

Mikrozensus Scientific Use File 2011 Dokumentation und Datenaufbereitung

Bernhard Schimpl-Neimanns, Andreas Herwig

GESIS-Technical Reports 2014|08

Mikrozensus Scientific Use File 2011

Dokumentation und Datenaufbereitung

Bernhard Schimpl-Neimanns, Andreas Herwig

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

Mikrozensus Scientific Use File 2011: Dokumentation und Datenaufbereitung

Bernhard Schimpl-Neimanns, Andreas Herwig

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 2011 skizziert. Diese Daten umfassen das Zusatzprogramm mit Informationen zur Krankenversicherung sowie zusätzlichen Angaben zur Erwerbstätigkeit und das Ad-hoc-Modul zum Thema Beschäftigung behinderter Personen 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 2011: 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 2011 will be discussed. The survey 2011 includes the supplementary programs on health insurance, additional information on employment, and the ad hoc module employment of disabled people 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. Von den 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 2011 zunächst die Datenaufbereitung und die Schritte zur Prüfung der Daten beschrieben bevor die wichtigsten Veränderungen im Vergleich zum SUF 2010 benannt und die Verwendung der amtlichen Bevölkerungskonzepte sowie die Hochrechnung beispielhaft erläutert werden.

¹ Weitere Informationen sind auf den Web-Seiten des Statistischen Bundesamtes (2014c; [Statistik von A bis Z](#)) und der Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder (FDZ 2014; 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 2011](#) des Statistischen Bundesamtes (2012g), der auch in den Fachserien zum Mikrozensus enthalten ist, sowie zwei Berichte zum LFS 2011, den [Qualitätsbericht](#) (Eurostat 2013) und die „[Main characteristics of national surveys, 2011](#)“ (Eurostat 2012)).

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 2011 umfasst Angaben zur Krankenversicherung sowie Zusatzangaben zur Erwerbstätigkeit.

Darüber hinaus wurde im Rahmen der EU-Arbeitskräfteerhebung (AKE) 2011 das Ad-hoc-Modul zum Thema „Beschäftigung behinderter Personen“ 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. 25 %). 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, sodass 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 2011 2,4 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 und des Ad-hoc-Moduls des EU Labour Force Surveys liegen für alle Merkmalsbereiche gleiche Hochrechnungsfaktoren vor.

3 Das Scientific Use File Mikrozensus 2011

Wie bereits einleitend erwähnt, ist das SUF eine faktisch anonymisierte 70 %-Substichprobe; i. d. R. eine Haushaltssubstichprobe. 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-Strukturvariablen 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 „Haushalts“-Nummer. Zur Stichprobenziehung dient die letzte (Einer-) Stelle der Haushaltsnummer. Alle Haushalte, deren Einerstelle nicht mit drei zufällig gezogenen einstelligen Zahlen übereinstimmen, gelangen nach diesem sogenannten Schlussziffernverfahren in die Stichprobe (70 % der Haushalte). 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 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. 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 2010

Um Nutzern, die bereits mit dem SUF früherer Erhebungszeitpunkte gearbeitet haben, die Auswertung zu erleichtern, werden im Folgenden die wichtigsten Veränderungen gegenüber dem MZ 2010 sowie

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 2014b).

zum Zusatzprogramm im MZ 2007 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 2011“ (Statistisches Bundesamt / GESIS 2014c) sowie im Datenhandbuch (Statistisches Bundesamt / GESIS 2014a), das in der Variablenbeschreibung ebenfalls erläuternde Kommentare mit Hinweisen zur zeitlichen Vergleichbarkeit sowie den Fragetext des Selbstausfüllerbogens enthält. Zur Vergleichbarkeit mit weiteren Erhebungen des Mikrozensus ab 1973 enthält die Variablen-Zeitpunkte-Matrix in MISSY zusätzliche Hinweise.

