodischen Problemen).
- Bei der Konstruktion des für die Merkmale des Ad-hoc-Moduls verwendeten Hochrechnungsfaktors EF956 wurden 2010 für die Schätzung der Antwortwahrscheinlichkeiten (1. Stufe) die Erhebungsmethode (EF7), das Bundesland (EF1), die Staatsangehörigkeit (EF371) und die Haushaltsgröße (EF663) verwendet. Wegen der Zielpopulation 15- bis 64-jähriger Personen, wurden keine Haushalts-, sondern Personenfaktoren geschätzt.
 Bei der Anpassung (2. Stufe) wurde auf den regionalen Ebenen Bund und Bundesländer für die Merkmale Erwerbstyp, Geschlecht und Altersklassen an hochgerechnete Stichprobenwerte des

Mikrozensus angepasst, ergänzend für Ost- (ohne Berlin) und Westdeutschland (inkl. Berlin) an Populationswerte nach Gruppen von Staatsangehörigkeiten.

- Wie bereits im SUF 2009 sind auch 2010 fehlende Werte nachträglich nach einzelnen Gründen bzw. Subpopulationen unterschieden und mit spezifischen negativen Werten ausgewiesen. Ausgenommen hiervon sind einige Variablen zu Mehrfachantworten, in denen eine 0/1-Kodierung umgesetzt wurde (EF254-EF262, EF268-EF274, EF403, EF406-EF410, EF413a, EF414, EF412, EF415-EF419, EF421-EF426, EF429-EF434, EF439, EF1000u1-EF1001u12, EF1112-EF1115, EF1118-1121).
- Zu ergänzen ist, dass bei der Überprüfung der Filterführung teilweise irreguläre fehlende Werte festgestellt wurden. Die betreffenden Fälle wurden im SUF in „keine Angabe“ recodiert (für Details siehe Statistisches Bundesamt / GESIS 2013c: 8-9).

4 Aufbereitung der Rohdaten zum Mikrozensus Scientific Use File 2010 und Überprüfung der Plausibilität des erstellten Systemfiles

Als Ergebnis der Datenaufbereitung liegen Rohdaten im numerischen ASCII-Format sowie sogenannte Systemfiles der Statistikprogramme SPSS, SAS und STATA im dokumentierten Format vor. Hinweise zur Nutzung der Setups und der Systemfiles werden in einem Dokument zu den Setups („[Readme](#)“) gegeben. Das Datenhandbuch enthält zusätzliche Informationen: Fragetext, Kommentare zu Änderungen (s. o.) und ungewichtete Häufigkeitsauszählungen. Somit können Einzelforscher nach Erhalt der Daten umgehend mit den eigentlichen Analysen beginnen. Bei der Erstellung der maschinenlesbaren Datendokumentation wurden die Variablen- und Wertelabels³ sowie die thematische Übersicht nach der Systematik des Mikrodaten-Informationssystems ([MISSY](#)) erstellt (GESIS 2013d). Die Variablenlabels enthalten die Nummer der entsprechenden Frage im Selbstausfüllerbogen, den grundsätzlichen inhaltlichen Bezug der Variable und eventuelle Spezifizierungen zu dieser Variablen; außerdem sind Freiwilligkeit und Zugehörigkeit zur Unterstichprobe gekennzeichnet. Diese Standardisierung und Anpassung an internationale Dokumentationsstandards für sozialwissenschaftliche Daten unterstützt auch die Replizierbarkeit von Analysen. Des Weiteren werden im [GESIS-Web](#) wie bisher Fragebogen und Klassifikationen sowie Mikrodaten-Tools zur Umsetzung internationaler Klassifikationen oder Skalen bereitgestellt.

Schließlich erfolgen im Rahmen der Datenaufbereitung Plausibilitätskontrollen, die nicht nur Einzelforschern zur Kontrolle dienen können, ob die erhaltenen Daten fehlerfrei zu verarbeiten sind. Ergänzende Beispiele im [GESIS-Web](#) zur Abgrenzung der amtlichen Bevölkerungsbegriffe und den für die Hochrechnung verwendeten Gewichtungsvariablen zeigen zugleich, wie mit Auswertungen des SUF zentrale Eckwerte aus den amtlichen Veröffentlichungen des Mikrozensus repliziert werden können. In diesem Zusammenhang ist aber auch darauf hinzuweisen, dass im Rahmen der Aufbereitung des Mikrozensus in den statistischen Ämtern zwar umfangreiche Datenbereinigungen erfolgen, jedoch nicht alle Angaben geprüft und ggf. plausibilisiert werden. Dies ist auch bei der Aufbereitung des Scientific Use Files nicht möglich. Einige Variablen sowie darauf aufbauende abgeleitete Variablen (Typisierungen) können daher noch Inkonsistenzen enthalten.⁴

