

Wichtige Informationen zur Nutzung des Mikrozensus Scientific Use File 2007

Diese Informationen sollen dabei helfen, das Scientific Use File sachgerecht auszuwerten. Ab dem Erhebungszeitpunkt 2007 werden zusätzlich zu den Rohdaten im ASCII-Format auch Systemfiles der Statistikprogramme PASW (früher SPSS), SAS und Stata im dokumentierten Format bereitgestellt.

Die Aufbereitung und Dokumentation des faktisch anonymisierten Einzelmaterials erfolgt durch das Statistische Bundesamt, Gruppe VIII C – Mikrozensus, in Kooperation mit dem German Microdata Lab (GML), dem Servicezentrum für amtliche Mikrodaten der GESIS. Ausführliche Informationen über den Mikrozensus (MZ) und die Mikrozensus Scientific Use Files (SUF), u. a. Fragebogen, Klassifikationen und Definitionen, stehen auf den [Web-Seiten von GESIS](#) zur Verfügung und werden im Folgenden zusammengefasst dargestellt. Hinweise zur Nutzung der Setups sowie Verteilungen zur Prüfung der Systemfiles bzw. zum korrekten Ablauf der Setups finden sich im Anhang.

- 1 [Allgemeine Informationen zum Mikrozensus](#)
- 2 [Die wichtigsten Veränderungen beim Mikrozensus 2007 in Stichworten](#)
- 3 [Kurzbeschreibung des Scientific Use Files:
Substichprobenziehung, Bevölkerungsbegriffe und Hochrechnung](#)
- 4 [Literaturhinweise](#)
[Anhang](#)

Weitere Informationen sind auf den Seiten des Statistischen Bundesamtes und der Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder abrufbar:

[Statistik von A bis Z | Was beschreibt der Mikrozensus?](#)
[Forschungsdatenzentren | Datenangebot | Mikrozensus](#)

Statistisches Bundesamt
Gruppe VIII C Mikrozensus
Graurheindorfer Str. 198
53117 Bonn
Tel.: 0228 99 / 643-8725
Fax: 0228 99 / 643-8990
<http://www.destatis.de>
Kontakt: Robert Herter-Eschweiler
E-Mail: mikrozensus@destatis.de

GESIS - Leibniz-Institut für Sozialwissenschaften
German Microdata Lab
B2, 1
68159 Mannheim
Tel.: 0621-12 46-265
Fax: 0621-12 46-100
<http://www.gesis.org/das-institut/gml/>
Kontakt: Bernhard Schimpl-Neimanns
E-Mail: gml@gesis.org

Inhalt

- Was ist der Mikrozensus?
- Das Wichtigste über den Mikrozensus 2007 im Überblick
- Datenerhebung
- Stichprobe
- Erhebungsprogramm
- Die wichtigsten Veränderungen beim Mikrozensus 2007 in Stichworten
- Kurzbeschreibung des Scientific Use Files
- Zur Abgrenzung der Bevölkerungsbegriffe und zu Gewichtung und Hochrechnung
- Zugang zum Mikrozensus Scientific Use File und Beratung
- Literaturhinweise
- Veröffentlichungen auf der Datengrundlage von Scientific Use Files
- Ergebnisse des Mikrozensus 2007
- Mikrozensusgesetz 2005
- EU-Verordnungen zur Arbeitskräfteerhebung
- Informationen zum Mikrozensus allgemein und seiner Konzeption
- Informationen zur Weitergabe faktisch anonymisierter Einzeldaten
- Informationen zur Fehlerrechnung im Mikrozensus
- Anhang: Informationen zur Nutzung der Setups; Verteilungen zur Prüfung der Systemfiles bzw. des Ablaufs der Setups

1 Allgemeine Informationen zum Mikrozensus

Was ist der Mikrozensus?

Der Mikrozensus ist die amtliche Repräsentativitätsstatistik über die Bevölkerung und den Arbeitsmarkt in Deutschland mit einem Auswahlatz von einem Prozent der Haushalte und der Bevölkerung. Sie wird seit 1957 in Westdeutschland und seit 1991 auch in den neuen Bundesländern durchgeführt. Die Arbeitskräftestichprobe der Europäischen Union (EU Labour Force Survey) ist in den Mikrozensus integriert.

Siehe auch:

Information für die Befragten [[pdf ca. 210kB](#)]

(Statistische Ämter des Bundes und der Länder, 2006)

Das Wichtigste über den Mikrozensus 2007 im Überblick

- Grundgesamtheit:** 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). Das Scientific Use File enthält Angaben zum Berichtsquartal.
- Stichprobe:** Flächenstichprobe (Klumpenstichprobe)
- Auswahleinheiten:** Auswahlbezirke (künstlich abgegrenzte Flächen bestehend aus zusammen liegenden Gebäuden oder Gebäudeteilen). Ein Auswahlbezirk verbleibt vier Jahre in der Stichprobe, in jedem Jahr wird 1/4 der Auswahlbezirke durch neue Auswahlbezirke ersetzt (rotierende Panelstichprobe).
Aufgrund des Prinzips der Flächenstichprobe werden die aus dem Auswahlbezirk wegziehenden Haushalte und Personen nicht weiter befragt, sondern durch zuziehende Haushalte und Personen ersetzt. Der Mikrozensus ist somit eine Wiederholungsbefragung mit teilweiser Überlappung der Erhebungseinheiten (partielle Rotation). Die Möglichkeit der Zusammenführung der Querschnittsdaten zu einem [Mikrozensuspanel](#) ist rechtlich seit dem Mikrozensusgesetz 1996 gegeben.
- Auswahlgrundlage:** Grundauswahl:
Alte Bundesländer: Ergebnisse der Volkszählung 1987.
Neue Bundesländer/Ost Berlin: Ergebnisse aus dem Bevölkerungsregister Statistik der DDR 1990.
Aktualisierung:
Die auf dieser Basis gebildete Grundauswahl wird jährlich durch Berücksichtigung der Neubautätigkeit über die Meldungen zur Bautätigkeitsstatistik aktualisiert.
- Auswahlverfahren:** Einstufig geschichtete Stichprobe
- Schichtungsmerkmale:** Bundesland, Regionalschicht, Gebäudeschicht
- Auswahlsatz:** 1 Prozent;
Ad-hoc Modul EU Labour Force Survey 0,1% (2007 „Arbeitsunfälle, Arbeit und Gesundheit“)
- Auswahltechnik:** (Grundauswahl)
Vor der Auswahl wurden die Auswahlbezirke regional angeordnet (je Schicht in der Sortierung nach regionaler Schichtungsgruppe, Kreis, Gemeindegrößenklasse, Gemeinde und Auswahlbezirksnummer).
- Zerlegung der Auswahlgesamtheit in 100 1%-Stichproben: Je

100 aufeinanderfolgende Auswahlbezirke („Zonen“) werden mit einer zufälligen Permutation der Zahlen 0 bis 99 versehen