- Um insbesondere die Erfassung des Erwerbsstatus der internationalen Arbeitsorganisation (ILO) zu verbessern, sind seit der Erhebung 2005 Fragen zur Arbeitsmarktbeteiligung und die Feldarbeit verändert worden. In der Erhebung 2011 wurden einige Fragen des Mikrozensus und der Arbeitskräfteerhebung im Vergleich zum Vorjahr verändert. Im Rahmen eines Methodenprojekts wurde dies mit einem Fragebogensplit durchgeführt. Die neuen Fragen wurden im Rahmen einer 0,9 %-Unterstichprobe (USP) gestellt. Die Fragen in der 0,1 %-USP entsprechen dagegen weitestgehend dem ursprünglichen Konzept. Von zentraler Bedeutung ist eine neue Frage, die das Themenfeld Beschäftigungssituation einleitet. Anhand dieser Frage zum sogenannten Hauptstatus (F18; EF66) wird der aktuelle soziale Status erfragt. (Die Variable EF66 ist jedoch nicht Bestandteil des SUF.³) Erst im Anschluss daran werden die für die Umsetzung des Labour-Force Konzepts der ILO relevanten Informationen über das Vorliegen einer bezahlten Tätigkeit erfasst (Statistisches Bundesamt 2012f). Diese neue Konzeption hatte zusätzliche Veränderungen im Fragebogen zur Folge. So wurden einige der in der 0,1 %-USP noch enthaltenen Fragen in der 0,9 %-USP überflüssig (z. B. F20a; EF81), andererseits mussten neue Fragen hinzugefügt (z. B. F19; EF67) oder neu platziert werden (F22/F18a; EF77). Darüber hinaus wurden Fragen teilweise neu formuliert (z. B. F23/F22a; EF85) oder vereinfacht (z. B. F22/F18a; EF77). Die Änderungen haben insgesamt betrachtet zu einer besseren Erfassung der Erwerbstätigkeit bzw. Erwerbslosigkeit geführt und die Unterschiede zu Referenzstatistiken verringert. Das neue Konzept wird deshalb im Mikrozensus ab 2012 fortgeführt. Aus methodischen Gründen sind die Ergebnisse für 2011 mit den Vorjahren nur eingeschränkt vergleichbar. Zu beachten ist des Weiteren, dass die Angaben zu diesen verschiedenen Fragebogenversionen in die einheitlichen Typisierungen (z. B. EF29 Erwerbstyp oder EF38 Nichterwerbstyp) eingeflossen sind. Siehe hierzu die Gegenüberstellung der Fragebogenversionen (Herter-Eschweiler 2014) sowie die methodischen Hinweise im Datenhandbuch. In den Systemfiles sind die betreffenden Variablen durch Angabe der Fragenummern (F.../...) und des Auswahlsatzes der Unterstichprobe (0,9 % / 0,1 %) im Variablenlabel gekennzeichnet. Es hat sich gezeigt, dass bei einigen Haushalten, deren Angaben per Selbstausfüllerbogen erhoben wurden, der falsche Fragebogen zum Einsatz kam. Haushalte, die für die 0,1 %-USP vorgesehen waren, erhielten den Fragebogen für die 0,9 %-USP und umgekehrt. Dies war bei insgesamt 0,2 % der befragten Haushalte der Fall. Bei den Variablen des Ad-hoc-Moduls werden die falschen Zuordnungen als Unit-Nonresponse behandelt, bei den AKE-Strukturvariablen als Item-Nonresponse.
- Im Rahmen der in 4-jährigen Abständen wiederholten Zusatzprogramme sind im Mikrozensus 2011 folgende Variablen neu:

EF172-EF176 *Zusatzangaben zur Erwerbstätigkeit (z. B. EF175 Stellung im Betrieb)*

EF456-EF476 *Krankenversicherung*

³ Nach Auffassung der Rechtsabteilung des Statistischen Bundesamtes dient Frage 18 zum sogenannten Hauptstatus ausschließlich der Abgrenzung der Erwerbstätigen von den Nichterwerbstätigen im Sinne der ILO-Definition. Eine solche Frage führe die Befragten nur durch den Fragebogen, um so die Antwortqualität zu verbessern. Da die in EF66 erfassten Antworten nicht zu den gesetzlichen Erhebungsmerkmalen zählten, kann EF66 weder ausgewertet noch veröffentlicht werden.