Als Referenz zur Plausibilitätsprüfung der Mikrozensus SUF dienen die in den Fachserien des Statistischen Bundesamtes veröffentlichten Ergebnisse des jeweiligen Mikrozensus. Für den Mikrozensus 2010 sind dies insbesondere die Fachserie 1 Bevölkerung und Erwerbstätigkeit: Reihe 3 (Haushalte und Familien), Reihe 4.1.1 (Stand und Entwicklung der Erwerbstätigkeit), Reihe 4.1.2 (Beruf, Ausbildung und Arbeitsbedingungen der Erwerbstätigen), Reihe 2.2 (Bevölkerung mit Migrationshintergrund) und Fachserie 5, Heft 1. Bei diesen Vergleichen sind die von den Statistischen Ämtern verwendeten Bevölkerungs- und Erwerbskonzepte sowie die Hochrechnung zu beachten (vgl. die Seiten [Abgrenzung und Hochrechnung der Bevölkerungsbegriffe im Mikrozensus 2010](#) sowie den [Definitionskatalog im GESIS-Web](#) (2013a, b)).

Bedingt durch die Ziehung der 70 %-Substichprobe ergeben sich Abweichungen zwischen den Häufigkeiten der Variablen des Mikrozensus SUF und den in den Fachserien des Statistischen Bundesamtes veröffentlichten Zahlen bzw. der Original-Mikrozensusdaten; siehe die Tabelle [Überprüfung der Randverteilung](#) (Statistisches Bundesamt 2013b). Die meisten Variablen des SUF weichen nur in sehr gering-

³ Geschlechtsspezifische Berufsbezeichnungen sind in den Setups zum Mikrozensus ab 2006 – im Unterschied zu den Setups früherer Erhebungen – ausschließlich aus Platzgründen und zur besseren Lesbarkeit in der männlichen Form genannt.

⁴ Siehe dazu die Hinweise im [Datenhandbuch](#) (Statistisches Bundesamt / GESIS 2013a: 5).

gem Maße von den veröffentlichten Daten ab. Größere relative Abweichungen sind nur bei Merkmalen festzustellen, die mit sehr geringen Fallzahlen besetzt sind.

Die folgenden Tabellen 1 und 2 enthalten zentrale Bevölkerungsbegriffe des Mikrozensus und ausgewählte Randverteilungen. Zum Grundprogramm des Mikrozensus werden die Stellung im Beruf und der Erwerbstyp, die Haushaltsgröße und der Migrationsstatus tabelliert. Beispielhaft für Analysen der Substichprobe mit Merkmalen des EU Labour Force Survey werden Häufigkeiten der AKE-Strukturvariable Leitungstätigkeit und der Variable Betreuung von Kindern Anderer des Ad-hoc-Moduls berichtet. Zum 2010 durchgeführten Zusatzprogramm skizzieren Verteilungen zur Anzahl der Wohnungen im Gebäude und zur Nutzung der Wohnung weitere Anwendungsmöglichkeiten. Im Anhang dieses Berichts befindet sich die SPSS-Syntax zu diesen Tabellen. Sie kann ebenfalls für Überprüfungen der Randverteilungen des SUF oder als Startpunkt für eigene Analysen zu diesen in der Tabelle berichteten Themen verwendet werden.

Tabelle 1: Bevölkerungsbegriffe im Mikrozensus Scientific Use File 2010 (Jahresdurchschnitt in 1.000)

Bevölkerung am Haupt- und Nebenwohnsitz^a (EF30>=1 & EF30<=3)		
82.695		
Bevölkerung am Hauptwohnsitz^b (EF30>=1 & EF30<=2)	81.715	Bev. am Nebenwohnsitz ^c (EF30=3) 979
Bevölkerung in Privathaushalten^d (EF31=1)	81.776	Bev. in Gemeinschaftsunterk. ^e (EF31=2) 918
Bevölkerung in Lebensformen am Hauptwohnsitz^f [Bezugsperson (BP) am Hauptwohnsitz] (EF761>=1 & EF761<=2)	80.964	- [BP am Nebenwohns.] (EF761>2) 812