- Zerlegung der Auswahlgesamtheit in vier Rotationsviertel zu je 0,25%: Je vier aufeinanderfolgende Zonen werden mit einer zufälligen Permutation der Zahlen 1 bis 4 versehen.
- Zerlegung der Auswahlgesamtheit in 48 „Wochenteile“
- Zusammenfassung der 48 Wochenteile zu einem Monatsvolumen (gleichmäßige Verteilung der Stichprobe über das Jahr)
- Aufteilung der 1%-Stichprobe nach Zufallsverfahren in zwölf gleiche Unterstichproben (gleichmäßige Verteilung der Stichproben über das Jahr)

Bei der Neubausauswahl, die in einer eigenen Gebäudeschicht abgegrenzt ist, gibt es gegenüber der Grundausswahl im Stichprobenplan einige Modifikationen. Während ein Auswahlbezirk der Grundausswahl (ohne Gemeinschaftsunterkünfte) durchschnittlich neun Wohnungen umfasst, ist die Richtgröße bei der Neubausauswahl sechs Wohnungen. Für Auswahlbezirke der Gebäudeschicht mit Gemeinschaftsunterkünften beträgt die Richtgröße 15 Personen.

| | |
|---|--|
| Erhebungseinheiten: | Personen (in Privathaushalten und Gemeinschaftsunterkünften), Haushalte und Wohnungen |
| Analyseeinheiten: | Personen, Lebensgemeinschaften, Familien, Haushalte, Wohnungen |
| Zuordnung Person/Haushalt: | Personen mit mehreren Haushalten werden dem Haushalt am Ort ihrer Hauptwohnung zugerechnet, gleichzeitig können Personen am Ort ihrer Nebenwohnung erfasst werden, sodass Mehrfachzählungen möglich sind. |
| Erhebungsverfahren: | I. d. R. mündliche Befragung (face-to-face), die dezentral von den Statistischen Landesämtern mit Hilfe von Interviewern – ausgestattet mit Laptops (CAPI) – durchgeführt wird, zum Teil auch schriftlich (Selbstaussfüllerbogen). Zulässig sind außerdem Proxy-Interviews, d. h., ein erwachsenes Haushaltsmitglied darf stellvertretend für andere Haushaltsmitglieder antworten (Fremdauskünfte für Personen ab 15 Jahren: ca. 27 %). |
| Scientific Use File (SUF): | Das SUF ist eine faktisch-anonymisierte 70%-Substichprobe der Haushalte bzw. Wohnungen des Mikrozensus, wobei alle Personen eines ausgewählten Haushalts bzw. einer Wohnung enthalten sind. In den Erhebungsjahren, in denen das Zusatzprogramm zur Wohnsituation (1998, 2002, 2006) durchgeführt wird, wird eine Wohnungssubstichprobe gezogen. |
| Hochrechnung / GewichtungsvARIABLEN im SUF: | Die Hochrechnung des Mikrozensus basiert auf einem zweistufigen Verfahren: (1) Zunächst erfolgt ein Ausgleich der nicht erreichbaren Haushalte durch Kompensationsfaktoren anhand von Informationen |

über diese Haushalte. Der Unit-Nonresponse beträgt im Mikrozensus seit 2005 ca. fünf Prozent der Haushalte. (Bei einzelnen Fragen bzw. Merkmalen liegt die Ausfallquote (Item-Nonresponse) i. d. R. unter zehn Prozent.)

(2) Anschließend werden die mit dem Kompensationsfaktor für Unit-Nonresponse gewichteten Stichprobenwerte an Eckwerte aus der laufenden Bevölkerungsfortschreibung und dem Ausländerzentralregister nach Altersklassen, Staatsangehörigkeit bzw. Staatsangehörigkeitsgruppen und Geschlecht angepasst. Für Zeit- und Berufssoldaten einschließlich Bundes- und Bereitschaftspolizei sowie Grundwehrdienstleistende stammen die Eckwerte aus Bestandsdaten des Bundesministeriums der Verteidigung, des Bundesministeriums des Innern bzw. der Bundespolizei.

Die Anpassung erfolgt quartalsweise auf unterschiedlichen regionalen Ebenen.

Mit Ausnahme des Ad-hoc-Moduls liegen für alle Merkmalsbereiche gleiche Hochrechnungsfaktoren vor.

Das SUF enthält die bereits mit dem Kehrwert der Auswahlwahrscheinlichkeit der Substichprobe (100/70) multiplizierten Hochrechnungsvariablen des Mikrozensus-Originalfiles für Quartalsauswertungen und Jahresauswertungen des Grundprogramms (EF951 bzw. EF952) sowie Jahresauswertungen des Ad-hoc-Moduls (EF955). Die Hochrechnungsfaktoren rechnen auf 1.000 der Population hoch.

Regionalinformationen: Das SUF enthält an Regionalinformationen Angaben zum Bundesland und zur (vergrößerten) Gemeindegrößenklasse (Stand 31. 12. 2006).

Bei Auswertungen nach Landesteilen wird von den Statistischen Ämtern i. d. R. nicht mehr zwischen dem früheren Bundesgebiet inkl. West-Berlin und den neuen Bundesländern inkl. Ost-Berlin unterschieden, sondern nach früherem Bundesgebiet ohne Berlin vs. neuen Ländern einschließlich Berlin. Dies hängt auch damit zusammen, dass nach einer Gebietsreform die ehemalige Ost-West-Aufteilung Berlins nicht mehr trennscharf abzubilden ist (siehe <http://www.statistik-berlin.de/statistiken/rbs/gebietsreform-internet.pdf>).

Die Auswahlbezirke des Mikrozensus können nach wie vor den früheren Gemeindeteilen West- und Ost-Berlin korrekt zugeordnet werden. Für Ost-West-Vergleiche ist dies mit dem SUF wie folgt umsetzbar (PASW- bzw. SPSS-Syntax):

* Beachte im SUF MZ 2007: Die Länder sind in EF1 nach der Sortierung des amtlichen Gemeindegchlüssels codiert.

```
DO IF (EF563 <> 9).
```

```
    RECODE EF1 (1 thru 11 = 1) (12 thru 16 = 2) into WESTOST.
```

```
ELSE IF (EF563 = 9).
```

```
    RECODE EF1 (11=2) into WESTOST.
```

```
END IF.
```

```
VALUE LABEL WESTOST
```

1 "Früheres Bundesgebiet inkl. West-Berlin"

2 "Neue Bundesländer inkl. Ost-Berlin".

Für die Anpassung der Mikrozensusergebnisse an die laufende Bevölkerungsfortschreibung liegen allerdings nur Daten zu den neuen Gebietsabgrenzungen (12 statt 23 Bezirke) vor. Deshalb ist bei der sogenannten gebundenen Hochrechnung (Gewichtungsvariablen EF951 ff.) mit regionalen Unschärfen zu rechnen, die aber nicht gravierend sein dürften.