Zusätzliche Hinweise zur Erfassung der Krankenversicherung und zu ihrer Vergleichbarkeit mit 2007 liefern die Anmerkungen zum Themenkomplex „Krankenversicherung“ (Statistisches Bundesamt 2014a).

- Das Thema des Ad-hoc-Moduls 2011 ist „Beschäftigung behinderter Personen“. Die dazu gehörigen Variablen sind EF1084-EF1096. Der Auswahlsatz beträgt 0,1 %, die Fragen sind freiwillig zu beantworten. Fragen zu diesem Thema wurden bereits im Ad-hoc-Modul 2002 gestellt (siehe die Variablen EF420-EF425 im SUF 2002). Dabei ist auf die Besonderheit hinzuweisen, dass die Variablen EF1090u1-EF1096 auch Angaben zu hypothetischen Sachverhalten enthalten. Während sich die Fragen 161-167 zum gesundheitlichen Zustand und zur Erwerbstätigkeit an 15- bis 64-jährige Erwerbstätige richten, d. h., Personen mit bezahlter Tätigkeit und unbezahlt mithelfende Familienangehörige, die wenigstens eine der Leitfragen zur Erfassung des ILO-Erwerbsstatus (F18a-F21a; EF77-EF79, EF81) mit „ja“ beantwortet haben, richten sich die Fragen 168-174 zum gesundheitlichen Zustand und möglicher Erwerbstätigkeit an Nichterwerbstätige, d. h. an sonstige Personen ohne bezahlte Tätigkeit. Beispielsweise enthalten die Variablen EF1090u1-EF1090u5 sowohl (Mehrfach-) Antworten von Erwerbstätigen zur Frage 161 „Ist der Umfang Ihrer wöchentlichen Arbeitsstunden eingeschränkt?“ als auch (Mehrfach-) Antworten von Nichterwerbstätigen auf die Frage 168 „Wäre bei einer Erwerbstätigkeit der Umfang Ihrer wöchentlichen Arbeitsstunden eingeschränkt?“
- Bei der Konstruktion des für die Merkmale des Ad-hoc-Moduls verwendeten Hochrechnungsfaktors EF956 wurden 2011 für die Schätzung der Antwortwahrscheinlichkeiten (1. Stufe) die Erhebungsmethode (EF7), der Erwerbstyp (EF29) differenziert nach Geschlecht (EF46) und Altersklassen (EF44) auf Bundesebene, der Erwerbstyp (EF29) differenziert nach Geschlecht (EF46) auf Landesebene (EF1) sowie die zwei Nationalitätengruppen Deutsche/Ausländer in den Gebieten Ost/West eingesetzt. Zielpopulation sind die 15- bis 64-Jährigen.

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.
- Die Fragen des computergestützten Interviews (CAPI) unterscheiden sich teilweise von den Fragen im Selbstausfüllerbogen. Die Dokumentation der Abweichungen zwischen Selbstausfüller und Blaiseprogramm (Statistische Ämter des Bundes und der Länder 2010) macht deutlich, welche Fragen betroffen sind und worin die Unterschiede bestehen.
- 2011 wurden die Antwortkategorien zum Schul- und Hochschulbesuch (EF289) und zum beruflichen Ausbildungsabschluss (EF312 usw.) erweitert, sodass ein direkter Vergleich mit früheren Zeitpunkten nur eingeschränkt möglich ist.
- In Bezug auf Fragestellungen zu Migration und Integration ist auf eine konzeptionelle Änderung hinzuweisen. Während bis 2010 Befragte, die nach einem ersten Zuzug die Bundesrepublik Deutschland für mehr als sechs Monate verlassen hatten, das Jahr des letzten Zuzugs angegeben haben, wird nun das Jahr des ersten Zuzugs (EF367) erfragt, auch wenn der Aufenthalt in Deutschland zwischenzeitlich unterbrochen wurde. Zusätzlich wird mit einer neuen Frage (F129; EF384) das Jahr der Rückkehr nach Deutschland nach der letzten Unterbrechung erfasst. Dies wirkt sich auch auf die darauf aufbauenden Typisierungen zum Migrationsstatus aus.
- Ab dem Mikrozensus 2011 wird von den statistischen Ländern eine über alle vier Berichtsjahre feste Personennummer vergeben. Die Variable EF63 ist im SUF 2011 enthalten. Dies verbessert künftig die Möglichkeiten der Zusammenführung der Rotationsviertel zu Panels.
- In der EU-Arbeitskräfteerhebung wurde ab 2011 die International Standard Classification of Occupations, Ausgabe 2008 (ISCO-08), eingesetzt. Entsprechend wurde diese neue Fassung im Mikro-