- a Zur Bevölkerung am Haupt- und Nebenwohnsitz (früher: „wohnberechtigte Bevölkerung“), zählen alle in der Gemeinde wohnhaften Personen. Mehrfachzählungen sind möglich.
- b Zur Bevölkerung am Hauptwohnsitz („Bevölkerung“) zählen alle Personen mit nur einer Wohnung (EF30=1) sowie Personen mit mehreren Wohnungen am Ort ihrer Hauptwohnung (EF30=2).
- c Zur Bevölkerung am Nebenwohnsitz zählen Personen mit mehreren Wohnungen, die am Nebenwohnsitz befragt wurden.
- d Zur Bevölkerung in Privathaushalten zählen alle Personen, die am Haupt- oder Nebenwohnsitz allein oder zusammen mit anderen Personen eine wirtschaftliche Einheit (Privathaushalt) bilden. Mehrfachzählungen sind möglich.
- e Zur Bevölkerung in Gemeinschaftsunterkünften gehören alle Personen, die dort wohnen, bzw. gemeldet sind, und nicht für sich wirtschaften, d. h. keinen eigenen Haushalt führen. (Ohne Personen, die sich nur vorübergehend (weniger als drei Monate) dort aufhalten.)
- f Zur Bevölkerung in Lebensformen am Hauptwohnsitz in Privathaushalten zählen alle Mitglieder einer Familie/ Lebensform (z. B. nichteheliche Lebensgemeinschaft mit Kindern), deren Bezugsperson am Ort der Hauptwohnung lebt. Die Lebensform wird als zusammengehörige Einheit betrachtet, unabhängig davon, ob einzelne Personen eventuell einen weiteren Wohnsitz haben. Für die Typisierung ist der (Haupt-) Wohnsitz der Bezugsperson relevant. Potenzielle Mehrfachzählungen werden dadurch vermieden.

Tabelle 2: Ausgewählte Verteilungen des Mikrozensus Scientific Use Files 2010 (Jahresdurchschnitt in 1.000)

Merkmal	Hochrechnung (in 1.000)
Bevölkerungsgruppen	
Bevölkerung am Haupt- und Nebenwohnsitz (Fallzahl insgesamt: n = 489.630)	82.695
Bevölkerung in Privathaushalten	81.776
Bevölkerung am Ort der Hauptwohnung	81.715
Bevölkerung in Lebensformen am Hauptwohnsitz	80.964
Bevölkerung am Ort der Hauptwohnung, Erwerbstätige	38.965
EF117: Stellung im Beruf x EF29: Erwerbstyp	
Bevölkerung am Ort der Hauptwohnung (siehe Fachserie 1, Reihe 4.1.1, Tabelle 1.1)	
Selbstständige	4.256
Mithelf. Familienangeh.	217
Beamte (inkl. Soldaten)	2.079
Angestellte (inkl. kfm./ techn. Auszubildende, Zivildienstleistende, geringfügig beschäftigte Schüler, Studenten, Rentner)	22.586
Arbeiter (inkl. gew. Auszubildende)	9.827
Erwerbstätige insgesamt	38.965
Erwerbslose	2.934
Erwerbspersonen	41.899
Nichterwerbspersonen	39.816
Arbeitssuchende Nichterwerbspersonen	469
Sonstige Nichterwerbspersonen	39.347
Insgesamt	81.715
EF663: Personen im Haushalt: Anzahl (Haushaltstypisierung)	
Bevölkerung in Privathaushalten, Haushaltsbezugsperson [Haupteinkommensbezieher] (siehe Fachserie 1, Reihe 3, Tabelle 1.1)	
1 Person	16.184
2 Personen	13.786
3 Personen	5.090
4 Personen	3.844
5 Personen und mehr	1.382
Haushalte insgesamt	40.286

Merkmal	Hochrechnung (in 1.000)
EF2001: Migrationsstatus (differenz. Angabe)	
Bevölkerung am Ort der Hauptwohnung (siehe Fachserie 1, Reihe 2.2, Tabelle 2)	
Person ohne Migrationshintergrund	65.971
Personen mit Migrationshintergrund im engeren Sinn	15.745
Bevölkerung mit Migrationshintergrund und eigener Migrationserfahrung	10.605
Ausländer	5.583
Deutsche	5.022
(Spät-) Aussiedler	3.282
durch Einbürgerung	1.740
Personen ohne eigene Migrationserfahrung	5.140
Ausländer	1.555
Deutsche	3.584
durch Einbürgerung	397
mit mind. einem zugewanderten oder als Ausländer in Deutschland geborenen Elternteil	3.187
mit beidseitigem Migrationshintergrund	1.654
mit einseitigem Migrationshintergrund	1.533
Insgesamt	81.715
EF118: Leitungstätigkeit (FB) (0,1 %)	
Unterstichprobe, erwerbstätige Bevölkerung nach EU-Definition	
Gewichtungsvariable: EF955 Hochrechnungsfaktor AKE-Strukturvariablen: Jahr (in 1000)	
Ja	5.556
Nein	21.272
Keine Angabe	7.562
Trifft nicht zu, da selbstständig	4.405
Insgesamt	38.796
EF1118: Regelmäßige, nicht gewerbl. Betreuung: Kinder Anderer (unter 15 Jahren) (FB) (0,1%)	
Unterstichprobe, Bevölkerung nach EU-Definition	
Gewichtungsvariable: EF956 Hochrechnungsfaktor Ad-hoc-Substichprobe: Jahr (in 1000)	
Betreuung Kinder Anderer (unter 15 Jahren) genannt	3.637