Datenerhebung

Der Mikrozensus ist eine Haushaltsstichprobe, d. h. alle Personen in einem Haushalt werden entweder direkt befragt oder eine erwachsene Person gibt Auskunft über die weiteren (alle minderjährigen) Haushaltsmitglieder (Proxy-Interviews). Einige Fragen sind freiwillig zu beantworten. Der Mikrozensus umfasst auch ausländische Haushalte. Zur Erhebung ihrer Angaben wird der gleiche Erhebungsbogen verwendet wie bei deutschen Haushalten. Die Interviewer leiten die Daten an die Statistischen Landesämter weiter. Haushalte, die sich dafür entschieden haben, den Fragebogen selbst auszufüllen, senden diesen an das jeweilige Statistische Landesamt. Von den Interviewern nicht erreichbare Haushalte werden von den Statistischen Landesämtern angeschrieben und erhalten einen Fragebogen für die schriftliche Befragung. Die Daten werden von den Statistischen Landesämtern aufbereitet und beim Statistischen Bundesamt zu einem Datensatz zusammengespielt.

Berichtszeitraum

Der Berichtszeitraum ist ab 2005 die jeweils letzte Kalenderwoche vor der Befragung (siehe auch [Berichtswochen der Mikrozensen 1989-2004](#)).

Stichprobe

Das stichprobenmethodische Grundkonzept des Mikrozensus ist die einstufige Klumpenstichprobe. Die Auswahlseinheiten (die sogenannten Auswahlbezirke) sind künstlich abgegrenzte Flächen, die in der Regel mehrere räumlich beieinanderliegende Gebäude (bei großen Gebäuden nur ein Gebäude oder ein Gebäudeteil) umfassen. Bis zum Mikrozensus 1989 umfasste ein Auswahlbezirk durchschnittlich 23 Wohnungen. Zur Verbesserung der Regionalisierbarkeit wurde 1990 die durchschnittliche Größe der Auswahlbezirke verkleinert. Ein Auswahlbezirk der Grundausswahl (ohne Gemeinschaftsunterkünfte) umfasst seitdem durchschnittlich neun Wohnungen. Für Auswahlbezirke der eigenen Gebäudeschicht mit Gemeinschaftsunterkünften beträgt die Richtgröße 15 Personen.

Als Auswahlgrundlage dient im früheren Bundesgebiet das Datenmaterial der Volkszählung (ab dem Mikrozensus 1990 die Volkszählung 1987) und in den neuen Bundesländern das „Bevölkerungsregister Statistik“ 1990 der DDR. Die Stichprobe wird jährlich durch eine Neubausauswahl auf Grundlage der Bautätigkeitsstatistik aktualisiert, wobei die Auswahlbezirke der Neubausauswahl im Unterschied zur Grundausswahl sechs Wohnungen umfassen.

Befragt werden alle in den ausgewählten Bezirken wohnenden Haushalte (alle Haushaltsmitglieder).

Erhebungsprogramm

Das Fragenprogramm des Mikrozensus besteht aus einem gleich bleibenden Grundprogramm, das sozioökonomische Grundinformationen liefert und die laufende Beobachtung des Arbeitsmarktes ermöglicht. Im **Grundprogramm** sind im Mikrozensus 2007 folgende Themenbereiche enthalten:

- Merkmale der Person, Familien-, Haushaltszusammenhang, Haupt- und Nebenwohnung, Aufenthaltsdauer, Staatsangehörigkeit (Wohn-/ Lebensgemeinschaft freiwillig)
- Fragen für eingebürgerte Personen
- Fragen für Ausländer (freiwillig)
- Quellen des Lebensunterhalts, Höhe des Einkommens
- Angaben zur Pflegeversicherung (Leistungen aus der Pflegeversicherung nach Pflegestufen)
- Angaben zur Rentenversicherung
- Besuch von Schule und Hochschule, allgemeiner Schul- und beruflicher Ausbildungsabschluss, allgemeine und berufliche Weiterbildung
- Erwerbstätigkeit, Arbeitslosigkeit und Arbeitsuche, frühere Erwerbstätigkeit, Nichterwerbstätige, Nichterwerbspersonen
- Situation ein Jahr vor der Erhebung (freiwillig)

Neben den jährlich zu erhebenden Merkmalen des Grundprogramms werden in vierjährigen Zusatzprogrammen weitere Angaben erhoben. Im Unterschied zu den [Zusatzprogrammen bis 2004](#), die für Substichproben mit variablem Auswahlsatz erhoben wurden, werden Zusatzprogramme seit 2005 mit dem vollen Auswahlsatz von einem Prozent durchgeführt. Das **Zusatzprogramm** umfasst 2007:

- Angaben zur Krankenversicherung (zuletzt im Mikrozensus 2003 erhoben) sowie
- Zusatzangaben zur Erwerbstätigkeit: Überwiegend ausgeübte Tätigkeit; Betriebs-, Werksabteilung; Stellung im Betrieb (zuletzt im Mikrozensus 2004 erhoben)

Im Rahmen des EU Labour Force Survey wurde 2007 das **Ad-hoc Modul** „Arbeitsunfälle, Arbeit und Gesundheit“ mit einem Auswahlsatz von 0,1 % erhoben. Die Themen der nächsten Ad-hoc Module 2008 bis 2012 sind:

- 2008: Migration und Arbeitsmarkt
- 2009: Übergänge von der Schule ins Erwerbsleben
- 2010: Vereinbarkeit von Beruf und Familie
- 2011: Beschäftigung behinderter Personen
- 2012: Übergang vom Erwerbsleben in den Ruhestand

Gesamtüberblick der Erhebungsprogramme:

- [ab 2005](#)
- [1996-2004](#)
- [1985-1995](#)
- [1973-1985](#)

2 Die wichtigsten Veränderungen beim Mikrozensus 2007 in Stichworten

- Der MZ 2007 enthält im Unterschied zur Erhebung 2006 keine sogenannten Jahresüberhänge (das waren Haushalte bzw. Personen, die im Erhebungszeitraum (2005) i. d. R. aufgrund schwerer Erreichbarkeit nicht, sondern erst im Folgejahr (2006) befragt werden konnten).
- Um insbesondere die Erfassung des ILO-Erwerbsstatus zu verbessern, sind seit 2005 einige Fragen zur Arbeitsmarktbeteiligung (siehe insbes. EF77-EF86) und die Feldarbeit verändert worden. Verteilungsunterschiede können deshalb sowohl mit Veränderungen des Arbeitsmarktes als auch methodisch mit geänderten Fragen zusammenhängen (siehe Köhne-Finster und Lingnau 2008; Körner und Puch 2009).
- Auch im Vergleich zwischen 2007 und 2006 weisen einige Fragen einen anderen Wortlaut auf bzw. enthalten andere Antwortkategorien.
- Mit dem Ziel, die Analysemöglichkeiten des SUF weiter zu verbessern, wurden 2007 bei verschiedenen Merkmalen die Kriterien der Zusammenfassung modifiziert.
- Diese Punkte betreffen folgende Variablen, deren Änderungen in eckigen Klammern und kursiver Schrift kurz kommentiert sind:

| Variablen im Mikrozensus 2007 und 2006 mit Änderungen | | Fragennummer | |
|---|---|--------------|------|
| | | 2007 | 2006 |
| EF1 | Bundesland [Code im MZ 2007 (wie bis 2004) gemäß amtlichem Gemeindegemeinschaftsschlüssel; im MZ 2005 und 2006 nach alphabetischer Sortierung.] | | |
| EF10 | Baujahr der Wohnung [Im MZ 2006 (Zusatzprogramm Wohnsituation: EF494) differenzierter erhoben.] | 19 | 122 |
| EF49 | Familienstand [Neue Merkmalskategorien: "eingetragener Lebenspartner verstorben"; "eingetragene Lebenspartnerschaft aufgehoben"] Siehe auch EF735, EF757, EF765, EF823, EF847 | 8 | 8 |
| EF77 | Bezahlte Tätigkeit (Berichtswoche) | 21 | 20 |
| EF78 | Elternzeit, (Sonder-) Urlaub o.Ä. (Berichtswoche) | 21b | 21 |
| EF79 | Mithelfender Familienangehöriger oder Gelegenheits-tätigkeit (Berichtswoche) | 22 | 22 |
| EF80 | Sonstige bezahlte Tätigkeit | 23 | 23 |
| EF81 | Geringfügige Beschäftigung (Berichtswoche) | 24 | 24 |
| EF82 | Geringfügige Beschäftigung (Berichtsw.): einzige oder hauptsächl. Erwerbstätig. | 25 | 24a |
| EF83 | Frühere Erwerbstätigkeit | 27 | 26 |
| EF85 | Erwerbs- oder Berufstätigkeit (Berichtsw.) | 21a | |
| EF86 | Geringfügige Beschäftigung (Berichtsw.): Elternzeit, (Sonder-) Urlaub o.Ä. | 24a | |
| EF172 | Überw. ausgeübte Tätigkeit [Zusatzprogramm 2007] | 55 | |

| Variablen im Mikrozensus 2007 und 2006 mit Änderungen | | Fragennummer | |
|---|--|--------------|------|
| | | 2007 | 2006 |
| EF173 | Betrieb: Gliederung in Abteilungen [Zusatzprogramm 2007] | 56 | |
| EF174 | Abteilung des Arbeitsplatzes [Zusatzprogramm 2007] | 56a | |
| EF175 | Stellung im Betrieb [Zusatzprogramm 2007] | 56b | |
| EF176 | Überw. ausgeübte Tätigkeit: Computernutzung [Zusatzprogramm 2007] | 55a | |
| EF190 | Arbeitsstätte: Staat (teilw. Zusammenfassung von Unterregionen) [Neue Zusammenfassungen im SUF nach den Kriterien: 1. Aufnahmedatum in die EU, 2. Räumliche Lage] Siehe auch EF455 | 57c | 54c |
| EF203 | 2. Erwerbstätig.: Beruf (KldB92) [Neue Zusammenfassungen im SUF zur Vermeidung von Zusammenfassungen über Berufsgruppen hinweg. Sofern eine Berufsgruppe zusammengefasst wurde, wird der Code der Berufsgruppe (Zweisteller) angegeben.] | 61 | 58 |
| EF214 | 2. Erwerbstätig.: Wirtschaftszweig Betrieb (WZ03) [Neue Zusammenfassungen im SUF zur Vermeidung von Zusammenfassungen über WZ-Abteilungen hinweg. Sofern eine WZ-Abteilung zusammengefasst wurde, wird der Code der Abteilung (Zweisteller) angegeben.] | 62 | 59 |
| EF244 | Arbeitslosengeld I oder II [Neuer Fragentext: Hartz IV] | 76 | 73 |
| EF312 | Höchster berufl. Ausbildungs- oder Hochschul-/ Fachhochschulabschluss [Kategorie 11 "Vorbereitungsdienst für den mittleren Dienst in der öffentlichen Verwaltung" 2007 extra; bis 2006 in Kategorie 3 "Abschluss einer Lehrausbildung ..." enthalten.] Siehe auch EF316, EF744, EF832, EF856 | 91a | 88a |
| EF313 | Höchster berufl. Abschluss: Hauptfachrichtung (HFR03) [Neue Zusammenfassungen im SUF nach den Kriterien: 1. ISCED-Ausbildungsfelder, 2. ISCED-Bildungsfelder] | 91b | 88b |
| EF328 | Allg./Berufl. Weiterbildung (letzte): Fachrichtung [Neue Zusammenfassungen im SUF nach den Kriterien: 1. ISCED-Ausbildungsfelder, 2. ISCED-Bildungsfelder, 3. ISCED breitere Felder.] | 97 | 94 |
| EF330 | Allg./Berufl. Weiterbildung (im letzten Jahr): Stunden [Differenziertere Darstellung im SUF 2007] | 95a | 92a |
| EF369 | 1. Staatsangehörigkeit [Neue Zusammenfassungen im SUF nach den Kriterien: 1. Aufnahmedatum in die EU, 2. Räumliche Lage] Siehe auch EF374, EF733, EF755, EF821, EF845 | 101a | 98a |
| EF372 | Erwerb deutsche Staatsangehörigk.: Geburt/Einbürgerung/Spätaussiedler [Änderung der Frage und Merkmalsausprägungen] | 103 | 100 |

| Variablen im Mikrozensus 2007 und 2006 mit Änderungen | | Fragennummer | |
|---|---|--------------|------|
| | | 2007 | 2006 |
| EF373 | Einbürgerung/Zuzug als Spätaussiedler: Jahr [Änderung der Frage] | 103a | 100a |
| EF401 | Überwiegender Lebensunterhalt [Neue Antwortkategorien] Siehe auch EF741, EF829, EF853 | 105 | 104 |
| EF421 | Öffentl. Zahlung: Wohngeld [Neue Antwortkategorien] Siehe auch EF422-EF426, EF439 | 109 | 108 |
| EF521 | Letzte Erwerbstätigk.: Beruf nach ISCO-88 COM [Für die Umschlüsselung von der Klassifikation der Berufe zu ISCO fehlen für den früheren Beruf Angaben zur Betriebsgröße. Siehe zum Schätzverfahren die Ausführungen in Klassifikation der Berufe (KldB-92) in der Signierung für den Mikrozensus - Statistisches Bundesamt, Ausgabe 1992 - in Verbindung mit dem Umsteigerschlüssel für die Internationale Standardklassifikation der Berufe zur Anwendung in der Europäischen Gemeinschaft (ISCO-88 COM) [pdf] .] | | |

- Im Rahmen der 0,1%-Substichprobe (siehe EF6) des Ad-hoc-Moduls „Arbeitsunfälle, Arbeit und Gesundheit“ (EF1040-EF1051) sind 2007 weitere ergänzende Fragen ohne Auskunftspflicht zur Erwerbstätigkeit gestellt worden:
 - EF118 Leitungstätigkeit (F127)
 - EF177 Arbeitsvertrag: Zeitarbeitsvermittlungsfirma (F126)
 - EF178 Arbeitsvermittlung: Agentur für Arbeit (F128)
 - EF179 Nichtarbeitsuche oder Teilzeitbeschäftigung: Hauptgrund (F129)
- Weitere Hinweise zur [Vergleichbarkeit des Mikrozensus bis 2004 vs. Mikrozensus ab 2005](#) finden sich im GESIS Web.