zensus 2011 umgesetzt (siehe EF517, EF521, EF739, EF827, EF851). Die Berufe nach ISCO-08 wurden auf Basis eines Umsteigeschlüssels (Statistisches Bundesamt 2014e) der Klassifikation der Berufe 1992 codiert. Die ISCO-88 COM wurde im Originalfile nicht mehr umgesetzt. Für das SUF wurde ISCO-88 COM nochmals generiert. Dies betrifft auch die abgeleiteten Variablen für den Haupteinkommensbezieher (EF739a), die Bezugsperson der Lebensform (EF827a) und den Lebenspartner der Bezugsperson der Lebensform (EF851a).

- Einkommensangaben werden im Mikrozensus nicht genau („spitz“), sondern gruppiert erfasst. Zusätzlich zu den herkömmlichen Variablen zum Nettoeinkommen (EF436, EF707, EF806) werden im SUF 2011 erstmals Variablen bereitgestellt (EF438, EF708, EF807), die Schätzungen der „spitzen“ Einkommen enthalten (siehe für detaillierte Hinweise Statistisches Bundesamt 2014b). Diese können für Einkommens- oder Armutsanalysen herangezogen werden.

4 Aufbereitung der Rohdaten zum Mikrozensus Scientific Use File 2011 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 Wertelabels⁴ sowie die thematische Übersicht nach der Systematik des Mikrodaten-Informationssystems ([MISSY](#)) erstellt (GESIS 2014d). 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.⁵

Als Referenz zur Plausibilitätsprüfung der Mikrozensus SUF dienen die in den Fachserien des Statistischen Bundesamtes veröffentlichten Ergebnisse des jeweiligen Mikrozensus. Für den Mikrozensus 2011 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 13, Reihe 1.1. 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 2011](#) sowie den [Definitionskatalog](#) im [GESIS-Web](#) (2014a, 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 2014d). Die meisten Variablen des SUF weichen nur in sehr gering-

⁴ 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.

⁵ Siehe dazu die Hinweise im [Datenhandbuch](#) (Statistisches Bundesamt / GESIS 2014a: 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 einer der AKE-Strukturvariablen zur Leiharbeit und der Variable Dauerhafte gesundheitl. Probleme des Ad-hoc-Moduls berichtet. Zum 2011 durchgeführten Zusatzprogramm skizzieren Verteilungen ob und wie eine Person krankenversichert ist und zur Stellung im Betrieb 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 2011 (Jahresdurchschnitt in 1.000)

Bevölkerung am Haupt- und Nebenwohnsitz^a (EF30>=1 & EF30<=3)	
82.655	
Bevölkerung am Hauptwohnsitz^b (EF30>=1 & EF30<=2)	Bev. am Neben- wohnsitz ^c (EF30=3) 891
81.764	
Bevölkerung in Privathaushalten^d (EF31=1)	
81.717	
Bevölkerung in Lebensformen am Hauptwohnsitz^f [Bezugsperson (BP) am Hauptwohnsitz] (EF761>=1 & EF761<=2)	- [BP am Neben- wohns.] (EF761>2) 764
80.953	
	Bev. in Ge- mein- sch.un- terk. ^e (EF31=2) 938

- 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 2011 (Jahresdurchschnitt in 1.000)

