

# **Wichtige Informationen zur Nutzung des Mikrozensus Scientific Use Files 2016**

**Dokumentation und  
Datenaufbereitung**

**Herausgeber:**

---

Statistisches Bundesamt  
Referat H 306 Auswertung und Analyse der Haushaltserhebungen, Wohnen  
Graurheindorfer Str. 198  
53117 Bonn  
Tel.: 0228 / 99 643 - 8725  
<http://www.destatis.de>  
Kontakt: Robert Herter-Eschweiler  
E-Mail: [Referat-h306@destatis.de](mailto:Referat-h306@destatis.de)  
Stand: Juni 2020

GESIS - Leibniz-Institut für Sozialwissenschaften  
German Microdata Lab  
  
B2, 1  
68159 Mannheim  
Tel.: 0621 / 12 46 - 265  
<http://www.gesis.org/gml>  
Kontakt: Tobias Roth  
E-Mail: [gml@gesis.org](mailto:gml@gesis.org)

## Inhalt

Einleitung.....	4
1 Allgemeine Informationen zum Mikrozensus.....	5
1.1 Was ist der Mikrozensus? .....	5
1.2 Das Wichtigste über den Mikrozensus 2016 im Überblick .....	5
1.2.1 Grundgesamtheit.....	5
1.2.2 Berichtszeitraum .....	5
1.2.3 Stichprobe .....	5
1.2.4 Erhebungs- und Analyseeinheiten .....	7
1.2.5 Erhebungsprogramm und Erhebungsverfahren .....	7
1.2.6 Hochrechnung / GewichtungsvARIABLEN.....	8
2 Kurzbeschreibung des Scientific Use Files Mikrozensus 2016 .....	9
2.1 Auswahl der Substichprobe.....	9
2.2 Hochrechnungs- / GewichtungsvARIABLEN im SUF .....	10
2.3 Anonymisierungsmaßnahmen, Regionalangaben.....	11
2.4 Die wichtigsten Veränderungen gegenüber 2015 in Stichworten .....	11
2.5 Möglichkeit der Längsschnittverknüpfung .....	12
2.6 Zur Abgrenzung der Bevölkerungsbegriffe und zur Hochrechnung.....	13
2.7 Zugang zum Mikrozensus Scientific Use File und Beratung .....	14
3 Literaturhinweise .....	15
3.1 Allgemein .....	15
3.2 Ergebnisse des Mikrozensus 2016 .....	19
3.3 Mikrozensusgesetz, EU-Verordnungen.....	19
3.3.1 Mikrozensusgesetz 2005 .....	19
3.3.2 EU-Verordnungen zur Arbeitskräfteerhebung .....	19
A Anhang.....	20
A1 Informationen zur Nutzung der Setups.....	20
A2 Verteilungen zur Prüfung der Systemfiles bzw. des Ablaufs der Setups .....	22

## Einleitung

Für Forscherinnen und Forscher, die sich erstmals für die Daten interessieren als auch für jene, die bereits mit den Daten anderer Erhebungszeitpunkte Erfahrungen gesammelt haben, gibt dieser Bericht einen Überblick zum Erhebungsprogramm und den Erhebungsverfahren des Mikrozensus sowie zur Aufbereitung und Dokumentation des Scientific Use Files 2016. Diese Informationen sollen dabei helfen, die Daten sachgerecht auszuwerten. Im Folgenden werden deshalb zunächst Informationen zum Mikrozensus allgemein sowie zum Erhebungs- und Stichprobendesign genannt. Anschließend werden das Scientific Use File 2016 und die Veränderungen gegenüber dem Vorjahr beschrieben. Nach den Literaturhinweisen finden sich im Anhang Informationen zur Nutzung der Setups zum Einlesen der Rohdaten und zur Erzeugung von Systemfiles, die ab dem Erhebungszeitpunkt 2007 zusätzlich zu den Rohdaten im ASCII-Format für die Statistikprogramme SPSS, SAS und Stata im dokumentierten Format bereitgestellt werden. Der Anhang enthält ebenfalls Verteilungen zur Prüfung der Systemfiles. Um die Ergebnisqualität zu prüfen, werden im Rahmen der Datenaufbereitung für ausgewählte Merkmalsverteilungen Ergebnisse des Scientific Use Files mit Ergebnissen des Mikrozensus verglichen; siehe u. a. den [Vergleich der Randverteilungen](#).

Ausführliche Informationen über den Mikrozensus (MZ) und die Mikrozensus Scientific Use Files (SUF), u. a. Fragebogen, Klassifikationen und Definitionen, stehen auf den [Web-Seiten der GESIS](#) zur Verfügung.

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

[Was ist der Mikrozensus?](#)

[Häufige Fragen zum Mikrozensus](#)

[Forschungsdatenzentren | Datenangebot | Mikrozensus](#)

Die Aufbereitung und Dokumentation des faktisch anonymisierten Einzelmaterials erfolgt durch das Statistische Bundesamt, Referat H 306 – Auswertung und Analyse der Haushalts-erhebungen, Wohnen, in Kooperation mit dem German Microdata Lab (GML) der GESIS.

# 1 Allgemeine Informationen zum Mikrozensus

## 1.1 Was ist der Mikrozensus?

Der Mikrozensus ist die amtliche Repräsentativstatistik 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. Die Arbeitskräftestichprobe der Europäischen Union (EU Labour Force Survey) ist in den Mikrozensus integriert. Im Jahr 2016 wurde die gesamte Stichprobe erneuert, das Frageprogramm basiert aber weiterhin auf dem Mikrozensusgesetz von 2005.

Siehe auch:

[Informationen zum Mikrozensus \[Informationen für die Befragten Mikrozensus 2016\]](#) (Statistische Ämter des Bundes und der Länder, 2015).

Statistisches Bundesamt (2017): [Qualitätsbericht Mikrozensus 2016](#). Wiesbaden.<sup>1</sup>

## 1.2 Das Wichtigste über den Mikrozensus 2016 im Überblick

### 1.2.1 Grundgesamtheit

Zur Grundgesamtheit gehört 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“<sup>2</sup> keine Erfassungschance.

### 1.2.2 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 SUF enthält Angaben zum Berichtsquartal.

### 1.2.3 Stichprobe

Der Mikrozensus ist eine einstufige Klumpenstichprobe („Flächenstichprobe“).

## Auswahlgrundlage

Während bis einschließlich 2015 die Grundausswahl für die alten Bundesländer auf Ergebnissen der Volkszählung 1987, für die neuen Bundesländer und Ost-Berlin auf Ergebnissen aus dem Bevölkerungsregister Statistik der DDR 1990 basierte, dient ab dem Berichtsjahr 2016 der Zensus 2011 als Auswahlgrundlage für den Mikrozensus. Durch die neue Auswahlgrundlage der Stichprobe ergeben sich Einschränkungen in der Vergleichbarkeit zu den Vorjahren. Die Aktualisierung der Grundausswahl erfolgt jährlich durch Berücksichtigung der Neubautätigkeit über die Meldungen zur Bautätigkeitsstatistik (abweichend hiervon wurden im Jahr 2016 die Neubausauswahlen der Jahre 2011 bis 2016 kumuliert).

---

<sup>1</sup> Weite Teile der hier dargestellten allgemeinen Informationen zum Mikrozensus orientieren sich am Qualitätsbericht Mikrozensus 2016 (Statistisches Bundesamt, 2017).

<sup>2</sup> Vereinfachend wird der Mikrozensus zwar als „Flächenstichprobe“ bezeichnet, konkret sind es aber Anschriften mit Wohnraum.

## **Auswahleinheiten**

Die Auswahleinheiten sind Auswahlbezirke (Anschriften mit Wohnraum, die nach der Anzahl der an der Anschrift befindlichen Wohnungen in sogenannte Anschriftengrößenklassen unterteilt werden. So kann ein Auswahlbezirk mehrere (benachbarte) Anschriften umfassen, alle Wohnungen einer Anschrift oder auch nur Teile einer Anschrift. Bei Anschriften ohne Wohnungen wurden die an der Anschrift gemeldeten Personen zur Größenbestimmung des Auswahlbezirks genutzt).