3 Kurzbeschreibung des Scientific Use Files

Das Scientific Use File (SUF) ist eine faktisch anonymisierte 70%-Substichprobe; i. d. R. eine Haushaltssubstichprobe. In Erhebungsjahren des Zusatzprogramms mit Fragen zur Wohnsituation wird im Unterschied zu den anderen Erhebungen nicht eine Haushaltssubstichprobe, sondern eine 70%-Substichprobe der Wohnungen gezogen. Das Ziehungsverfahren orientiert sich eng am Stichprobendesign des Mikrozensus.

Die Haushaltssubstichprobe ist eine systematische Zufallsauswahl aus den Originaldaten. Die Daten werden dazu nach Berichtsquartal, Bundesland, Regierungsbezirk, Gemeindegrößenklasse, Anzahl der Personen im Haushalt, Gebäudegrößenklasse, Auswahlbezirksnummer und Haushaltsnummer sortiert. (Das Merkmal Gebäudegrößenklasse wurde im SUF 2006 erstmals verwendet.) Durch diese Anordnung ist gewährleistet, dass die Substichprobe hinsichtlich dieser Merkmale nur geringe zufallsbedingte Abweichungen zum Originalmaterial 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).

Weiterhin unterscheidet sich das SUF vom Originalfile des Mikrozensus dadurch, dass bestimmte Variablen im SUF, bedingt durch die Anonymisierung, in klassierter und vergrößerter Form verfügbar sind. So sind im SUF die Regionalangaben Bundesland und eine vergrößerte Gemeindegrößenklasse, die nicht für die kleinen Bundesländer 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 Fälle 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.

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 finden Sie am Ende der [thematischen Übersicht](#).

Zur Abgrenzung der Bevölkerungsbegriffe und zu Gewichtung und Hochrechnung

Zum Vergleich mit den vom Statistischen Bundesamt veröffentlichten Zahlen sind die jeweiligen [Bevölkerungs- und Erwerbskonzepte](#) und Hochrechnungsfaktoren zu berücksichtigen; siehe dazu die Webseite von GESIS unter [Mikrozensus 2007](#)

- Definitionen, Erläuterungen und methodische Hinweise zum Mikrozensus ab 2005 ([Definitions-katalog 2005](#), [Definitions-katalog 2006](#); [Definitions-katalog 2007](#))
- [Definitionen zum Zusatzprogramm Krankenversicherung 2007](#)

Die Hochrechnungsfaktoren

- EF951 Hochrechnungsfaktor Quartal
- EF952 Hochrechnungsfaktor Jahr
- EF955 Hochrechnungsfaktor Ad-hoc-Substichprobe: Jahr

des Mikrozensus SUF stammen aus dem vollen Mikrozensus. Sie sind bereits mit dem Kehrwert der Auswahlwahrscheinlichkeit der Substichprobe multipliziert (100/70) und rechnen auf 1.000 Personen bzw. Wohnungen, Haushalte, Familien usw. der Population hoch. Um auf 1 % der Bevölkerung, also auf den vollen Mikrozensus hochzurechnen, sind die auf 1.000 normierten Hochrechnungsfaktoren mit 10 zu multiplizieren.

Zugang zum Mikrozensus Scientific Use File und Beratung

Die Mikrozensus [Scientific Use Files](#) können von unabhängigen inländischen Forschungseinrichtungen gegen eine Bereitstellungsgebühr von 95 EURO bei den Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder bestellt werden.

Interessenten und Nutzer des Mikrozensus können sich auch an das GML bei GESIS wenden und Beratungsleistungen in Anspruch nehmen.

4 Literaturhinweise

- Afentakis, A.; Bihler, W. (2005): Das Hochrechnungsverfahren beim unterjährigen Mikrozensus ab 2005. [Wirtschaft und Statistik \(10\): 1039-1048](#).
- Berke, P. (2009): Ergebnisse einer nordrhein-westfälischen Zusatzbefragung der Interviewerinnen und Interviewer im Mikrozensus zu ihren Erfahrungen in der Feldarbeit. S. 15-35 in: Information und Technik Nordrhein-Westfalen (IT.NRW), Geschäftsbereich Statistik (Hrsg.): [Statistische Analysen und Studien NRW, Band 55](#).
- Gerhardt, A.; Habenicht, K.; Munz, E. (2009): Analysen zur Einkommensarmut mit Datenquellen der amtlichen Statistik. In: Information und Technik Nordrhein-Westfalen (Hrsg.): [Statistische Analysen und Studien NRW, Band 58](#).
- Iversen, K. (2007): Das Mikrozensusgesetz 2005 und der Übergang zur Unterjährigkeit. [Wirtschaft und Statistik \(1\): 38-44](#).
- Köhne-Finster, S.; Güllner, G. (2009): Ergebnisse der Interviewerbefragung im Mikrozensus. [Wirtschaft und Statistik \(5\): 397-405](#).
- Köhne-Finster, S.; Lingnau, A., (2009): Untersuchung der Datenqualität erwerbsstatistischer Angaben im Mikrozensus. Ergebnisse des Projekts „Nachbefragung im Mikrozensus/ LFS“. [Wirtschaft und Statistik 12/2008: 1067-1088](#).
- Körner, T. Puch, K. (2009): Der Mikrozensus im Kontext anderer Arbeitsmarktstatistiken - Ergebnisunterschiede und ihre Hintergründe. [Wirtschaft und Statistik \(6\): 528 - 552](#).
- Lotze, S.; Breiholz, H. (2002): Zum neuen Erhebungsdesign des Mikrozensus. Teil 1. [Wirtschaft und Statistik \(5\): 359-366](#).
- Lotze, S.; Breiholz, H. (2002): Zum neuen Erhebungsdesign des Mikrozensus. Teil 2. [Wirtschaft und Statistik \(6\): 454-459](#).
- Nöthen, M. (2005): Von der „traditionellen Familie“ zu „neuen Lebensformen“. Neuerungen in der Familienberichterstattung des Mikrozensus. *Wirtschaft und Statistik* (1): 25-40.
- Schäfer, T.; Brückner, G. (2008): Soziale Homogenität der Bevölkerung bei alternativen Definitionen für Migration. Eine Analyse am Beispiel von Bildungsbeteiligung, Erwerbstätigkeit und Einkommen auf der Basis von Mikrozensusdaten. [Wirtschaft und Statistik \(12\): 1046-1066](#).
- Statistisches Bundesamt (2009): Mikrozensus: [Haushaltszahlen ab 2005](#). Wiesbaden.
- Wolf, C.; Lüttinger, P. (2009): Proxy-Angaben im deutschen Mikrozensus, S. 395-412 in: Johann Bacher und Christof Wolf. Herausforderungen und Grenzen der Umfrageforschung. Sonderheft 7 der Österreichischen Zeitschrift für Soziologie. Wiesbaden: VS Verlag.
- Zühlke, S. Auswirkungen von Proxy-Interviews auf die Datenqualität des Mikrozensus. S. 3-10 in: Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen (Hrsg.): [Statistische Analysen und Studien NRW, Band 53](#).