Merkmal	Hochrechnung (in 1.000)
EF635: Anzahl der Wohnungen im Gebäude	
Privathaushalte ohne Jahresüberhang, Wohnungen in Wohngebäuden	
Gewichtungsvariable: EF960 Hochrechnungsfaktor Wohnsituation: Jahr (in 1000)	
(Fachserie 5 Heft 1, Tab. 1; siehe auch dort den folgenden Hinweis:	
"Die Ergebnisse dieser Veröffentlichung werden voraussichtlich revidiert werden. Nach Abschluss der Datenaufbereitung zeigte sich ein Fehler, der nach ersten Analysen jedoch nur eine geringe Zahl der befragten Haushalte betrifft. Die exakten Auswirkungen auf die Ergebnisse dieser Veröffentlichung sind allerdings derzeit noch nicht quantifizierbar."	
1 Wohnung	10.612
2 Wohnungen	6.625
3-6 Wohnungen	7.555
7-12 Wohnungen	7.466
13-20 Wohnungen	1.632
21 oder mehr Wohnungen	2.134
Insgesamt	36.024
EF491: Nutzung der Wohnung	
Privathaushalte ohne Jahresüberhang, Wohngebäude	
Gewichtungsvariable: EF952 Standardhochrechnungsfaktor Jahr (in 1000)	
(Fachserie 5 Heft 1, Tab. 21)	
Eigentümer	16.921
Hauptmieter	20.375
Untermieter	886
Insgesamt	38.182

Literatur

- Afentakis, Anja, und Wolf Bihler, 2005: Das Hochrechnungsverfahren beim unterjährig Mikrozensus ab 2005. *Wirtschaft und Statistik* (10): 1039-1048. URL: https://www.destatis.de/DE/Publikationen/WirtschaftStatistik/AllgemeinesMethoden/Hochrechnungunterjaehrig.pdf?__blob=publicationFile
- Eurostat, 2012a: Labour force survey in the EU, Candidate and EFTA countries - Main characteristics of the national surveys - 2010. URL: <http://www.eds-destatis.de/downloads/publ/KS-RA-12-013-EN-N.pdf>
- Eurostat, 2012b: Quality report of the European Union Labour Force Survey 2010. Luxembourg: Publications Office of the European Union. URL: <http://www.eds-destatis.de/downloads/publ/KS-RA-12-021-EN-N.pdf>
- Eurostat, 2013: Evaluation of the LFS ad hoc module 2010 on reconciliation between work and family life. Final report. URL: http://epp.eurostat.ec.europa.eu/portal/page/portal/employment_unemployment_lfs/documents/Evaluation%20report%20AHM%202010.pdf
- Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder [FDZ], 2013: Datenangebot, Mikrozensus. URL: <http://www.forschungsdatenzentrum.de/bestand/mikrozensus/index.asp>
- Gauckler, Britta, und Thomas Körner, 2011: Measuring the Employment Status in the Labour Force Survey and the German Census 2011. Insights from Recent Research at Destatis. *Methoden - Daten - Analysen*, 5 (2): 181-205. URL: http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/mda/Vol.5_Heft_2/MDA_2011_2_Fb_2_Gaukler.pdf
- GESIS, 2013a: Abgrenzung und Hochrechnung der Bevölkerungsbegriffe im Mikrozensus 2010. Mannheim. URL: <http://www.gesis.org/missy/studie/erhebung/studienbeschreibung/hochrechnung/zur-abgrenzung-der-bevoelkerungsbegriffe-und-zur-hochrechnung-im-mikrozensus-suf-2010/>
- GESIS, 2013b: Definitionskatalog. Mannheim. URL: <http://www.gesis.org/missy/studie/klassifikationen/konzepte-und-definitionen/definitionskatalog/>
- GESIS, 2013c: Informationen zum Mikrozensus 2010. Mannheim. URL: <http://www.gesis.org/missy/missy-home/auswahl-datensatz/mikrozensus-2010/>
- GESIS, 2013d: MISSY - Mikrodaten-Informationssystem. Mannheim. URL: <http://www.gesis.org/missy/>
- Schimpl-Neimanns, Bernhard, 2011: Schätzung des Stichprobenfehlers in Mikrozensus Scientific Use Files ab 2005, *AStA Wirtschafts- und Sozialstatistisches Archiv* 5 (1): 19-38. (Siehe auch die unter <http://www.gesis.org/dienstleistungen/tools-standards/mikrodaten-tools/bereitgestellten/Programme>).
- Statistisches Bundesamt, 2011a: Qualitätsbericht Mikrozensus 2010. Wiesbaden. URL: https://www.destatis.de/DE/Publikationen/Qualitaetsberichte/Bevoelkerung/Mikrozensus2010.pdf?__blob=publicationFile
- Statistisches Bundesamt, 2011b: Fachserie 1 Bevölkerung und Erwerbstätigkeit, Reihe 4.1.1 Stand und Entwicklung der Erwerbstätigkeit. Mikrozensus 2010. URL: https://www.destatis.de/DE/Publikationen/Thematisch/Arbeitsmarkt/Erwerbstaetige/StandEntwicklungErwerbstaetigkeit2010411107004.pdf?__blob=publicationFile