Ein Auswahlbezirk verbleibt vier Jahre in der Stichprobe, in jedem Jahr wird 1/4 der Auswahlbezirke durch neue Auswahlbezirke ersetzt (rotierende Panelstichprobe). Infolge der Änderung der Auswahlgrundlage wurde im Jahr 2016 abweichend hiervon die gesamte Stichprobe erneuert (zu beachten ist allerdings, dass auch im Mikrozensus 2016 Jahresüberhänge enthalten sind). Die aus dem 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). Die Möglichkeit der Zusammenführung der Querschnittsdaten zu einem [Mikrozensuspanel](#) ist rechtlich seit dem Mikrozensusgesetz 1996 gegeben. Aufgrund der Erneuerung der kompletten Stichprobe im Jahr 2016 sind Verknüpfungen der Querschnittsdaten ab dem Erhebungszeitpunkt 2012 nur bis einschließlich der Erhebung 2015 möglich. Zudem werden aufgrund der Neuregelung des Mikrozensus ab 2020 Verknüpfungen der Querschnittsdaten ab dem Erhebungszeitpunkt 2016 nur bis einschließlich der Erhebung 2019 möglich sein.

## **Schichtungsmerkmale**

Neben der regionalen Schichtung (Bundesland und Regionalschicht) wird fachlich nach der Anschriftengrößenklasse geschichtet.

## **Auswahlsatz**

Der Auswahlsatz des Mikrozensus beträgt 1 %. Die Ad-hoc-Module des EU Labour Force Survey werden als Unterstichprobe mit einem Auswahlsatz von 0,1 % durchgeführt.

## **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 00 bis 99 versehen. Vier aufeinanderfolgende Zonen bilden einen Block.

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.

Die geplante Befragungswoche eines Auswahlbezirks ergibt sich durch eine Kombination des „Quartalskennzeichens“ (je vier aufeinanderfolgende Blöcke werden mit einer zufälligen Permutation der Zahlen 1 bis 4 versehen) und des Wochenkennzeichens (die Auswahlbezirke werden mit einer zufälligen Permutation der Zahlen 1 bis 13 versehen).

Neubauauswahl:

Bei der Neubauauswahl, die in einer eigenen Anschriftengrößenklasse abgegrenzt ist, gibt es gegenüber der Grundauswahl im Stichprobenplan einige Modifikationen. Während ein Auswahlbezirk der Grundauswahl (ohne Gemeinschaftsunterkünfte) durchschnittlich neun

Wohnungen umfasst, ist die Richtgröße bei der Neubausauswahl sechs Wohnungen. Für Auswahlbezirke der Anschriftengrößenklasse mit Gemeinschaftsunterkünften beträgt die Richtgröße 15 Personen.

#### **1.2.4 Erhebungs- und Analyseeinheiten**

Erhebungseinheiten sind Personen (in Privathaushalten und Gemeinschaftsunterkünften), Haushalte und Wohnungen.

Analyseeinheiten sind Personen, Lebensgemeinschaften, Familien, Haushalte und Wohnungen.

Bei der Zuordnung von Personen und Haushalten ist zu beachten, dass Personen mit mehreren Wohnsitzen am Ort ihrer Haupt- und ihrer Nebenwohnung erfasst werden können. Um Mehrfachzählungen zu vermeiden, werden in Veröffentlichungen üblicherweise nur Personen am Hauptwohnsitz ausgewählt.

#### **1.2.5 Erhebungsprogramm und Erhebungsverfahren**

##### **Erhebungsprogramm**

Das Fragenprogramm des Mikrozensus besteht aus einem gleichbleibenden Grundprogramm, das sozioökonomische Grundinformationen liefert und die laufende Beobachtung des Arbeitsmarktes ermöglicht. In den vierjährigen Zusatzprogrammen, die seit 2005 mit dem vollen Auswahlsatz von einem Prozent durchgeführt werden, werden weitere Angaben erfragt. Daneben werden im Rahmen des EU Labour Force Survey jährlich wechselnde Ad-hoc-Module eingesetzt und weitere Fragen zur Erwerbstätigkeit (sog. Strukturvariablen) gestellt. Die Ad-hoc-Module werden mit einem Auswahlsatz von 0,1 % erhoben. Die Strukturvariablen sind ab 2012 nicht mehr an die Ad-hoc-Module gekoppelt, sondern werden mit einem Auswahlsatz von 1 % erhoben.

##### **Grundprogramm:**

Im Grundprogramm sind im Mikrozensus 2016 folgende Themenbereiche enthalten: Merkmale der Person, Familien- und Haushaltszusammenhang, Haupt- und Nebenwohnung, Aufenthaltsdauer, Staatsangehörigkeit, Wohn-/ Lebensgemeinschaft (freiwillig), Fragen für eingebürgerte Personen, Fragen für Ausländer, Quellen des Lebensunterhalts, Höhe des Einkommens, Angaben zur Renten- und Pflegeversicherung, Besuch von Schule und Hochschule, allgemeiner Schul- und beruflicher Ausbildungsabschluss, allgemeine und berufliche Weiterbildung, Erwerbstätigkeit, Arbeitslosigkeit und Arbeitssuche, frühere Erwerbstätigkeit, Situation ein Jahr vor der Erhebung (freiwillig).

##### **Zusatzprogramme:**

Die Zusatzprogramme 2016 umfassen Angaben zu Pendlerströmen für Schülerinnen und Schüler, Studierende und Erwerbstätige sowie Angaben zu der Zahl lebend geborener Kinder für Frauen im Alter zwischen 15 und 75 Jahren. Die Erteilung der Auskunft in den Zusatzprogrammen war freiwillig. (Für Auskunftspflichtige, die gemäß Stichprobenplan 2015 hätten befragt werden sollen, aber erst 2016 mit dem Fragebogen 2015 befragt wurden, liegen keine Angaben zu den Zusatzprogrammen vor. Zudem liegen für diese Befragten-gruppe keine Angaben für 2016 neu aufgenommene Fragen des Grundprogramms vor.)

##### **Ad-hoc-Modul, AKE-Strukturvariablen:**

Im Rahmen der in den Mikrozensus integrierten Arbeitskräfteerhebung (AKE) werden normalerweise jährlich Ad-hoc-Module durchgeführt, die dazu dienen, detaillierte Informationen zu politisch relevanten Themen der Europäischen Union (EU) zu sammeln, welche nicht Teil des Standardfrageprogramms der Arbeitskräfteerhebung sind. Im Jahr 2016 wurde das Ad-hoc-Modul zum „Arbeitsmarkteintritt junger Menschen“ durchgeführt. Zielpopulation des Ad-

hoc-Moduls waren 15- bis 34-Jährige. Der Auswahlsatz betrug 0,1 % und die Erteilung der Auskunft war freiwillig. Die AKE-Strukturvariablen (Arbeitsvertrag: Zeitarbeitsvermittlungsfirma (EF177), Arbeitsvermittlung: Agentur für Arbeit (EF178), Gründe für Nichtarbeitsuche oder Teilzeitbeschäftigung (EF179u1 bis EF179u5) und Arbeitsplatzanforderungen: Leitungstätigkeit (EF116u1 bis EF116u6)), die bis einschließlich 2011 nur mit einem Auswahlsatz von 0,1 % erhoben wurden, werden ab 2012 mit dem vollen Auswahlsatz von 1 % ermittelt. Die Fragen sind freiwillig zu beantworten.

## Erhebungsverfahren

I. d. R. erfolgt eine 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<sup>3</sup> oder telefonisch.

Zulässig sind außerdem Proxy-Interviews, d. h., ein erwachsenes Haushaltsmitglied darf stellvertretend für andere Haushaltsmitglieder antworten.

### 1.2.6 Hochrechnung / GewichtungsvARIABLEN

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 2016 2,9 % der Haushalte (ohne Jahresüberhänge; s. u.). Bei Pflichtfragen liegt die Ausfallquote (Item-Nonresponse) i. d. R. unter 3 %.

(2) Anschließend werden die mit dem Kompensationsfaktor für Unit-Nonresponse gewichteten Stichprobenwerte an Populationswerte aus der laufenden Bevölkerungsfortschreibung und dem Ausländerzentralregister nach Altersklassen, Staatsangehörigkeitsgruppen und Geschlecht angepasst.<sup>4</sup> (Die Hochrechnung basiert auf den Ergebnissen der laufenden Bevölkerungsfortschreibung mit dem Stand der Aktualisierung durch den Zensus 2011.)