Merkmal	Hochrechnung (in 1.000)
Bevölkerungsgruppen	
Bevölkerung am Haupt- und Nebenwohnsitz (Fallzahl insgesamt: n = 486.100)	82.655
Bevölkerung in Privathaushalten	81.717
Bevölkerung am Hauptwohnsitz	81.764
Bevölkerung in Lebensformen am Hauptwohnsitz	80.953
Bevölkerung am Hauptwohnsitz, Erwerbstätige	39.850
EF117: Stellung im Beruf x EF29: Erwerbstyp	
Bevölkerung am Ort der Hauptwohnung (siehe Fachserie 1, Reihe 4.1.1, Tabelle 1.1)	
Selbstständige	4.413
Mithelf. Familienangeh.	241
Beamte (inkl. Soldaten)	2.091
Angestellte (inkl. kfm./ techn. Auszubildende, Zivildienstleistende, geringfügig beschäftigte Schüler, Studenten, Rentner)	22.660
Arbeiter (inkl. gew. Auszubildende)	10.446
Erwerbstätige insgesamt	39.850
Erwerbslose	2.508
Erwerbspersonen	42.358
Nichterwerbspersonen	39.406
Arbeitssuchende Nichterwerbspersonen	442
Sonstige Nichterwerbspersonen	38.964
Insgesamt	81.764
EF663: Personen im Haushalt: Anzahl (Haushaltstypisierung)	
Bevölkerung in Privathaushalten, Haushaltsbezugsperson [Haupteinkommensbezieher] (siehe Fachserie 1, Reihe 3, Tabelle 1.1)	
1 Person	16.351
2 Personen	13.881
3 Personen	5.075
4 Personen	3.790
5 Personen und mehr	1.353
Haushalte insgesamt	40.450

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	65.804
Personen mit Migrationshintergrund im engeren Sinn	15.960
Bevölkerung mit Migrationshintergrund und eigener Migrationserfahrung	10.681
Ausländer	5.685
Deutsche	4.997
(Spät-) Aussiedler	3.215
durch Einbürgerung	1.782
Personen ohne eigene Migrationserfahrung	5.279
Ausländer	1.527
Deutsche	3.751
durch Einbürgerung	425
mit mind. einem zugewanderten oder als Ausländer in Deutschland geborenen Elternteil	3.326
mit beidseitigem Migrationshintergrund	1.725
mit einseitigem Migrationshintergrund	1.601
Insgesamt	81.764
EF177: Arbeitsvertrag: Zeitarbeitsvermittlungsfirma (FB) (0,1 %)	
Unterstichprobe, erwerbstätige Bevölkerung nach EU-Definition GewichtungsvARIABLE: EF955 Hochrechnungsfaktor AKE-Strukturvariablen: Jahr (in 1000)	
Ja	872
Nein	29.777
Keine Angabe	4.556
Insgesamt	35.204
EF1084: Dauerhafte gesundheitl. Probleme (FB) (0,1%)	
Unterstichprobe, Bevölkerung nach EU-Definition GewichtungsvARIABLE: EF956 Hochrechnungsfaktor Ad-hoc-Substichprobe: Jahr (in 1000)	
Ja	9.387
Nein	35.503
Keine Angabe	8.929
Insgesamt	53.819

Merkmal	Hochrechnung (in 1.000)
EF456: Krankenversicherung	
Bevölkerung am Hauptwohnsitz ohne Jahresüberhang mit Fragebogen MZ 2010	
Gewichtungsvariable: EF952 Standardhochrechnungsfaktor Jahr (in 1000)	
(siehe Fachserie 13, Reihe 1.1, Tabelle 1.3)	
In einer gesetzlichen Krankenversicherung selbst pflichtversichert	49.524
In einer gesetzlichen Krankenversicherung selbst freiwillig versichert	3.330
In einer gesetzlichen Krankenversicherung als Familienangehöriger versichert	17.724
In einer privaten Krankenversicherung selbst versichert	6.929
In einer privaten Krankenversicherung als Familienangehöriger versichert	2.468
Nein, nicht krankenversichert ⁶	350
Keine Angabe	257
Insgesamt	80.583
EF175: Stellung im Betrieb	
Bevölkerung am Hauptwohnsitz	
Gewichtungsvariable: EF952 Standardhochrechnungsfaktor Jahr (in 1000)	
(siehe Fachserie 1, Reihe 4.1.2, Tabelle 3.11)	
Selbstständige	4.413
Mithelfende Familienangehörige	241
Abhängig Beschäftigte	35.197
Einfache Angestellte und Beamte, angelernte Arbeiter	6.809
Mittlere Angestellte und Beamte, Facharbeiter	11.003
Gehobene Angestellte und Beamte, Meister	9.106
Höhere Angestellte und Beamte	5.919
Auszubildende in anerkannten Ausbildungsberufen	1.617
Differenz zu „Abhängig Beschäftigte“	743
Grundwehr-, Zivildienstleistende	43
Zeit-/Berufssoldaten	192
Keine Angabe	507
Erwerbstätige insgesamt	39.850