- Statistisches Bundesamt, 2011c: Fachserie 1 Bevölkerung und Erwerbstätigkeit, Reihe 4.1.2 Beruf, Ausbildung und Arbeitsbedingungen der Erwerbstätigen. Mikrozensus 2010. Wiesbaden. URL: https://www.destatis.de/DE/Publikationen/Thematisch/Arbeitsmarkt/Erwerbstaetige/BerufArbeitsbedingungErwerbstaetigen2010412107004.pdf?__blob=publicationFile
- Statistisches Bundesamt, 2011d: Fachserie 1 Bevölkerung und Erwerbstätigkeit, Reihe 2.2 Bevölkerung mit Migrationshintergrund. Ergebnisse des Mikrozensus 2010. Wiesbaden. URL: https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/MigrationIntegration/Migrationshintergrund2010220107004.pdf?__blob=publicationFile
- Statistisches Bundesamt, 2011e: Fachserie 1 Bevölkerung und Erwerbstätigkeit, Reihe 3 Haushalte und Familien. Ergebnisse des Mikrozensus 2010. Wiesbaden. URL: https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/HaushalteMikrozensus/HaushalteFamilien2010300107004.pdf?__blob=publicationFile
- Statistisches Bundesamt, 2012a: Fachserie 5, Heft 1. Bauen und Wohnen. Mikrozensus-Zusatzerhebung 2010. Bestand und Struktur der Wohneinheiten. Wohnsituation der Haushalte. Wiesbaden. URL: https://www.destatis.de/DE/Publikationen/Thematisch/EinkommenKonsumLebensbedingungen/Wohnen/WohnsituationHaushalte2055001109004.pdf?__blob=publicationFile
- Statistisches Bundesamt, 2012b: Liste der Staaten und Regionen im Mikrozensus 2010. Bonn. URL: http://www.gesis.org/missy/fileadmin/missy/klassifikationen/Amtliche_Klassifikationen/SL/Staatenliste_2010.pdf
- Statistisches Bundesamt, 2012c: Mikrozensus: Haushaltszahlen ab 2005. Wiesbaden. URL: https://www.destatis.de/DE/Methoden/MikrozensusHaushaltszahlen.pdf?__blob=publicationFile
- Statistisches Bundesamt, 2013a: Statistik von A bis Z: Mikrozensus. URL: <https://www.destatis.de/DE/Meta/AbisZ/Mikrozensus.html>
- Statistisches Bundesamt, 2013b: Überprüfung der Randverteilung. Bonn. URL: http://www.gesis.org/missy/fileadmin/missy/daten/Randverteilung/Randverteilung_2010.pdf
- Statistisches Bundesamt, GESIS (Hrsg.), 2013a: Datenhandbuch zum Mikrozensus Scientific Use File 2010. URL: http://www.gesis.org/missy/fileadmin/missy/erhebung/datenhandbuch/DHB_2010.pdf
- Statistisches Bundesamt, GESIS (Hrsg.), 2013b: Thematische Übersicht zum Mikrozensus Scientific Use File 2010. URL: http://www.gesis.org/missy/fileadmin/missy/daten/thematische_uebersicht/Thematische_Uebersicht_MZ2010.pdf
- Statistisches Bundesamt, GESIS (Hrsg.), 2013c: Wichtige Informationen zur Nutzung des Mikrozensus Scientific Use Files 2010. [Readme zu den Setups] URL: http://www.gesis.org/missy/fileadmin/missy/daten/setups/mz2010/spss_setup_10.zip

Letzter Zugriff bei allen Web-Dokumenten: 04. 06. 2013.