Die Anpassung erfolgt bei den Standardhochrechnungsfaktoren (EF951 und EF952) quartalsweise auf unterschiedlichen regionalen Ebenen. Alle Personen in einem Haushalt erhalten den gleichen Gewichtungsfaktor. Mit Ausnahme des Hochrechnungsfaktors für Merkmale des Ad-hoc-Moduls (EF956), liegen für alle Merkmalsbereiche gleiche Hochrechnungsfaktoren vor. Die Anpassung des Hochrechnungsfaktors für das Ad-hoc-Modul (EF956) ist auf der Personenebene vorgenommen worden, da (1) Auswertungen im Haushaltzusammenhang mit dem Ad-hoc-Modul von untergeordneter Bedeutung sind und (2) Personenfaktoren die Qualität der Hochrechnung verbessern. Bei der Konstruktion des für die Merkmale des Ad-hoc-Moduls verwendeten Hochrechnungsfaktors EF956 wurden 2016 für die Schätzung der Antwortwahrscheinlichkeiten die Erhebungsmethode (EF7), der Erwerbstyp (EF29) differenziert nach Geschlecht (EF46) und Altersklassen (EF44) auf Bundesebene, der Erwerbstyp (EF29) nach Geschlecht (EF46) auf der Landesebene, Altersklassen (EF44) ebenfalls auf der Landesebene sowie die zwei Nationalitätengruppen Deutsche/Ausländer in den Gebieten Ost/West eingesetzt.

---

<sup>3</sup> Die CAPI-Fragen unterscheiden sich teilweise von den Fragen im Selbstausfüllerbogen.

<sup>4</sup> Für Zeit- und Berufssoldaten einschließlich Bundes- und Bereitschaftspolizei sowie Personen im freiwilligen Wehrdienst stammen die Eckwerte aus Bestandsdaten des Bundesministeriums der Verteidigung, des Bundesministeriums des Innern, für Bau und Heimat bzw. der Bundespolizei.

## 2 Kurzbeschreibung des Scientific Use Files Mikrozensus 2016

Das SUF ist eine faktisch anonymisierte 70 %-Substichprobe des Mikrozensus.

### 2.1 Auswahl der Substichprobe

Auswahleinheiten waren bis 2011 Haushalte bzw. Wohnungen des Mikrozensus, wobei alle Personen eines ausgewählten Haushalts bzw. einer Wohnung in der Substichprobe enthalten waren. Wohnungen wurden in den Erhebungsjahren gezogen, in denen das Zusatzprogramm zur Wohnsituation (1998, 2002, 2006, 2010, 2014) durchgeführt wurde. Ab 2012 werden als Auswahleinheiten für die Substichprobe die Auswahlbezirke innerhalb eines Rotationsviertels herangezogen. Damit wird die Möglichkeit geschaffen selbstständig mit den Scientific Use Files Paneldatensätze zu erzeugen. Infolge der Erneuerung der gesamten Mikrozensusstichprobe im Jahr 2016, sind Verknüpfungen der Querschnittsdaten ab dem Erhebungszeitpunkt 2012 nur bis einschließlich der Erhebung 2015 möglich. Das für den SUF 2012 bis 2015 verwendete Verfahren kann weitestgehend auch für die neue Stichprobe des SUF 2016 übernommen werden. Es wird folgende Auswahltechnik angewendet:

Sortierung der Auswahlbezirke nach

1. Kennung des Rotationsviertels
2. Bundesland
3. Größenklasse (Dezile) der Auswahlbezirke
4. Kennung über Grundauswahl/Aktualisierung der Auswahlbezirke
5. Kennung über den Einbezug eines Auswahlbezirks in das Ad-hoc-Modul
6. Regierungsbezirk
7. Regionale Anpassungsschicht
8. Regionale Schicht
9. Regionale Untergruppe
10. Gemeindegrößenklasse
11. Nummer des Auswahlbezirks

Die Sortierung folgt damit weitgehend der Schichtung des Originalfiles. Ab 2012 wurden in die Sortierung neu aufgenommen die Größenklasse (Dezile) der Auswahlbezirke sowie die Kennung über den Einbezug eines Auswahlbezirks in das Ad-hoc-Modul.

Zwar wurde bei der Grundauswahl bzw. wird bei der Aktualisierung der Auswahlbezirke eine einheitliche Größe der jeweiligen Anschriftengrößenklassen angestrebt, sie unterscheiden sich jedoch fluktuationsbedingt erheblich. Um eine daraus resultierende Erhöhung der Fehlervarianz zu begrenzen, erfolgt die zusätzliche Aufnahme einer Sortierung nach der Anzahl der befragten Personen in einem Auswahlbezirk.

Mit der Aufnahme der Kennung eines Auswahlbezirks in das Ad-hoc-Modul soll die Unterstichprobeneigenschaft des Ad-hoc-Moduls in der Substichprobenziehung kontrolliert werden.

Beide Merkmale wurden für so wichtig gehalten, dass sie in der Sortierung weit oben, an dritter bzw. fünfter Stelle platziert werden.

### Schichtung:

Jeweils zehn in der Reihenfolge der Sortierung aufeinanderfolgende Auswahlbezirke bilden eine Schicht. Eine heterogene Zusammensetzung der Schicht in Bezug auf die Sortiermerkmale ist ungünstig, da sie tendenziell mit einem höheren Varianzbeitrag verbunden ist. Im Hinblick auf die Analysemöglichkeit nach Bundesländern sowie die anzustrebende Homogenität der Auswahlbezirke innerhalb einer Schicht, die varianzreduzierend wirkt, werden bei Sortierwechseln von Rotationsviertel, Bundesland und Größenklassen (Dezile) der Auswahlbezirke unvollständige Schichten mit weniger als zehn Auswahlbezirken zugelassen. Die letzte Schicht kann ebenfalls weniger als zehn Auswahlbezirke umfassen.

### Auswahlsatz:

Im „Startjahr“ 2016 werden für vollständige Schichten mit Hilfe eines „Pseudozufallsgenerators“ sieben unterschiedliche Zahlen im Intervall  $\alpha$  mit  $(1 \leq \alpha \leq 10, \alpha \in \mathbb{N})$  gezogen. Für jede vollständige Schicht werden die Auswahlbezirke nach einer zufälligen Permutation ihrer Ordnungsnummer in der Schicht sortiert. In den so angeordneten Daten gilt jeder Auswahlbezirk mit gezogener Zufallszahl als ausgewählt. Der Stichprobenumfang jeder vollständigen Schicht beträgt 7.

Für die Auswahlbezirke unvollständiger Schichten werden gleichverteilte unabhängige Zufallszahlen  $z$  im Intervall  $(0, 1)$  gezogen. Es gelten diejenigen Auswahlbezirke als gezogen, deren Zufallszahl im Intervall  $(0 \leq z \leq 0,7)$  liegt. Die Anzahl gezogener Auswahlbezirke in unvollständigen Schichten ist zufällig, die Ziehungswahrscheinlichkeit von 7/10 bleibt aber erhalten.

Mit der so gezogenen einfachen Stichprobe ohne Zurücklegen werden gleiche Auswahlwahrscheinlichkeiten für alle Auswahlbezirke des SUF erreicht.

In den Folgejahren werden zunächst die Auswahlbezirke der Substichprobe des vorangegangenen Jahres ( $t-1$ ) übernommen, die auch in der Stichprobe des aktuellen Jahres ( $t$ ) vertreten sind. Für die Auswahlbezirke des im aktuellen Jahr neu hinzugekommenen Rotationsviertels sowie der Neubausauswahl wird die Auswahltechnik gleichermaßen eingesetzt. Die entsprechend gezogenen Auswahlbezirke werden anschließend mit den Auswahlbezirken der Substichprobe des Vorjahres zusammengeführt und bilden das SUF des aktuellen Jahres.

## 2.2 Hochrechnungs- / GewichtungsvARIABLEN im SUF

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 des Ad-hoc-Moduls (EF956). Die Hochrechnungsfaktoren rechnen auf 1.000 der Population hoch. Ihnen liegt der neue Hochrechnungsrahmen zugrunde, mit dem eine Anpassung an die laufende Bevölkerungsfortschreibung basierend auf dem Zensus 2011 vorgenommen wird.