[Veröffentlichungen](#) auf der Datengrundlage von Scientific Use Files finden sich auf der Webseite des GML

Ergebnisse des Mikrozensus 2007

- Statistisches Bundesamt (2009): Fachserie 1 Bevölkerung und Erwerbstätigkeit, Reihe 2.2 Bevölkerung mit Migrationshintergrund. Ergebnisse des Mikrozensus 2007. Wiesbaden. [\[Download\]](#)
- Statistisches Bundesamt (2008): Fachserie 1 Bevölkerung und Erwerbstätigkeit, Reihe 3 Haushalte und Familien. Ergebnisse des Mikrozensus 2007. Wiesbaden. [\[Download\]](#)
- Statistisches Bundesamt (2008): Fachserie 1 Bevölkerung und Erwerbstätigkeit, Reihe 4.1.1 Stand und Entwicklung der Erwerbstätigkeit. Mikrozensus 2007. [verschiedene Bände; Band 1 [Download](#), Band 2 [Download](#)]
- Statistisches Bundesamt (2008): Fachserie 1 Bevölkerung und Erwerbstätigkeit, Reihe 4.1.2 Beruf, Ausbildung und Arbeitsbedingungen der Erwerbstätigen. Band 1: Allgemeine und methodische Erläuterungen. Mikrozensus 2007. Wiesbaden. verschiedene Bände; Band 1 [Download](#), Band 2 [Download](#)]
- Statistisches Bundesamt (2008): Fachserie 13 Sozialleistungen, Reihe 1.1 Angaben zur Krankenversicherung. Ergebnisse des Mikrozensus 2007. Wiesbaden. [\[Download\]](#)
- Statistisches Bundesamt (2008): Qualitätsbericht Mikrozensus 2007. Wiesbaden [\[pdf ca.212kB\]](#).

Unter [Statistisches Bundesamt | Publikationsservice | Detailinformation](#) (Schnellsuche: Mikrozensus) kann auf zahlreiche kostenlose Publikationen aus dem Bereich Mikrozensus zugegriffen werden.

Mikrozensusgesetz 2005

- Gesetz zur Durchführung einer Repräsentativstatistik über die Bevölkerung und den Arbeitsmarkt sowie die Wohnsituation der Haushalte [\[pdf ca.435kB\]](#).

EU-Verordnungen zur Arbeitskräfteerhebung

- Verordnung (EG) Nr. 577/98 des Rates vom 9. März zur Durchführung einer Stichprobenerhebung über Arbeitskräfte in der Gemeinschaft [\[pdf ca.111kB\]](#).
- Verordnung (EG) Nr. 1575/2000 der Kommission vom 19. Juli 2000 zur Durchführung der Verordnung (EG) Nr. 577/98 des Rates zur Durchführung einer Stichprobenerhebung über Arbeitskräfte in der Gemeinschaft hinsichtlich der von 2001 an für die Datenübermittlung zu verwendenden Codierung [\[pdf ca.167kB\]](#).
- Verordnung (EG) Nr. 2257/2003 des Europäischen Parlaments und des Rates vom 25. November 2003 zur Änderung der Verordnung (EG) Nr. 577/98 des Rates zur Durchführung einer Stichprobenerhebung über Arbeitskräfte in der Gemeinschaft im Hinblick auf die Anpassung der Liste der Erhebungsmerkmale [\[hier\]](#).
- Verordnung (EG) Nr. 1897/2000 der Kommission vom 7. September 2000 zur Umsetzung der Verordnung (EG) Nr. 577/98 des Rates zur Durchführung einer Stichprobenerhebung über Arbeitskräfte in der Gemeinschaft bezüglich der Arbeitsdefinition der Arbeitslosigkeit [\[pdf ca. 151kB\]](#).
- Verordnung (EG) Nr. 246/2003 der Kommission vom 10. Februar 2003 zur Annahme des Programms von Ad-hoc-Modulen für die Erhebung über Arbeitskräfte nach der Verordnung (EG) Nr. 577/98 des Rates für den Zeitraum 2004-2006 [\[pdf ca.469kB\]](#).

- Verordnung (EG) Nr. 29/2004 der Kommission vom 8. Januar 2004 zur Annahme der Spezifikationen des Ad-hoc-Moduls 2005 „Vereinbarkeit von Beruf und Familie“ nach der Verordnung (EG) Nr. 577/98 des Rates [\[pdf ca.159kB\]](#).
- Verordnung (EG) Nr. 384/2005 der Kommission vom 7. März 2005 zur Annahme des Programms von Ad-hoc-Modulen für die Jahre 2007 bis 2009 für die Stichprobenerhebung über Arbeitskräfte gemäß der Verordnung (EG) Nr. 577/98 des Rates [\[pdf ca. 100kB\]](#).
- Verordnung (EG) Nr. 388/2005 der Kommission vom 8. März 2005 zur Annahme der Spezifikationen des Ad-hoc-Moduls 2006 - Übergang vom Erwerbsleben in den Ruhestand - nach der Verordnung (EG) Nr. 577/98 des Rates und zur Änderung der Verordnung (EG) Nr. 246/2003 [\[pdf ca. 112kB\]](#).
- Verordnung (EG) Nr. 430/2005 der Kommission vom 15. März 2005 zur Durchführung der Verordnung (EG) Nr. 577/98 des Rates zur Durchführung einer Stichprobenerhebung über Arbeitskräfte in der Gemeinschaft im Hinblick auf die ab 2006 für die Datenübermittlung zu verwendende Kodierung und die Nutzung einer Teilstichprobe für die Datenerhebung zu Strukturvariablen (Gültig bis 31. Dezember 2008) [\[pdf ca. 113kB\]](#).
- Verordnung (EG) Nr. 365/2008 der Kommission vom 23. April 2008 zur Annahme des die Jahre 2010, 2011 und 2012 umfassenden Programms von Ad-hoc-Modulen für die Stichprobenerhebung über Arbeitskräfte gemäß der Verordnung (EG) Nr. 577/98 des Rates [\[pdf ca. 65kB\]](#).
- Verordnung (EG) Nr. 377/2008 der Kommission vom 25. April 2008 zur Durchführung der Verordnung (EG) Nr. 577/98 des Rates zur Durchführung einer Stichprobenerhebung über Arbeitskräfte in der Gemeinschaft im Hinblick auf die ab 2009 für die Datenübermittlung zu verwendende Kodierung, die Verwendung einer Teilstichprobe für die Datenerhebung zu Strukturvariablen und die Definition der Referenzquartale [\[pdf ca. 848kB\]](#).