Anhang

```
* SPSS-Syntax zu Tabellen 1 und 2.

* IBM SPSS Statistics 21, Version 21.0.0.1

* In spitzen Klammern < > stehende Verzeichnisse und Dateinamen ersetzen.
* CD '<Arbeitsverzeichnis>'.

get file='<mz2010.sav>'.

* MZ-GRUNDPROGRAMM, BEVÖLKERUNGSBEGRIFFE.

temporary.
* Alle Missing Value Deklarationen aufheben.
missing values all().

compute bev = (EF30>=1 & EF30<=3).
var lab bev 'Bevölkerung am Haupt- und Nebenwohnsitz (EF30=1-3)'.

* Siehe Fachserie 1 Reihe 4.1.1, Tab. 1.1.
compute bev_hw = (EF30>=1 & EF30<=2).
var lab bev_hw 'Bevölkerung am Hauptwohnsitz (EF30=1-2)'.

* Siehe Fachserie 1 Reihe 4.1.1, Tab. 1.2.
compute bev_et = (EF30>=1 & EF30<=2 & EF29=1).
var lab bev_et 'Bevölkerung am Hauptwohnsitz, Erwerbstätige
(EF30=1-2 & EF29=1)'.

compute bev_nw = (EF30=3).
var lab bev_nw 'Bevölkerung am Nebenwohnsitz (EF30=3)'.

* Siehe Fachserie 1 Reihe 3, Tab. 2.1.
compute bev_ph = (EF31=1).
var lab bev_ph 'Bevölkerung in Privathaushalten (EF31=1)'.

compute bev_gu = (EF31=2).
var lab bev_gu 'Bevölkerung in Gemeinschaftsunterkünften (EF31=2)'.

* Siehe Fachserie 1 Reihe 3, Tab. 4.1.
compute bev_lf = (EF761>=1 & EF761<=2).
var lab bev_lf 'Bevölkerung in Lebensformen am Hauptwohnsitz (EF761=1-2)'.

compute bev_lfnw = (EF761>2).
var lab bev_lfnw 'Bevölkerung in Lebensformen am Nebenwohnsitz (EF761>2)'.

recode bev bev_hw bev_et bev_nw bev_ph bev_gu bev_lf bev_lfnw (0=sysmis).
formats bev bev_hw bev_et bev_nw bev_ph bev_gu bev_lf bev_lfnw (f1.0).
* Fallzaehlung.
compute fz = 1.
formats fz (f1.0).
* Hochrechnung: EF952 Jahresdurchschnitt .
means tables EF952 fz
      by bev bev_ph bev_hw bev_lf bev_et bev_nw bev_gu bev_lfnw
      /cells sum.

* MZ-GRUNDPROGRAMM: Erwerbsbeteiligung .
* EF117: Stellung im Beruf x EF29: Erwerbstyp, .
* Filter: Bevölkerung am Ort der Hauptwohnung.
* Hochrechnung: Jahresdurchschnitt EF952.
* Fachserie 1, Reihe 4.1.1, Tabelle 1.1.
```

```

temporary.
* Alle Missing Value Deklarationen aufheben.
missing values all().
do if (EF29=1).
    recode EF117 (2 1=1) (3=2) (4 9 10=3) (5 7 11 12=4) (6 8=5) (else=-1)
into stib.
else if (EF29=2).
    compute stib = 6.
else if (EF29=3).
    compute stib = 7.
else if (EF29=4).
    compute stib = 8.
end if.
var lab stib 'Stellung im Beruf (EF117) x Erwerbsstatus (EF29)'.
value label stib
    1 'Selbstständige'
    2 'Mithelfende Familienangehörige'
    3 'Beamate (inkl. Zeit-/Berufssoldat (einschl. BGS und
        Bereitschaftspolizei), Grundwehrdienstleistender)'
    4 'Angestellte (inkl. kaufm./techn. Auszubildende, Zivildienstleistender)'
    5 'Arbeiter (inkl. gewerbl. Auszubildende)'
    6 'Erwerbslose'
    7 'Nichterwerbspersonen: Arbeitssuchend'
    8 'Nichterwerbspersonen: Sonstige'
    -1 'Entfällt'.
missing value stib (6 7 8).
recode EF29 (1 2=1) (3 4=2) into erwtyp.
var label erwtyp 'Erwerbstyp (EF29 rekodiert)'.
val label erwtyp 1 'Erwerbspersonen'
                2 'Nichterwerbspersonen'.
formats stib erwtyp (f1.0).
* Filter: Bevölkerung am Ort der Hauptwohnung.
select if (EF30>=1 & EF30<=2).
* Hochrechnung: Jahresdurchschnitt EF952.
weight by EF952.
fre stib erwtyp.

* MZ-GRUNDPROGRAMM: Haushalte.
* Haushaltsgröße: EF663.
* Fachserie 1, Reihe 3, Tabelle 1.1 .
* HAUSHALTE := Bevölkerung in Privat-HH (EF31=1) .
*           Haushaltsbezugsperson = Haupteinkommensbezieher (EF662=1-4).
temporary.
* Alle Missing Value Deklarationen aufheben.
missing values all().
recode EF663 (5 thru 9=5) (else=copy) into v663.
var lab v663 'Haushaltsgröße'.
formats v663 (f1.0).
val lab v663
    1 '1 Person'
    2 '2 Personen'
    3 '3 Personen'
    4 '4 Personen'
    5 '5 und mehr Personen'.
* Mehrpersonenhaushalte.
recode EF663 (1=1) (2 thru 9=2) into v663m.
var lab v663m 'Haushaltstyp Ein-/Mehrpers.HH.'.
formats v663m (f1.0).
val lab v663m
    1 '1 Personenhaushalt'
    2 'Mehrpersonenhaushalt'.
* Fallzaehlung.
compute fz = 1.
formats fz (f1.0).