Um eine verbesserte Anpassung der Verteilungen des SUF an die veröffentlichten Ergebnisse zu erreichen, wurden die mit den Hochrechnungsfaktoren der Originaldaten (EF951, EF952) gewichteten Ergebnisse des SUF nachträglich nach Geschlecht, Altersgruppen (unter 15 Jahre, 15 bis unter 45 Jahre, 45 bis unter 65 Jahre, 65 Jahre oder älter) sowie den Erwerbsstatus (erwerbstätig, nicht erwerbstätig) innerhalb eines jeden Bundeslandes an Verteilungen der gewichteten Originaldaten angepasst. Diese poststratifizierten Hochrechnungsfaktoren des SUF wurden anschließend so modifiziert, dass sie für alle Personen innerhalb eines Haushaltes gleich sind.

## 2.3 Anonymisierungsmaßnahmen, Regionalangaben

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. Im Originalfile ist die Gemeindegrößenklasse zum Stand 30.06.2016 umgesetzt. Im Scientific Use File wurde die Gemeindegrößenklasse mit Stand 31.12.2015 umgesetzt, um die Kompatibilität zu anderen Regionalklassifikationen zu wahren. Neben den aus Datenschutzgründen notwendigen Vergrößerungen können daher Verteilungsunterschiede zwischen Ergebnissen des Scientific Use Files und den Originaldaten methodisch mit dieser Anpassung zusammenhängen.

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. Die Auswahlbezirke des Mikrozensus können nach wie vor den früheren Gemeindeteilen West- und Ost-Berlin korrekt zugeordnet werden (Bömermann/Gruber 2010: 23). Im SUF kann mithilfe eines speziellen Codes der Variable Gemeindegrößenklasse weiterhin zwischen West- und Ost-Berlin (EF563 = 9 Berlin-Ost) unterschieden werden. Für die Anpassung des Mikrozensus an die laufende Bevölkerungsfortschreibung liegen nur Daten zu den neuen Gebietsabgrenzungen (12 statt 23 Bezirke) vor. Deshalb ist bei der Hochrechnung (Gewichtungsvariablen EF951 ff.) mit regionalen Unschärfen zu rechnen.

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 Staatsangehörigkeit werden so aggregiert, dass jede Gruppe von Staatsangehörigkeiten in der Grundgesamtheit wenigstens 50.000 Einwohner umfasst. Im Fall von Vergrößerungen wird im SUF in der Regel 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](#).

## 2.4 Die wichtigsten Veränderungen gegenüber 2015 in Stichworten

Auf eine detaillierte Darstellung der betreffenden Variablen und Fragen wird aus Platzgründen verzichtet. Im Folgenden werden nur die bedeutendsten Änderungen genannt. Ausführliche Informationen zu den Änderungen des Mikrozensus 2016 im Vergleich zum Mikrozensus 2015 (bzw. 2012 für die Zusatzprogramme) enthalten sowohl das [Datenhandbuch](#) als auch die Datenbeschreibungen im Mikrodaten-Informationssystem [MISSY](#). Die [Variablen-Zeitpunkte-Matrix](#) in MISSY gibt zusätzliche Hinweise zur zeitlichen Vergleichbarkeit von Variablen des Mikrozensus ab 1973.

- Neue Variablen des Zusatzprogramms Pendlerströme für Schülerinnen und Schüler, Studierende (EF297-EF303) sowie Erwerbstätige (EF191-EF194), die letztmals 2012 erhoben wurden, sind:
  - EF191 F45 Arbeitsstätte: Hinweg ab Wohnsitz
  - EF192 F46 Arbeitsstätte: Entfernung

- EF193 F47 Arbeitsstätte: Zeitaufwand Hinweg
- EF194 F48 Arbeitsstätte: Verkehrsmittel Hinweg
- EF297 F131 Schule/Hochschule: Hinweg ab Wohnsitz
- EF298 F130 Schule/Hochschule: Lage im selben Bundesland
- EF299 F129 Schule/Hochschule: Lage in der Wohnsitzgemeinde
- EF300 F130 Schule/Hochschule: Bundesland
- EF301 F132 Schule/Hochschule: Entfernung
- EF302 F133 Schule/Hochschule: Zeitaufwand Hinweg
- EF303 F134 Schule/Hochschule: Verkehrsmittel Hinweg

- Variable EF 195 neu aufgenommen.
- Neue Variablen des Zusatzprogramms zur Zahl geborener Kinder von Frauen im Alter von 15 bis 75 Jahren (EF618-EF621), die letztmals 2012 erhoben wurden, sind:
  - EF618 F11 Geburten
  - EF619 F12 Geburten: Anzahl
  - EF620 Geburten (Imputationskennung)
  - EF621 Geburten: Anzahl (Imputationskennung)

In diesem Zusatzprogramm wurden teilweise die Angaben zur Zahl geborener Kinder imputiert. Hinweise zur Imputation können dem [Skript zum Imputationsverfahren zu den Geburtenangaben](#) aus der Erhebung 2008 entnommen werden, in 2016 wurde das gleiche Imputationsverfahren angewendet.

- Im Vergleich des Ad-hoc-Moduls „Arbeitsmarkteintritt junger Menschen“ (EF1080 – EF1106) zu dem von 2009 gab es Änderungen in den Fragen und Antwortkategorien, sodass die Vergleichbarkeit eingeschränkt ist; siehe dazu Eurostat (2018).
- Eine Reihe von Typisierungen sind zur Vereinfachung der Datenaufbereitung seit 2015 nicht mehr enthalten. Der Bericht „Einführung in die eigenständige Erstellung von Typisierungen am Beispiel des Mikrozensus Scientific Use Files 2014“ (Börlin 2020) zeigt am Beispiel des MZ SUF 2014, wie diese Typisierungen mithilfe der in den Daten vorliegenden Angaben selbst erstellt werden können.
- Um die Umsetzung des Labour-Force-Konzepts der internationalen Arbeitsorganisation (ILO) zu verbessern, wurden in den vergangenen Jahren einige Fragen verändert. Auch wenn seit 2011 keine wesentlichen Änderungen mehr stattgefunden haben, sind auch 2016 teilweise weiterhin Abweichungen zwischen Ergebnissen des Mikrozensus und anderen arbeitsmarktstatistischen Daten festzustellen. Dies betrifft z. B. kleine und geringfügige Beschäftigungen. Insbesondere sind bei Vergleichen mit früheren Erhebungen des Mikrozensus methodisch bedingte Ursachen zu berücksichtigen (siehe Körner/Puch 2009; Körner/Marder-Puch 2015).

## 2.5 Möglichkeit der Längsschnittverknüpfung

Im Abschnitt 1.2.3 wird darauf hingewiesen, dass der Mikrozensus eine Wiederholungsbefragung mit teilweiser Überlappung der Erhebungseinheiten ist. Nach der Umstellung der Substichprobenziehung ab dem SUF2012 und mit der Bereitstellung längsschnittkonsistenter Ordnungsnummern besteht die Möglichkeit, die Querschnittsdaten bis einschließlich des

SUF 2015 zusammenzuführen. Eine Zusammenführung mit den Querschnittsdaten des SUF 2016 ist aufgrund der Erneuerung der kompletten Stichprobe im Jahr 2016 nicht möglich (die Querschnittsdaten des SUF 2016 können in Zukunft mit den Querschnittsdaten der SUF bis einschließlich 2019 zusammengeführt werden). Für leichtere Auswertungen im Längsschnitt sind seit 2015 folgende Identifikatoren enthalten: idpers (Längsschnittorientierter Personenidentifikator), idpersx (Querschnittsorientierter Personenidentifikator), idhh (Längsschnittorientierter Haushaltsidentifikator) und idhxx (Querschnittsorientierter Haushaltsidentifikator). Wie ein Kurzzeitpanel erstellt werden kann und was hierbei zu beachten ist, entnehmen Sie bitte dem Skript „[Möglichkeiten der Verknüpfung von Mikrozensus-Querschnittserhebungen ab 2012 zu Panels](#)“.

## 2.6 Zur Abgrenzung der Bevölkerungsbegriffe und zur Hochrechnung

Zum Vergleich mit den vom Statistischen Bundesamt veröffentlichten Zahlen sind die jeweiligen Abgrenzungen (siehe Anhang A2) sowie [Definitionen](#) und Hochrechnungsfaktoren zu berücksichtigen.