Informationen zum Mikrozensus allgemein und seiner Konzeption

- Emmerling, D.; Riede, T. (1997): 40 Jahre Mikrozensus. *Wirtschaft und Statistik* (3), S. 160-174.
- Hartmann, P. (1989): Der Mikrozensus als Datenquelle für die Sozialwissenschaften. [ZUMA-Nachrichten 24, S. 6-25](#).
- Heidenreich, H.-J. (1994): Hochrechnung des Mikrozensus ab 1990. In: Gabler, S.; Hoffmeyer-Zlotnik, J.; Krebs, D. (Hrsg.): *Gewichtung in der Umfragepraxis*. Westdeutscher Verlag, Opladen, S. 112-123
- Lengerer, A.; Bohr, J.; Janßen, A. (2005): Haushalte, Familien und Lebensformen im Mikrozensus - Konzepte und Typisierungen. [ZUMA-Arbeitsbericht 2005/05](#).
- Lüttinger, P.; Riede, T. (1997): Der Mikrozensus: amtliche Daten für die Sozialforschung. [ZUMA-Nachrichten 4, S. 19-43](#).
- Meyer, K. (1994): Zum Auswahlplan des Mikrozensus ab 1990. In: Gabler, S.; Hoffmeyer-Zlotnik, J.; Krebs, D. (Hrsg.): *Gewichtung in der Umfragepraxis*. Westdeutscher Verlag, Opladen, S. 106-111.
- Statistisches Bundesamt (2004): *Arbeitsunterlagen zum Mikrozensus. Das Erhebungsprogramm des Mikrozensus seit 1957*. Bonn.
- Statistisches Bundesamt (1989): *Inhaltliche Fragen bevölkerungsstatistischer Stichproben am Beispiel des Mikrozensus. Bericht zur Konferenz vom 21. und 22. Okto-*

ber 1988. Heft 10 der Schriftenreihe Ausgewählte Arbeitsunterlagen zur Bundesstatistik. Wiesbaden.

- Schmidt, S. (2000): Erwerbstätigkeit im Mikrozensus. Konzept, Definition, Umsetzung. [ZUMA-Arbeitsbericht 00/01](#).
-

Informationen zur Weitergabe faktisch anonymisierter Einzeldaten

- Müller, W.; Blien, U.; Knoche, P.; Wirth, H. (1991): Die Faktische Anonymität von Mikrodaten, (Band 19 der Schriftenreihe Forum der Bundesstatistik). Metzler-Poeschel, Stuttgart.
 - Wirth, H. (1992): Die faktische Anonymität von Mikrodaten: Ergebnisse und Konsequenzen eines Forschungsprojektes. ZUMA-Nachrichten Nr. 30, S. 7-65.
-

Informationen zur Fehlerrechnung im Mikrozensus

- Krug, W.; Nourney, M.; Schmidt, J.: (1994): Wirtschafts- und Sozialstatistik. Gewinnung von Daten. München: Oldenbourg (3. völlig Neubearb. Auflage). [für die Originaldaten des Mikrozensus]
 - Rendtel, U.; Schimpl-Neimanns, B. (2001): Die Berechnung der Varianz von Populationsschätzern im Scientific Use File des Mikrozensus ab 1996. ZUMA-Nachrichten 48: 85-116 [[pdf ca.530kB](#)]. [siehe auch [Mikrodaten-Tools](#)]
 - Schimpl-Neimanns, B. (2010): Varianzschätzung für Mikrozensus Scientific Use Files ab 2005. [Gesis-Technical Report 2010/03](#). Mannheim.
-

Anhang

Informationen zur Nutzung der Setups

Dieser Abschnitt dokumentiert, wie die Systemfiles auf Basis der Rohdaten erstellt wurden bzw. mittels der Setups des German Microdata Lab (GML) der GESIS selbst erstellt werden können.

Die vom GML bereitgestellten Setups für das Mikrozensus SUF 2007 dienen zum Einlesen des Rohdatenmaterials und zum Erstellen von Systemfiles für die Programmpakete PASW (früher: SPSS), SAS und STATA. Sie beinhalten Programmanweisungen zur Definition von fehlenden Werten sowie zum Versehen der Variablen und ihrer Ausprägungen mit entsprechenden Labels.

Spezifika der Statistikprogramme:

- Am Anfang der Setups sind in der **Configuration Section** die vollständigen lokalen Dateinamen (einschließlich Laufwerkskennzeichen und Verzeichnis) zu nennen. Ansonsten sind im Setup keine weiteren Änderungen vorzunehmen.

PASW Statistics 17.0 Release
17.0.2 [SPSS]

STATA/SE 10.1 for Windows

SAS 9.1 for Windows

```
file handle raw_data / name =  
"RAW_DATA_FILENAME" /  
lrecl = 687.
```

```
local log_file "LOG_FILENAME.log" ;  
local raw_data_file  
"RAW_DATA_FILENAME" ;
```

```
%let raw_data_filename =  
'RAW_DATA_FILENAME' ;
```

```
file handle spss_reference /  
name = "SPSS_FILENAME.sav"
```

```
local stata_file "STA-  
TA_FILENAME.dta" ;
```

```
%let sas_file_dirname =  
'SAS_FILE_DIRECTORY_PATH' ;
```

```
%let sas_file_basename =  
SAS_FILENAME ;
```

- **Stata:** Aufgrund der Größe des Setups können die o. g. lokalen Dateinamen nicht mit dem internen **Do-File Editor** eingetragen werden, sondern hierfür muss ein anderer Editor verwendet werden. Danach kann das Setup-Programm mit dem Stata-Kommando `do setup07.do` gestartet werden. (Bei Verwendung älterer Stata-Versionen kann die Zeichenzahl der Labels dieses Setups die maximal zulässige Zeichenzahl überschreiten. Ggf. müssen die Labels dieses Setups manuell angepasst werden.)
- **PASW** (bzw. SPSS) benötigt für temporäre Operationen ca. das Dreifache des Plattenplatzes des Rohdatenfiles, sodass ausreichend Speicherplatz zur Verfügung stehen sollte. (Bei Verwendung einer früheren Version ist „set unicode“ auszukommentieren.)
- **Missing Values:** In PASW bzw. SPSS können fehlende Werte als benutzerdefinierte Missings deklariert und mit entsprechenden Werten und Labels ausgewiesen werden. Sie zählen bei Analysen nicht als gültige Werte. SAS und STATA kennen keine benutzerdefinierten Missings bzw. fehlende Werte, sondern nur System Missings. Im PASW-Setup werden benutzerdefinierte Missings spezifiziert. Diese Definitionen sind zwar auch in den SAS- und STATA-Setups enthalten, im Unterschied zum PASW-Setup je-

doch auskommentiert, da sonst die Werte des Rohdatenfiles im Systemfile durch das System Missing Zeichen (.) ersetzt würden. Bei Bedarf können diese im Setup unter dem Kommentar „Definition of system-missing values“ stehenden Programmanweisungen durch Entfernen der Kommentarzeichen (/*, */) am Beginn und Ende des Anweisungsblocks aktiviert werden.