⁶ Zu beachten ist, dass für die Abgrenzung nicht krankenversicherter Personen neben der Angabe in EF456 (8 Nein, nicht krankenversichert) außerdem Variable EF460 (8 Kein sonstiger Anspruch auf Krankenversicherung) verwendet werden sollte: EF456=8 & EF460=8 (siehe Statistisches Bundesamt (2012e), Fachserie 13, Reihe 1.1, S. 22).

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, 2012: Labour force survey in the EU, Candidate and EFTA countries - Main characteristics of the national surveys - 2011. Luxembourg: Publications Office of the European Union. URL: http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-12-025/EN/KS-RA-12-025-EN.PDF
- Eurostat, 2013: Quality report of the European Union Labour Force Survey 2011. Luxembourg: Publications Office of the European Union. URL: http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-13-008/EN/KS-RA-13-008-EN.PDF
- Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder [FDZ], 2014: 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, 2014a: Abgrenzung und Hochrechnung der Bevölkerungsbegriffe im Mikrozensus 2011. Mannheim. URL: <http://www.gesis.org/missy/studie/erhebung/studienbeschreibung/hochrechnung/zur-abgrenzung-der-bevoelkerungsbegriffe-und-zur-hochrechnung-im-mikrozensus-suf-2011/>
- GESIS, 2014b: Definitionskatalog. Mannheim. URL: <http://www.gesis.org/missy/studie/klassifikationen/konzepte-und-definitionen/definitionskatalog/>
- GESIS, 2014c: Informationen zum Mikrozensus 2011. Mannheim. URL: <http://www.gesis.org/missy/missy-home/auswahl-datensatz/mikrozensus-2011/>
- GESIS, 2014d: MISSY - Mikrodaten-Informationssystem. Mannheim. URL: <http://www.gesis.org/missy/>
- Herter-Eschweiler, Robert, 2014: Gegenüberstellung der Fragebogenversionen der Mikrozensus-erhebung 2011. URL: http://www.gesis.org/missy/fileadmin/missy/erhebung/fragebogen/fragebogenvergleich_mz2011.pdf
- 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/unser-angebot/daten-analysieren/amtliche-mikrodaten/mikrozensus/mikrodaten-tools/> bereitgestellten Programme).
- Statistische Ämter des Bundes und der Länder, 2010: Dokumentation der Abweichungen zwischen Selbstausfüller und Blaise-Programm. Mikrozensus 2011. Hilfen zur Erfassung. URL: http://www.gesis.org/missy/fileadmin/missy/erhebung/fragebogen/Selbstausfueller_Blaise.pdf
- Statistisches Bundesamt, 2012a: Fachserie 1 Bevölkerung und Erwerbstätigkeit, Reihe 2.2 Bevölkerung mit Migrationshintergrund. Ergebnisse des Mikrozensus 2011. Wiesbaden. URL: https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/MigrationIntegration/Migrationshintergrund2010220107004.pdf?__blob=publicationFile
- Statistisches Bundesamt, 2012b: Fachserie 1 Bevölkerung und Erwerbstätigkeit, Reihe 3 Haushalte und Familien. Ergebnisse des Mikrozensus 2011. Wiesbaden. URL:

https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/HaushalteMikrozensus/HaushalteFamilien2010300117004.pdf?__blob=publicationFile

Statistisches Bundesamt, 2012c: Fachserie 1 Bevölkerung und Erwerbstätigkeit, Reihe 4.1.1 Stand und Entwicklung der Erwerbstätigkeit. Mikrozensus 2011. URL:

https://www.destatis.de/DE/Publikationen/Thematisch/Arbeitsmarkt/Erwerbstaetige/StandEntwicklungErwerbstaetigkeit2010411117004.pdf?__blob=publicationFile