Die Hochrechnungsfaktoren EF951 „Hochrechnungsfaktor Quartal“ und EF952 „Hochrechnungsfaktor 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 basierend auf der laufenden Bevölkerungsfortschreibung nach dem Zensus 2011 hoch.

Für deskriptive Querschnittsauswertungen zum Erhebungsjahr werden Sätze nicht benötigt, die in EF30 sowie weiteren Variablen mit -7 „Auswahlbezirk ohne befragten Haushalt“ und -8 „Entfällt (leerstehende Wohnung, ausgefallene Privathaushalte und Gemeinschaftsunterkünfte)“ codiert sind. Bei der Berechnung des Stichprobenfehlers sind jedoch auch Auswahlbezirke ohne erhobene Personen, sogenannte Nullbezirke, zu berücksichtigen (siehe Statistisches Bundesamt 2016: Anhang B).

Infolge der kontinuierlichen Erhebung des Mikrozensus treten sogenannte Jahresüberhänge auf. Dies sind Haushalte, die in  $t-1$  (mit  $t$  = aktuelles Erhebungsjahr) hätten befragt werden sollen, in der Regel aufgrund schwerer Erreichbarkeit nicht befragt werden konnten und deren Befragung erst in  $t$  erfolgte. Je nach Befragungszeitpunkt durch einen Interviewer in  $t$  werden Haushalte der Jahresüberhänge mit dem Erhebungsprogramm des Vorjahres oder mit dem Erhebungsprogramm des aktuellen Jahres befragt. Ebenso zu den Jahresüberhängen zählen Selbstausfüller, die den schriftlichen Fragebogen aus  $t-1$  erst in  $t$  nach einem festgelegten Abschlussstermin für  $t-1$  an das zuständige Statistische Landesamt zurückschicken. Inhaltlich bedeutet dies, dass für Haushalte aus dem Jahresüberhang, die mit dem Erhebungsprogramm des Vorjahres befragt wurden, erhebungsbedingt keine Angaben zu dem aktuellen vierjährigen Zusatzprogramm vorliegen und ebenso keine Angaben zu den Fragen, die neu in das Erhebungsprogramm aufgenommen wurden. Im Scientific Use File werden bei Fragen des vierjährigen Zusatzprogramms bzw. bei neu in das Erhebungsprogramm aufgenommenen Fragen die zugehörigen Variablen mit einem eigenen Missing-Code (-6) gekennzeichnet. Im Originalfile werden diese Fälle der Kategorie „Ohne Angabe“ zugeordnet.

Im Fall von Jahresüberhängen (EF5b=1, 2) ist EF4 mit Haushalten des Jahresbestandes bzw. ohne Jahresüberhang (EF5b=0) identisch. Bei eigenen Abgrenzungen auf der Ebene von Haushalten, Lebensformen oder Familien muss für eine eindeutige Identifikation der Einheiten die folgende Sortierfolge verwendet werden: EF1, EF3, EF12, EF5b und EF4 (sowie nach Bedarf Nr. der Lebensform im Haushalt (EF27) oder Nr. der Familie im Haushalt (EF25)).

## **2.7 Zugang zum Mikrozensus Scientific Use File und Beratung**

Die Mikrozensus Scientific Use Files können von unabhängigen inländischen Forschungseinrichtungen gegen eine Bereitstellungsgebühr von 250 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.

### 3 Literaturhinweise

#### 3.1 Allgemein

- 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.](#)
- Bihler, W.; Zimmermann, D. (2016): Die neue Mikrozensusstichprobe ab 2016. [Wirtschaft und Statistik \(6\): 20-29.](#)
- Bömermann, H.; Gruber, G. (2010): Gebietsgliederungen in Berlin. [Zeitschrift für amtliche Statistik 5+6/2010: 22-25.](#)
- Börlin, S. (2020 – im Erscheinen): Einführung in die eigenständige Erstellung von Typisierungen am Beispiel des Mikrozensus Scientific Use Files 2014. [GESIS Papers 2020/01.](#)
- Bujard, M; Dorbritz, J.; Herter-Eschweiler, R.; Lux, L. (2015): [Das unterschätzte Potenzial hoher Fallzahlen – Stärken und Limitierungen des Mikrozensus am Beispiel von Fertilitätsanalysen.](#) Zeitschrift für Familienforschung, 27 (3): 343-372.
- Crößmann, A. (2016): [Methodeninformation. Auswirkungen des neuen Hochrechnungsrahmens im Mikrozensus - Analysen im Bereich „Arbeitsmarkt“.](#) Wiesbaden.
- Emmerling, D.; Riede, T. (1997): 40 Jahre Mikrozensus. [Wirtschaft und Statistik \(3\), S. 160-174.](#)
- Eurostat (2018): [Young people on the labour market. Evaluation of the 2016 EU labour force survey ad hoc module.](#) Luxembourg.
- Gauckler, B.; Körner, T. (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.
- 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.](#)
- Groß, F. (2006): Die Neuorganisation des Interviewereinsatzes im Mikrozensus – Erfahrungen nach dem ersten kontinuierlichen Erhebungsjahr. Statistische Hefte Mecklenburg-Vorpommern 1/2006: 12-19.
- Günther, L.; Körner, T. (2016): Erfassung der Langzeiterwerbslosigkeit in Mikrozensus und Arbeitskräfteerhebung. [Wirtschaft und Statistik \(2\): 9-25.](#)
- Günther, L., & Gawronski, K (2018): [Junge Menschen auf dem Arbeitsmarkt. Ergebnisse des Ad-hoc-Moduls zur Arbeitskräfteerhebung 2016.](#) Wirtschaft und Statistik (1): 9-22.
- Hammes, W. (2012): Haushalte und Lebensformen der Bevölkerung. Ergebnisse des Mikrozensus 2011. [Wirtschaft und Statistik \(11/2012\): 977-989.](#)
- Hartmann, P. (1989): Der Mikrozensus als Datenquelle für die Sozialwissenschaften. [ZUMA-Nachrichten 24](#), S. 6-25.
- Hungnes Lien, H. (2016): [Migrants on the European labour market. Wirtschaft und Statistik – Sonderheft Arbeitsmarkt und Migration: 55-68.](#)

- 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.
- Herter-Eschweiler, R; Schimpl-Neimanns, B. (2018): [Möglichkeiten der Verknüpfung von Mikrozensus-Querschnittserhebungen ab 2012 zu Panels](#). Bonn / Mannheim.
- Herter-Eschweiler, R. (2019): [Der Mikrozensus und die Möglichkeiten seiner Regionalisierung](#). In: Behrens, K.; Böltken, F. ; Dittmar, H. ; Götttsche, F. ; Gutfleisch, R. ; Habla, H. ; ... Wiese, K. (2019). Regionale Standards: Ausgabe 2019. Köln: GESIS - Leibniz-Institut für Sozialwissenschaften, S. 205-225.
- Hochgürtel, T. (2013): Das künftige System der amtlichen Haushaltsstatistiken. [Wirtschaft und Statistik \(7\): 457-466](#).
- Hochgürtel, T. (2017): [Familiengründung und -erweiterung im Kohortenvergleich](#). Wirtschaft und Statistik(6): 60-73.
- Hochgürtel, T. (2018): [Realisierte Erwerbstätigkeit zur Messung des Vereinbarkeitsarrangement von Familie und Beruf](#), Wirtschaft und Statistik(1): 23-34.
- Höhne, J. (2016): Migrantinnen und Migranten auf dem deutschen Arbeitsmarkt. [Wirtschaft und Statistik – Sonderheft Arbeitsmarkt und Migration: 24-40](#).
- Hundenborn, J., & Enderer, J. (2019): [Die Neuregelung des Mikrozensus ab 2020](#). Wirtschaft und Statistik (6): 9-17.
- International Labour Office (2008): [The concepts of employment and unemployment as set out by the 13th ICLS - Is there a need for revision? Some remarks from a European perspective](#). Genf.
- Inter-university Consortium for Political and Social Research (ICPSR) (2012): [Guide to social science data preparation and archiving: best practice throughout the data life cycle \(5th ed.\)](#). Ann Arbor, MI.
- Iversen, K. (2007a): Das Mikrozensusgesetz 2005 und der Übergang zur Unterjährigkeit. [Wirtschaft und Statistik \(1\): 38-44](#).
- Iversen, K. (2007b): Auswirkungen der neuen Hochrechnung für den Mikrozensus ab 2005. [Wirtschaft und Statistik \(8\): 739-747](#).
- Keller, M.; Haustein, T. (2014): Vereinbarkeit von Familie und Beruf. Ergebnisse des Mikrozensus 2013. [Wirtschaft und Statistik \(12\): 733-753](#).
- 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.; Günther, L. (2011): Frauen in Führungspositionen. Ansatzpunkte zur Analyse von Führungskräften in Mikrozensus und Arbeitskräfteerhebung. [Wirtschaft und Statistik \(5\): 434-451](#).
- Körner, T.; Marder-Puch, K. (2015): Der Mikrozensus im Vergleich mit anderen Arbeitsmarktstatistiken. Ergebnisunterschiede und Hintergründe seit 2011. [Wirtschaft und Statistik \(4\): 39-53](#).
- Körner, T.; Puch, K. (2013): Wer sind die ausschließlich geringfügig Beschäftigten? Eine Analyse nach sozialer Lebenslage. [Wirtschaft und Statistik \(1\): 42-61](#).
- Körner, T.; Puch, K.; Frank, T.; Meinken, H. (2011): Geringfügige Beschäftigung in Mikrozensus und Beschäftigungsstatistik. Neue Erkenntnisse zu den Hintergründen der Ergebnisunterschiede. [Wirtschaft und Statistik \(11\): 1065-1085](#).