```

```

* Haushalte:= Haushaltsbezugsperson = Haupteinkommensbezieher.
select if (EF31=1 & EF662>=1 & EF662<=4).
means tables EF952 fz by v663m v663
      /EF952 fz by v663m by v663
      /cells sum.

* MZ-GRUNDPROGRAMM: Migrationsstatus .
* Fachserie 1, Reihe 2.2, Tabelle 2.
* Definition D* entspricht „Lfd. Nr.“ in Tab. 2.
temporary.
* Alle Missing Value Deklarationen aufheben.
missing values all().
recode ef2001 (0 thru 64=1) (else=sysmis) into D001.
recode ef2001 (0 11=1) (else=sysmis) into D002.
recode ef2001 (21 thru 64=1) (else=sysmis) into D005.
recode ef2001 (21 22 25 29 31 41=1) (else=sysmis) into D036.
recode ef2001 (41=1) (else=sysmis) into D066.
recode ef2001 (21 22 25 29 31=1) (else=sysmis) into D095.
recode ef2001 (21 22 25 29=1) (else=sysmis) into D096.
recode ef2001 (31=1) (else=sysmis) into D109.
recode ef2001 (32 42 thru 43 52 thru 64=1) (else=sysmis) into D133.
recode ef2001 (42 43=1) (else=sysmis) into D140.
recode ef2001 (32 52 thru 64=1) (else=sysmis) into D147.
recode ef2001 (32=1) (else=sysmis) into D148.
recode ef2001 (52 thru 64=1) (else=sysmis) into D154.
recode ef2001 (62 63 64=1 ) (else=sysmis) into D155.
recode ef2001 (52 thru 59=1 ) (else=sysmis) into D156.
formats D001 to D156 (f1.0).
var lab D001 'Bevölkerung insgesamt'.
var lab D002 'Personen ohne Migrationshintergrund'.
var lab D005 'Personen mit Migrationshintergrund im engeren Sinn'.
var lab D036 'Personen mit eigener Migrationserfahrung'.
var lab D066 '- Ausländer'.
var lab D095 '- Deutsche'.
var lab D096 ' * (Spät-)Aussiedler'.
var lab D109 ' * durch Einbürgerung'.
var lab D133 'Personen ohne eigene Migrationserfahrung'.
var lab D140 '- Ausländer'.
var lab D147 '- Deutsche'.
var lab D148 ' * durch Einbürgerung'.
var lab D154 ' * mit mindestens einem zugewanderten oder als Ausländer in
      Deutschland geborenen Elternteil'.
var lab D155 ' + mit beidseitigem Migrationshintergrund'.
var lab D156 ' + mit einseitigem Migrationshintergrund'.
* Bevölkerung am Hauptwohnsitz.
select if ef30>=1 & ef30<=2 .
weight by ef952.
descriptives variables = D001 to D156.

* AKE-STRUKTURVARIABLEN (0,1%): EF118, EF177, EF178.
* Filter: EF6=2 Unterstichprobe, Fragenprogramm .
*       EF32=1 Bevölkerungsabgrenzung nach EU-Definition .
*       [Bevölkerung am Hauptwohnsitz ohne und mit weiterem
*       Wohnsitz ohne Wehr- und Zivildienstleistende und ohne
*       Bevölkerung in Gemeinschaftsunterkünften] .
* Gewichtung: EF955 Hochrechnungsfaktor AKE-Strukturvariablen Jahr (1.000).
temporary.
* Alle Missing Value Deklarationen aufheben.
missing values all().
* Beispiel EF118: Leitungstätigkeit (FB) (0,1%).
* Selektion: Unterstichprobe (EF6=2), erwerbstätige Bevölkerung (EF29=1),
      Bevölkerung nach EU-Definition (EF32=1).