- **Voreinstellungen:** Um eine reibungsfreie Aufbereitung der Daten zu gewährleisten, empfiehlt es sich, die in den Setups vorgesehenen Voreinstellungen nicht zu verändern und ggf. das Setup – je nach Leistung des verarbeitenden Systems – schrittweise durchführen zu lassen.

| PASW Statistics 17.0 | STATA/SE 10.1 for Windows | SAS 9.1 for Windows |
|--|---------------------------|-----------------------------|
| set decimal dot. | set more off | options compress = binary ; |
| * Bei Verwendung von SPSS vor V16 ggf. "set unicode" auskommentieren . | set memory 600m | |
| set unicode no. | set dp comma | |
| set olang = GERMAN. | #delimit ; | |

Ist das entsprechende Systemfile erstellt, kann die einfache Fallzahl $n = 483.595$ (ohne Gewichtung, ohne Selektion) zur Kontrolle, ob der Rohdatensatz fehlerfrei eingelesen wurde, mit der des erstellten Datensatzes verglichen werden. Unterscheiden sich die Fallzahlen, weist dies meist auf einen Fehler beim Einlesen hin. Zusätzlich können die folgenden Verteilungen zur Prüfung des fehlerfreien Ablaufs des Setups herangezogen werden. Weitere Randverteilungen ausgewählter Variablen des Mikrozensus SUF 2007 sind im GESIS Web zu finden.

Verteilungen zur Prüfung der Systemfiles bzw. des Ablaufs der Setups

| Bevölkerungskonzept | Bevölkerung in Privathaushalten | Bevölkerung am Ort der Hauptwohnung | Bevölkerung in Lebensformen am Hauptwohnsitz | Bevölkerung am Ort der Hauptwohnung, Erwerbstätige | Fallzahlen |
|---|---------------------------------|-------------------------------------|--|--|--------------|
| <i>Selektion</i> | <i>ef31=1</i> | <i>ef30>=1 & ef30<=2</i> | <i>ef761>=1 & ef761<=2</i> | <i>ef30>=1 & ef30<=2 & ef29=1</i> | <i>keine</i> |
| <i>Hochrechnung, Gewichtung</i> | <i>ef952</i> | <i>ef952</i> | <i>ef952</i> | <i>ef952</i> | <i>keine</i> |
| Insgesamt (in 1.000; außer Fallzahlen) | 82.373 | 82.250 | 81.544 | 38.151 | 483.595 |

EF117: Stellung im Beruf x EF29: Erwerbstyp; Selektion: Bevölkerung am Ort der Hauptwohnung (*ef30>=1 & ef30<=2*); (siehe Fachserie 1, Reihe 4.1.1, Band 2, Tabelle 1.1, S. 3)

| | | |
|---|---------------|----------------|
| Erwerbstätige | | |
| Selbstständige | 4.147 | 24.408 |
| Mithelf. Familienangeh. | 397 | 2.319 |
| Beamte (inkl. Soldaten) | 2.236 | 12.758 |
| Angestellte (inkl. kfm./techn. Auszubildende u. Zivildienstleistende) | 19.896 | 116.154 |
| Arbeiter (inkl. gew. Auszubildende) | 11.475 | 64.893 |
| Erwerbstätige insges. | 38.151 | 220.532 |
| Erwerbslose | 3.624 | 19.915 |
| Erwerbspersonen | 41.775 | 240.447 |
| Nichterwerbspersonen | 40.475 | 243.148 |
| Insgesamt | 82.250 | 483.595 |

EF663: Personen im Haushalt: Anzahl (Haushaltstypisierung); Selektion: Bevölkerung in Privathaushalten (*ef31=1*), *Haushaltsbezugsperson [Haupteinkommensbezieher]* (*ef662>=1 & ef662<=4*); (siehe Fachserie 1, Reihe 3, Tabelle 1.1, S. 26)

| | | |
|----------------------------|---------------|----------------|
| 1 Person | 15.384 | 88.522 |
| 2 Personen | 13.519 | 80.376 |
| 3 Personen | 5.318 | 30.360 |
| 4 Personen | 4.075 | 23.510 |
| 5 Personen und mehr | 1.443 | 8.258 |
| Haushalte insgesamt | 39.738 | 231.026 |

EF118: Leitungstätigkeit (FB) (0,1 %); Selektion: Unterstichprobe (*ef6=2*), *erwerbstätige Bevölkerung am Hauptwohnsitz* (*ef29=1 & ef30>=1 & ef30<=2*); Gewichtungsvariable: *ef955 Hochrechnungsfaktor Ad-hoc-Substichprobe: Jahr*

| | | |
|-----------------------------------|---------------|---------------|
| Ja | 4.141 | 2.438 |
| Trifft nicht zu, da selbstständig | 4.137 | 2.460 |
| Nein | 24.131 | 14.139 |
| Keine Angabe | 5.689 | 3.276 |
| Insgesamt | 38.097 | 22.313 |

Als Referenz zur Plausibilitätsprüfung der Mikrozensus SUF dienen die in den Fachserien des Statistischen Bundesamtes veröffentlichten Ergebnisse des jeweiligen Mikrozensus. Für den Mikrozensus 2007 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) und Reihe 2.2 (Bevölkerung mit Migrationshintergrund) sowie Fachserie 13 Sozialleistungen, Angaben zur Krankenversicherung. Bei diesen Vergleichen sind die von den Statistischen Ämtern verwendeten Bevölkerungs- und Erwerbskonzepte sowie die Hochrechnung zu beachten (vgl. die Seite [Abgrenzung und Hochrechnung der Bevölkerungsbegriffe im Mikrozensus 2007](#) im GESIS Web).

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](#)". Die meisten Variablen des SUF weichen nur in sehr geringem Maße von den veröffentlichten Daten ab. Größere relative Abweichungen sind nur bei Merkmalen festzustellen, die mit sehr geringen Fallzahlen besetzt sind.

Bei der Aufbereitung des Mikrozensus SUF 2007 wurden die Variablenlabels nach der neuen, erstmals für das Mikrozensus SUF 2003 verwendeten Systematik des Mikrodaten-Informationssystems ([MISSY](#)) erstellt. 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. Ein im Vergleich mit früheren Files abweichendes Label bedeutet also i. d. R. nicht, dass sich am Inhalt der entsprechenden Variable etwas geändert hat.

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.