- Körner, T.; Puch, K. (2009): Der Mikrozensus im Kontext anderer Arbeitsmarktstatistiken - Ergebnisunterschiede und ihre Hintergründe. [Wirtschaft und Statistik \(6\): 528 - 552](#).
- Körner, T.; Puch, K. (2011): Coherence of German labour market statistics: the labour force survey in comparison with the employment accounts and the unemployment register. In: Statistisches Bundesamt (Hrsg.), [Statistik und Wissenschaft, Bd. 19](#). Wiesbaden.
- Körner, T.; Schüller, F.; Götsche, F. (2016): Arbeitsmarkt und Migration in der amtlichen Statistik. [Wirtschaft und Statistik – Sonderheft Arbeitsmarkt und Migration: 5-23](#).
- Krug, W.; Nourney, M.; Schmidt, J. (1994): Wirtschafts- und Sozialstatistik. Gewinnung von Daten. München: Oldenbourg (3. völlig Neubearb. Auflage).
- Lengerer, A.; Schroedter, J. H.; Boehle, M.; Hubert, T.; Wolf, C. (2008): Harmonisierung der Mikrozensus 1962 bis 2005. Version 2. [GESIS-Methodenbericht 2008/12](#). Mannheim.
- Lengerer, A.; Bohr, J.; Janßen, A. (2005): Haushalte, Familien und Lebensformen im Mikrozensus - Konzepte und Typisierungen. [ZUMA-Arbeitsbericht 2005/05](#).
- Liersch, A. (2014): Arbeitsunfälle und arbeitsbedingte Gesundheitsprobleme. Ergebnisse einer Zusatzerhebung im Rahmen des Mikrozensus 2013. [Wirtschaft und Statistik \(9\): 561-574](#).
- Lotze, S.; Breiholz, H. (2002a): Zum neuen Erhebungsdesign des Mikrozensus. Teil 1. [Wirtschaft und Statistik \(5\): 359-366](#).
- Lotze, S.; Breiholz, H. (2002b): Zum neuen Erhebungsdesign des Mikrozensus. Teil 2. [Wirtschaft und Statistik \(6\): 454-459](#).
- Lüttinger, P.; Riede, T. (1997): Der Mikrozensus: amtliche Daten für die Sozialforschung. [ZUMA-Nachrichten 41](#), 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.
- Müller, W.; Blien, U.; Knoche, P.; Wirth, H. (1991): [Die Faktische Anonymität von Mikrodaten, \(Band 19 der Schriftenreihe Forum der Bundesstatistik\)](#). Metzler-Poeschel, Stuttgart.
- Nöthen, M. (2005): Von der „traditionellen Familie“ zu „neuen Lebensformen“. Neuerungen in der Familienberichterstattung des Mikrozensus. [Wirtschaft und Statistik \(1\): 25-40](#).
- Pöttsch, O. (2018): Aktueller Geburtenanstieg und seine Potenziale. [Wirtschaft und Statistik \(3\): 72-89](#).
- Rengers, M.; Bringmann, J.; Holst, E. (2017): Arbeitszeiten und Arbeitszeitwünsche: Unterschiede zwischen Mikrozensus und SOEP. [Wirtschaft und Statistik \(4\): 11-43](#).
- Rengers, Martina (2016): Ungenutztes Arbeitskräftepotenzial in der stillen Reserve. Ergebnisse für das Jahr 2015. [Wirtschaft und Statistik, Heft 6, S. 30-51](#).
- Rengers, M. (2015): Unterbeschäftigung, Überbeschäftigung und Wunscharbeitszeiten in Deutschland. Ergebnisse für das Jahr 2014. [Wirtschaft und Statistik \(6\): 22-42](#).
- Rengers, M. (2012): Ungenutztes Arbeitskräftepotenzial in der Stillen Reserve. Ergebnisse für das Jahr 2010. [Wirtschaft und Statistik \(4\): 299-319](#).
- Rengers, M. (2004): Das international vereinbarte Labour-Force-Konzept. [Wirtschaft und Statistik\(12\): 1369–1383](#).

- Schäfer, T.; Brückner, G. (2009): 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 \(12/2008\): 1046-1066](#).
- Schimpl-Neimanns, B. (2019): Varianzschätzung von Nettoveränderungen mit dem Mikrozensus ab 2012. Wirtschafts- und Sozialstatistisches Archiv 13(1): 73-85. DOI [10.1007/s11943-019-00240-0](#).
- Schimpl-Neimanns, B. (2011): Schätzung des Stichprobenfehlers in Mikrozensus Scientific Use Files ab 2005. AStA Wirtschafts- und Sozialstatistisches Archiv 5 (1): 19-38. DOI: <http://dx.doi.org/10.1007/s11943-011-0092-4> (Preprint).[siehe auch [Mikrodaten-Tools](#)]
- Schmidt, S. (2000): Erwerbstätigkeit im Mikrozensus. Konzepte, Definition, Umsetzung. [ZUMA-Arbeitsbericht 00/01](#).
- Statistische Ämter des Bundes und der Länder (2015): [Informationen zum Mikrozensus \[Informationen für die Befragten Mikrozensus 2016\]](#). Wiesbaden.
- Statistische Ämter des Bundes und der Länder (2014): [Dokumentation der Abweichungen zwischen Selbstausfüller und Blaise-Programm](#).
- Statistisches Bundesamt (ohne Jahr): [Was beschreibt die Internationale Arbeitsorganisation \(ILO\)-Arbeitsmarktstatistik?](#)
- Statistisches Bundesamt (1989a): [Methodische Fragen bevölkerungsstatistischer Stichproben am Beispiel des Mikrozensus](#). Bericht zur Konferenz vom 10. und 11. Juni 1988. Heft 9 der Schriftenreihe Ausgewählte Arbeitsunterlagen zur Bundesstatistik. Wiesbaden.
- Statistisches Bundesamt (1989b): [Inhaltliche Fragen bevölkerungsstatistischer Stichproben am Beispiel des Mikrozensus](#). Bericht zur Konferenz vom 21. und 22. Oktober 1988. Heft 10 der Schriftenreihe Ausgewählte Arbeitsunterlagen zur Bundesstatistik. Wiesbaden.
- Statistisches Bundesamt (2004): Arbeitsunterlagen zum Mikrozensus. Das Erhebungsprogramm des Mikrozensus seit 1957. Bonn.
- Statistisches Bundesamt (2017): [Qualitätsbericht Mikrozensus 2016](#). Wiesbaden.
- Statistisches Bundesamt (2012a): [Methodeninformation. Mikrozensus und Arbeitskräfteerhebung: Ergebnisse zur Erwerbstätigkeit ab dem Jahr 2011](#). Wiesbaden.
- Statistisches Bundesamt (2012b): [Mikrozensus: Haushaltszahlen ab 2005](#). Wiesbaden. ([https://www.destatis.de/DE/Methoden/MikrozensusHaushaltszahlen.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Methoden/MikrozensusHaushaltszahlen.pdf?__blob=publicationFile); letzter Zugriff: 26.10.18; Originaldatei nicht mehr vorhanden)
- Statistisches Bundesamt (2016a): [Anonymisierungskonzept zu den Berufsangaben der Klassifikation der Berufe 2010 ab den Mikrozensus Scientific Use Files 2012](#). Bonn.
- Statistisches Bundesamt (2016b): [Zum Umsteigeschlüssel von der Klassifikation der Berufe in der Ausgabe von 2010 zur internationalen Standardklassifikation der Berufe in der Ausgabe von 2008 in der Anwendung für den Mikrozensus](#). Bonn.
- Weinmann, J. (2013): Kind und Beruf: Nicht alle Mütter wollen beides. [STATmagazin: Bevölkerung](#).
- Wirth, H. (1992): Die faktische Anonymität von Mikrodaten: Ergebnisse und Konsequenzen eines Forschungsprojektes. [ZUMA-Nachrichten Nr. 30](#), S. 7-65.