```

```

select if (EF6=2 & EF29=1 & EF32=1).
* GewichtungsvARIABLE: EF955 Hochrechnungsfaktor Ad-hoc-Substichprobe.
weight by EF955.
missing values EF118(-5).
fre EF118.

* AD-HOC-MODUL VEREINBARKEIT VON FAMILIE UND BERUF .
* Filter: EF6 Unterstichprobe, Fragenprogramm .
* Gewichtung: EF956 Hochrechnungsfaktor Ad-hoc-Substichprobe: Jahr (in
1000).
* EF1118 F204 Regelmäßige, nicht gewerbl. Betreuung: Kinder Anderer (unter
15 Jahren) (FB) (0,1%)
* [1] Betreuung Kinder Anderer (unter 15 Jahren) genannt.
* EF6 Unterstichprobe, Fragenprogramm
* [2] Mit Ad-hoc-Modul .
* EF32 Bevölkerung: Eurostat-Definition
* [1] Bevölkerungsabgrenzung nach EU-Definition.
* EF956 Hochrechnungsfaktor Ad-hoc-Substichprobe: Jahr (in 1000).
* Siehe http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lfso\_10cregcar&lang=de .
temporary.
* Alle Missing Value Deklarationen aufheben.
missing values all().
* Substichprobe & Bevölkerung nach EU-Definition.
select if (EF6=2 & EF32=1).
select if (EF44>=15 & EF44<=64) /* Alter */ .
* EF956 Hochrechnungsfaktor Ad-hoc-Substichprobe: Jahr (in 1000) .
* EF1118 = 1 Betreuung Kinder Anderer (unter 15 Jahren) genannt .
weight by EF956.
fre EF1118.

* MZ-ZUSATZPROGRAMM WOHN-SITUATION DER HAUSHALTE:
* Anzahl der Wohnungen im Gebäude (EF635) x Nutzung der Wohnung (EF641).
* Siehe Fachserie 5 Heft 1 .
* Hinweis in Fachserie 5, Heft 1: .
* "Die Ergebnisse dieser Veröffentlichung werden voraussichtlich revidiert
* werden. Nach Abschluss der Datenaufbereitung zeigte sich ein Fehler, der
* nach ersten Analysen jedoch nur eine geringe Zahl der befragten Haushalte
* betrifft. Die exakten Auswirkungen auf die Ergebnisse dieser
* Veröffentlichung sind allerdings derzeit noch nicht quantifizierbar." .
temporary.
* Alle Missing Value Deklarationen aufheben.
missing values all().
* Fachserie 5 Heft 1, Tab. 1..
select if (EF31=1 /* Privathaushalte */
& EF5b=0 /* Jahresbestand */
& EF489=1 /* Wohngebäude */
& EF642>=1 & EF642<=2) /* Bezugsperson der Wohnung */ .
recode EF635 (3 thru 6=3) (else=copy) into v635.
formats v635 (f1.0).
var lab v635 'Wohnungen im Gebäude: Anzahl (Typisierung) '.
val lab v635
1 '1 Wohnung'
2 '2 Wohnungen'
3 '3-6 Wohnungen'
7 '7-12 Wohnungen'
8 '13 bis 20 Wohnungen'
9 '21 oder mehr Wohnungen'
-5 'Entfällt (Jahresüberhang) '
-1 'Entfällt (Gemeinschaftsunterkunft)'.
recode EF641 (1 2 = 1) (3 4 5 = 2) (else=copy) into v641.
var lab v641 'Nutzung der Wohnung'.

```

```
val lab v641
  1 'Eigentümerwohnungen'
  2 'Mietwohnungen'
-5 'Entfällt (Jahresüberhang)'
-1 'Entfällt (Gemeinschaftsunterkunft)'.
formats v635 (f1.0).
* EF960 Hochrechnungsfaktor Wohnsituation: Jahr (in 1000).
weight by EF960.
crosstabs v635 by v641.

* Haushaltgröße und Art der Nutzung der Wohnung .
* Siehe Fachserie 5 Heft 1, Tab. 21.
temporary.
* Alle Missing Value Deklarationen aufheben.
missing values all().
select if (EF31=1 /* Privathaushalte */
          & EF5b=0 /* Jahresbestand */
          & EF489=1 /* Wohngebäude */
          & EF662>=1 & EF662<=4) /* Haupteinkommensbezieher im Haushalt */ .
recode EF491 (1 2=1) (3=2) (4=3) (else=sysmis) into v491.
formats v491 (f1.0).
var lab v491 'Nutzung Wohnung'.
val lab v491
  1 'Eigentümer'
  2 'Hauptmieter'
  3 'Untermieter'.
recode EF663 (5 thru 9=5) (else=copy) into v663.
var lab v663 'Haushaltsgröße'.
formats v663 (f1.0).
val lab v663
  1 '1 Person'
  2 '2 Personen'
  3 '3 Personen'
  4 '4 Personen'
  5 '5 und mehr Personen'.
weight by EF952 /* Hochrechnung: Jahresdurchschnitt EF952 */ .
crosstabs v663 by v491.
```