Statistisches Bundesamt, 2012d: Fachserie 1 Bevölkerung und Erwerbstätigkeit, Reihe 4.1.2 Beruf, Ausbildung und Arbeitsbedingungen der Erwerbstätigen. Mikrozensus 2011. Wiesbaden. URL:

https://www.destatis.de/DE/Publikationen/Thematisch/Arbeitsmarkt/Erwerbstaetige/BerufArbeitsbedingungErwerbstaetigen2010412117004.pdf?__blob=publicationFile

Statistisches Bundesamt, 2012e: Fachserie 13 Sozialleistungen, Reihe 1.1 Angaben zur Krankenversicherung (Ergebnisse des Mikrozensus 2011). Wiesbaden. URL:

https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/HaushalteMikrozensus/KrankenversicherungMikrozensus2130110119004.pdf?__blob=publicationFile

Statistisches Bundesamt, 2012f: Methodeninformation. Mikrozensus und Arbeitskräfteerhebung: Ergebnisse zur Erwerbstätigkeit ab dem Jahr 2011. Wiesbaden. URL:

https://www.destatis.de/DE/Methoden/Methodenpapiere/Download/MethodenArbeitskraefterhebung.pdf?__blob=publicationFile

Statistisches Bundesamt, 2012g: Qualitätsbericht Mikrozensus 2011. Wiesbaden. URL:

https://www.destatis.de/DE/Publikationen/Qualitaetsberichte/Bevoelkerung/Mikrozensus2011.pdf?__blob=publicationFile

Statistisches Bundesamt, 2014a: Anmerkungen zur Erfassung des Themenkomplexes „Krankenversicherung“ in der Mikrozensus-Erhebung 2011. Bonn. URL:

<http://www.gesis.org/missy/fileadmin/missy/daten/methodenpapiere/Krankenversicherung.pdf>

Statistisches Bundesamt, 2014b: Schätzung des individuellen „spitzen“ Nettoeinkommens im Mikrozensus und die daraus abgeleiteten Variablen EF707 „Haushaltseinkommen (im letzten Monat)“ und EF806 „Nettoeinkommen (im letzten Monat): Lebensform“. Bonn. URL:

<http://www.gesis.org/missy/fileadmin/missy/daten/methodenpapiere/Nettoeinkommen.pdf>

Statistisches Bundesamt, 2014c: Statistik von A bis Z: Mikrozensus. URL:

<https://www.destatis.de/DE/Meta/AbisZ/Mikrozensus.html>

Statistisches Bundesamt, 2014d: Überprüfung der Randverteilung. Bonn. URL:

http://www.gesis.org/missy/fileadmin/missy/daten/Randverteilung/Randverteilung_2011.pdf

Statistisches Bundesamt, 2014e: Umsteigeschlüssel: KldB 1992 auf ISCO-08. Bonn. URL:

http://www.gesis.org/missy/fileadmin/missy/klassifikationen/Amtliche_Klassifikationen/ISCO/kldb_92_isco08.pdf

Statistisches Bundesamt, GESIS (Hrsg.), 2014a: Datenhandbuch zum Mikrozensus Scientific Use File 2011. URL:

http://www.gesis.org/missy/fileadmin/missy/erhebung/datenhandbuch/DHB_2011.pdf

Statistisches Bundesamt, GESIS (Hrsg.), 2014b: Thematische Übersicht zum Mikrozensus Scientific Use File 2011. URL:

http://www.gesis.org/missy/fileadmin/missy/daten/thematische_uebersicht/Thematische_Uebersicht_MZ2011.pdf

Statistisches Bundesamt, GESIS (Hrsg.), 2014c: Wichtige Informationen zur Nutzung des Mikrozensus Scientific Use Files 2011. [Readme zu den Setups] URL:

http://www.gesis.org/missy/fileadmin/missy/daten/setups/mz2011/spss_setup_11.zip

Letzter Zugriff bei allen Web-Dokumenten: 14. 05. 2014.