- Wolf, C.; Lüttinger, P. (2009): Proxy-Angaben im deutschen Mikrozensus, S. 395-412 in: Bacher, J.; Wolf, C. (Hrsg): [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**

### **3.2 Ergebnisse des Mikrozensus 2016**

- Statistisches Bundesamt (2016): [Qualitätsbericht Mikrozensus 2016](#). Wiesbaden.
- Statistisches Bundesamt (2017): [Fachserie 1 Reihe 3, Bevölkerung und Erwerbstätigkeit, Haushalte und Familien](#), Ergebnisse des Mikrozensus 2016. Wiesbaden.
- Statistisches Bundesamt (2017): [Fachserie 1 Reihe 2.2, Bevölkerung und Erwerbstätigkeit, Bevölkerung mit Migrationshintergrund](#), Ergebnisse des Mikrozensus 2016. Wiesbaden.
- Statistisches Bundesamt (2017): [Fachserie 1 Reihe 4.1, Bevölkerung und Erwerbstätigkeit, Erwerbsbeteiligung der Bevölkerung, Ergebnisse des Mikrozensus zum Arbeitsmarkt](#), Mikrozensus 2016. Wiesbaden. Wiesbaden.
- Statistisches Bundesamt (2017): [Kinderlosigkeit, Geburten und Familien](#), Ergebnisse des Mikrozensus 2016. Wiesbaden.

Unter [Themen](#) kann auf weitere kostenlose Publikationen aus dem Bereich Mikrozensus zugegriffen werden. Des Weiteren können aus der Datenbank Genesis-Online kostenlos [Tabellen](#) abgerufen werden.

### **3.3 Mikrozensusgesetz, EU-Verordnungen**

#### **3.3.1 Mikrozensusgesetz 2005**

- Mikrozensusgesetz 2005 (MZG 2005) vom 24. Juni 2004 (BGBl. I S. 1350), zuletzt geändert durch Artikel 1 des Gesetzes vom 2. Dezember 2014 (BGBl. I S. 1926). [\[Download\]](#).

#### **3.3.2 EU-Verordnungen zur Arbeitskräfteerhebung**

- Verordnung (EG) Nr. 577/98 des Rates vom 9. März 1998 zur Durchführung einer Stichprobenerhebung über Arbeitskräfte in der Gemeinschaft (ABl. L 77 vom 14. März 1998, S. 3), zuletzt geändert durch Verordnung (EG) Nr. 545/2014 des Europäischen Parlaments und des Rates vom 15. Mai 2014 (ABl. L 163 vom 29. Mai 2014, S. 10). [\[Download\]](#).

## A Anhang

### A1 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 2016 dienen zum Einlesen des Rohdatenmaterials und zum Erstellen von Systemfiles für die Programmpakete SPSS, SAS und Stata. Sie beinhalten Programmanweisungen zur Definition von fehlenden Werten sowie zum Versehen der Variablen und ihrer Ausprägungen mit entsprechenden Labels. Die Setups werden im Dateiformat PC, Dos/Windows angeboten. Eine Umsetzung auf das Dateiformat Unix kann zum Beispiel mit Notepad++ oder Textpad++ durchgeführt werden.

#### **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.

IBM SPSS Statistics 24.0.0.1	Stata/MP 15.1 for Windows	SAS 9.4 for Windows
	<code>local log_file "LOG_FILENAME.log" ;</code>	
<code>file handle raw_data / name = "RAW_DATA_FILE- NAME" / lrec1 = 1012.</code>	<code>local raw_data_file "RAW_DATA_FILENAME" ;</code>	<code>%let raw_data_filename = 'RAW_DATA_FILENAME' ;</code>
		<code>%let sas_file_dirname = 'SAS_FILE_DIREC- TORY_PATH' ;</code>
<code>file handle spss_refer- ence / name = "SPSS_FILENAME.sav"</code>	<code>local stata_file "STATA_FILENAME.dta"  ;</code>	<code>%let sas_file_basename = SAS_FILENAME ;</code>

- **Stata:** Das Setup steht mit der Zeichencodierung Unicode (UTF-8) bereit und ist mit Stata ab Version 14 ablauffähig. Für die Verwendung mit älteren Versionen kann es z. B. mithilfe von MS-Edit in die Zeichencodierung ANSI bzw. Windows-1252 umgesetzt werden.
- **SPSS:** Das Setup mit der Zeichencodierung Windows-1252 ist sowohl mit Version 24 als auch mit älteren Versionen ablauffähig, wenn entsprechend unter „Bearbeiten“ | „Optionen“ | „Sprache“ | „Zeichencodierung ...“ | „[x] Schriftsystem der Ländereinstellung ...“ eingestellt ist.
- **Missing Values:** In SPSS können vorliegende Werte als benutzerdefinierte Missings deklariert werden. Sie werden bei Auszählungen mit entsprechenden Werten und Labels ausgewiesen, zählen aber i. d. R. bei statistischen Modellen nicht als gültige Werte. In SAS und Stata können dagegen vorliegende Werte nur nach Recodierung als fehlende

Werte definiert werden. Im SPSS-Setup werden benutzerdefinierte Missings spezifiziert. Diese Definitionen sind zwar auch in den SAS- und Stata-Setups enthalten, im Unterschied zum SPSS-Setup jedoch auskommentiert, da sonst die Originalwerte der Rohdaten (-1, ..., -9) im Systemfile durch benutzerdefinierte Missing-Zeichen (.a, ..., .h) ersetzt würden. Bei Bedarf können diese im Setup unter dem Kommentar „Definition of usermissing 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 ausführen zu lassen.

IBM SPSS Statistics 24.0.0.1	Stata/MP 15.1for Windows	SAS 9.4 for Windows
Set decimal dot. * ggf. ändern in "... yes." oder "... no." und Kommentarzeichen * löschen. * set unicode no. set olang = GERMAN.	set more off set dp comma #delimiter ;	options compress = binary ;

Ist das entsprechende Systemfile erstellt, kann die einfache Fallzahl  $n = 546.785$  (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 2016 sind im [GESIS Web](#) zu finden.

Die Datei enthält insgesamt 26.465 Sätze, die in EF30 sowie weiteren Variablen mit -7 „Entfällt (Auswahlbezirk ohne befragten Haushalt)“ und -8 „Entfällt (Leerstehende Wohnung, ausgefallene Privathaushalte und Gemeinschaftsunterkünfte)“ codiert sind. Diese Sätze werden für Querschnittsauswertungen zum Erhebungsjahr nicht benötigt. Hierfür stehen  $n = 520.320$  ( $= 546.785 - 26.465$ ) Sätze zur Verfügung.