Anhang

```

* SPSS-Syntax zu Tabellen 1 und 2.

* IBM SPSS Statistics 22, Version 22.0.0.0

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

get file='<mz2011.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 'Beamte (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%): EF177, EF178, EF179u1-EF179u5, EF1097u1-
EF1097u6.
* 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.
* Beispiel EF177: Arbeitsvertrag: Zeitarbeitsvermittlungsfirma (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 AKE-Strukturvariablen.
weight by EF955.
fre EF177.
```

```
* AD-HOC-MODUL BESCHÄFTIGUNG BEHINDERTER PERSONEN .
* Filter: EF6 Unterstichprobe, Fragenprogramm .
* Gewichtung: EF956 Hochrechnungsfaktor Ad-hoc-Substichprobe: Jahr (in
1000).
* EF1084 F158 Dauerhafte gesundheitl. Probleme (FB) (0,1%)
* [1] Dauerhafte gesundheitl. Probleme 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).
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) .
* EF1084 = 1 Dauerhafte gesundheitl. Probleme genannt .
weight by EF956.
fre EF1084.
```

```
* MZ-ZUSATZPROGRAMM KRANKENVERSICHERUNG:
* Filter: EF30>=1 & EF30<=2 Bevölkerung am Hauptwohnsitz .
* EF456<>-5 Ohne Jahresüberhang mit Fragebogen MZ
* 2010 .
* Gewichtung: EF952 Standardhochrechnungsfaktor: Jahr (in 1000) .
temporary.
* Alle Missing Value Deklarationen aufheben.
missing values all().
* Beispiel: EF456 Krankenversicherung .
* Siehe Fachserie 13, Reihe 1.1, Tabelle 1.3 .
select if (EF30>=1 & EF30<=2 /* Hauptwohnsitz */
& EF456<>-5) /* Ohne Jahresüberhang mit Fragebogen MZ 2010*/ .
* Hochrechnung: Jahresdurchschnitt EF952 (in 1000) .
weight by EF952.
fre EF456.
```

```
* MZ-ZUSATZPROGRAMM: ERWERBSTÄTIGKEIT:
* Filter: EF29=1 Erwerbstätige .
* EF30>=1 & EF30<=2 Bevölkerung am Hauptwohnsitz .
* Gewichtung: EF952 Standardhochrechnungsfaktor: Jahr (in 1000) .
temporary.
* Alle Missing Value Deklarationen aufheben.
missing values all().
* Beispiel: EF175 Stellung im Betrieb .
* Siehe Fachserie 1, Reihe 4.1.2, Tabelle 3.11 .
select if (EF29=1 & EF30>=1 & EF30>=2) /* Erwerbstätige & Hauptwohnsitz */ .
recode EF175 (1 2=1) (19=2) (3 7 13=3) (4 8 14=4) (5 9 10 15 16=5) (6 17
18=6) (11 12=7) (21=8) (20=9) (99=99)(else=sysmis) into v175.
formats v175 (f1.0).
var lab v175 'Stellung im Betrieb (kategorisiert)'.
val lab v175
1 'Selbstständige'
2 'Mithelfende Familienangehörige'
3 'Einfache Angestellte und Beamte, angelernte Arbeiter'
4 'Mittlere Angestellte und Beamte, Facharbeiter'
```

```
5 'Gehobene Angestellte und Beamte, Meister'  
6 'Höhere Angestellte und Beamte'  
7 'Auszubildende in anerkannten Ausbildungsberufen'  
8 'Grundwehr-/Zivildienstleistende'  
9 'Zeit-/Berufssoldaten'  
99 'Keine Angabe'.  
weight by EF952 /* Hochrechnung: Jahresdurchschnitt EF952 */ .  
fre v175.
```

```
* Abhängig Beschäftigte.  
recode EF175 (3 thru 18=1) (20 thru 99=1) (else=0) into abh.  
formats abh (f1.0).  
missing values abh (0).  
var lab abh 'Abhängig Beschäftigte'.
```

```
temporary.  
select if (EF29=1 & EF30>=1 & EF30<=2) /* Erwerbstätige & Hauptwohnsitz */ .  
weight by ef952.  
fre abh.
```