## A2 Verteilungen zur Prüfung der Systemfiles bzw. des Ablaufs der Setups

<b>Merkmal</b>	<b>Selektion</b>	<b>Hochrechnung</b>	<b>in 1.000</b>
<b>Bevölkerungskonzepte</b>			
Bevölkerung am Haupt- und Nebenwohnsitz (ungewichtete Fallzahl (EF30>0): n = 520.320)	EF30>=1 & EF30<=3	EF952	83.375
Bevölkerung am Hauptwohnsitz	EF30>=1 & EF30<=2	EF952	82.425
Bevölkerung am Hauptwohnsitz, Erwerbstätige	EF30>=1 & EF30<=2 & EF29=1	EF952	41.334
Bevölkerung in Privathaushalten	EF31=1	EF952	82.389
Privathaushalte	EF662>=1 & EF662<=4 & EF31=1	EF952	41.019
Bevölkerung in Familien/Lebensformen am Hauptwohnsitz (Lebensformenkonzept)	EF761>=1 & EF761<=2	EF952	81.583
Lebensformen am Hauptwohnsitz (Lebensformenkonzept)	EF761>=1 & EF761<=2 & EF762=1	EF952	42.023
Bevölkerung am Familienwohnsitz (trad. Familienkonzept <sup>5</sup> )	EF863=1	EF952	81.592
Familien (trad. Familienkonzept)	EF863=1 & EF864=1	EF952	44.996
<b>Mikrozensus-Grundprogramm</b>			
<b>Stellung im Beruf (EF117) x Erwerbstyp (EF29)<sup>6</sup></b>	EF30>=1 & EF30<=2	EF952	
Erwerbstätige	EF29=1		
Selbstständige			4.131
Mithelfende Familienangehörige			159
Beamte (inkl. Zeit-/ Berufssoldaten (inkl. BGS und Bereitschaftspolizei), Beamtenanwärter, freiwillig Wehrdienstleistende)			2.043
Angestellte (inkl. geringfügig beschäftigte Personen, Personen im Freiwilligendienst)			25.844
Arbeiter			7.721
Auszubildende (inkl. Volontäre, Praktikanten)			1.484
Erwerbstätige insgesamt			41.334
Erwerbslose	EF29=2		1.784
Erwerbspersonen	EF29>=1 & EF29<=2		43.118
Nichterwerbspersonen	EF29>2		39.306
Insgesamt			82.425

<sup>5</sup> Für die bis einschließlich Mikrozensus 2004 verwendeten traditionellen Familienkonzepte werden in den aktuellen Fachserien keine Verteilungen veröffentlicht.

<sup>6</sup> Siehe Statistisches Bundesamt (2017), Fachserie 1 Reihe 4.1, Tab. 1.1.

<b>Merkmal</b>	<b>Selektion</b>	<b>Hochrechnung</b>	<b>in 1.000</b>
<b>Privathaushalte: Haushaltsgröße (EF20)<sup>7</sup></b>		EF952	
1 Person	EF31=1 & EF662>=1 & EF662<=4		16.857
2 Personen			13.964
3 Personen			5.032
4 Personen			3.803
5 Personen und mehr			1.363
Haushalte insgesamt			41.019
<b>Migrationsstatus (differenz. Angabe) (EF2001)<sup>8</sup></b>	EF30>=1 & EF30<=2 & EF2001 <sup>9</sup>	EF952	
Personen ohne Migrationshintergrund	0-16		63.878
Personen mit Migrationshintergrund im engeren Sinn insgesamt	21-64		18.547
Zugewanderte (Personen mit eigener Migrations- erfahrung) insgesamt	21-31, 41		12.714
Ausländer	41		7.622
Deutsche	21-31		5.092
(Spät-)Aussiedler	21-29		3.185
Eingebürgerte	31		1.907
Personen ohne eigene Migrationserfahrung	32, 42-43, 52- 54, 62-64		5.796
Ausländer	42-43		1.339
Deutsche	32, 52-54, 62- 64		4.457
Eingebürgerte	32		508
Als Deutsche Geborene	52-54, 62-64		3.949
mit beidseitigem Migrationshintergrund	62-64		2.099
mit einseitigem Migrationshintergrund	52-54		1.850
Bevölkerung insgesamt			82.425

<sup>7</sup> Siehe Statistisches Bundesamt (2017), Fachserie 1 Reihe 3, Tab. 1.1.

<sup>8</sup> Siehe Statistisches Bundesamt (2017), Fachserie 1 Reihe 2.2, Tab. 2.

<sup>9</sup> Der Übersichtlichkeit halber sind nachfolgend nur die Wertebereiche von EF2001 genannt.

<b>Merkmal</b>	<b>Selektion</b>	<b>Hochrechnung</b>	<b>in 1.000</b>
<b>Mikrozensus-Zusatzprogramm Pendlereigen-schaft, -merkmale</b>	EF30>=1 & EF30<=2 & EF29=1 &		
<b>Arbeitsstätte: Verkehrsmittel (EF194)<sup>10</sup></b>	EF44>=15 &	EF952	
Erwerbstätige Bevölkerung am Hauptwohnsitz, 15 Jahre und älter, Pendeln zur Arbeitsstätte von der hiesigen Wohnung (Hauptwohnsitz), mit Angabe zur Lage der Arbeitsstätte	EF191=1 & EF195<>9 & EF195>0 & EF194>0		
Bus			1.363
U-Bahn, Straßenbahn			1.567
Eisenbahn, S-Bahn			1.506
PKW-Selbstfahrer			20.720
PKW-Mitfahrer			1.034
Motorrad, Moped, Mofa, Motorroller			268
Fahrrad			2.873
Zu Fuß			2.634
Sonstiges Verkehrsmittel			122
Elektrofahrrad (Pedelec, E-Bike)			18
Keine Angabe			7.346
Insgesamt			39.451
<b>Mikrozensus-Zusatzprogram Zahl der geborenen Kinder von Frauen im Alter von 15 bis 75 Jahren</b>	EF30>=1 & EF30<=2 & EF46=2 &		
<b>Zahl der geborenen Kinder (EF619)<sup>11</sup></b>	EF47>=1941 & EF47<=2001 &	EF952	
Freiwillige Beantwortung	EF5b=0 &		
1 Kind	((EF618=1   EF618=8)		6.211
2 Kinder	EF619=-5) &		8.469
3 Kinder	EF619<99		2.665
4 oder mehr Kinder			1.047
Zusammen			18.392
Ohne Kind			10.321
Insgesamt (mit Angaben zur Geburt von Kindern)			28.713
<b>EU-Labour Force Survey – Strukturmerkmale</b>			
<b>Arbeitsvertrag: Zeitarbeitsvermittlungsfirma (EF177), freiwillige Beantwortung – Arbeitnehmer, erwerbstätige Bevölkerung nach EU-Definition</b>	EF29=1 & EF32=1 & EF177>=1 & EF177<=9	EF952	
Ja			771
Nein			26.877
Keine Angabe			4.558
Insgesamt			32.206

<sup>10</sup> Siehe Statistisches Bundesamt (2017), Fachserie 1 Reihe 4.1, Tab. 9.1.1.

<sup>11</sup> Siehe Statistisches Bundesamt (2017), Kinderlosigkeit, Geburten und Familien – Ergebnisse des Mikrozensus 2016, Tab. 3.1.

Merkmal	Selektion	Hochrechnung	in 1.000
<b>EU-Labour Force Survey Ad-hoc-Modul Arbeitsmarkteintritt junger Menschen</b>			
<b>Bezahlte Tätigkeit während des Bildungsgangs zum höchsten Abschluss (EF1080)</b> 0,1%-Unterstichprobe, freiwillige Beantwortung - Bevölkerung nach EU-Definition, Personen im Alter von 15 bis 34 Jahren, die bereits einen Bildungsabschluss haben	EF6=2 & EF32=1 & EF44>=15 & EF44<=34 & EF1080>=1 & EF1080<=9	EF956	
Ja			4.192
Nein			11.606
Keine Angabe			650
Insgesamt			16.448

Als Referenz zur Plausibilitätsprüfung der Mikrozensus SUF dienen die in den Fachserien des Statistischen Bundesamtes veröffentlichten Ergebnisse des jeweiligen Mikrozensus. Für den Mikrozensus 2016 sind dies insbesondere die Fachserie 1 Bevölkerung und Erwerbstätigkeit: Reihe 3 (Haushalte und Familien), Reihe 4.1 (Bevölkerung und Erwerbstätigkeit) und Reihe 2.2 (Bevölkerung mit Migrationshintergrund). Bei diesen Vergleichen sind die von den Statistischen Ämtern verwendeten Bevölkerungs- und Erwerbskonzepte sowie die Hochrechnung zu beachten.

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 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 2016 wurden die Variablenlabels nach der 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; 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.