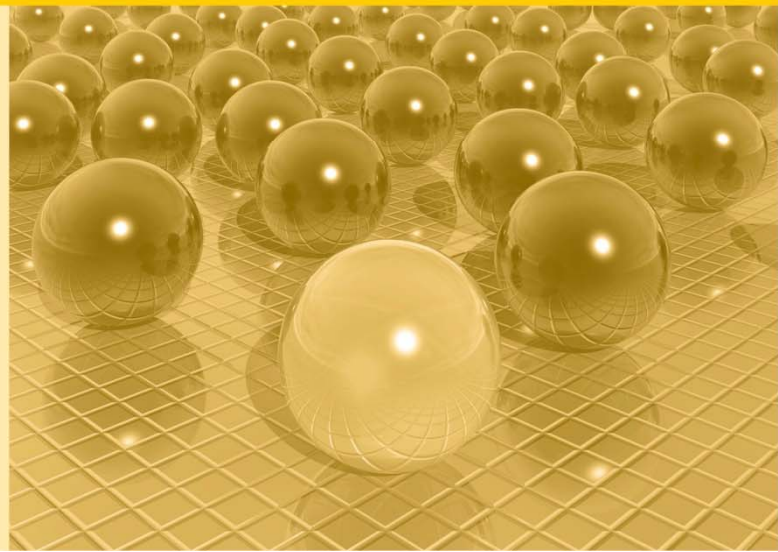


# Metadatenreport



Teil II: Produktspezifische Informationen zur Nutzung des Mikrozensus 2013 am  
Gastwissenschaftsarbetsplatz sowie per kontrollierter Datenfernverarbeitung  
(sog. On-Site-Zugangswege/-Nutzungen)

DOI: 10.21242/12211.2013.00.00.1.1.1

Version 2

## Impressum

Herausgeber: Statistische Ämter des Bundes und der Länder  
Herstellung: Information und Technik Nordrhein-Westfalen  
Telefon 0211 9449-01 • Telefax 0211 9449-8000  
Internet: [www.forschungsdatenzentrum.de](http://www.forschungsdatenzentrum.de)  
E-Mail: [forschungsdatenzentrum@it.nrw.de](mailto:forschungsdatenzentrum@it.nrw.de)

### Fachliche Informationen zu dieser Veröffentlichung:

Forschungsdatenzentrum der  
Statistischen Ämter der Länder  
– Düsseldorf –  
Tel.: 0211 9449-2877  
Fax: 0211 9449-8087  
[forschungsdatenzentrum@it.nrw.de](mailto:forschungsdatenzentrum@it.nrw.de)

### Informationen zum Datenangebot:

Statistisches Bundesamt  
Forschungsdatenzentrum  
Tel.: 0611 75-2420  
Fax: 0611 72-3915  
[forschungsdatenzentrum@destatis.de](mailto:forschungsdatenzentrum@destatis.de)

Forschungsdatenzentrum der  
Statistischen Ämter der Länder  
– Geschäftsstelle –  
Tel.: 0211 9449-2873  
Fax: 0211 9449-8087  
[forschungsdatenzentrum@it.nrw.de](mailto:forschungsdatenzentrum@it.nrw.de)

Erscheinungsfolge: unregelmäßig  
Erschienen im September 2020

Diese Publikation wird kostenlos als PDF-Datei zum Download unter [www.forschungsdatenzentrum.de](http://www.forschungsdatenzentrum.de) angeboten.

© Information und Technik Nordrhein-Westfalen, Düsseldorf, 2020  
(im Auftrag der Herausbergemeinschaft)

Vervielfältigung und Verbreitung, nur auszugsweise, mit Quellenangabe gestattet. Alle übrigen Rechte bleiben vorbehalten.

Fotorechte Umschlag: © artSILENCE –Fotolia.com

### Empfohlene Zitierung:

Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder: Metadatenreport. Teil II: Produktspezifische Informationen zur Nutzung des Mikrozensus 2013 am Gastwissenschaftsarbeitsplatz sowie per kontrollierter Datenfernverarbeitung (EVAS-Nummer: 12211, 12212, 12213). Version 2. DOI: 10.21242/12211.2013.00.00.1.1.1. Düsseldorf 2020.

# Metadatenreport

Teil II: Produktspezifische Informationen zur Nutzung des Mikrozensus 2013 am Gastwissenschaftsarbetsplatz sowie per kontrollierter Datenfernverarbeitung (sog. On-Site-Zugangswege/-Nutzungen)

DOI: 10.21242/12211.2013.00.00.1.1.1

Version 2



## Inhalt

<b>1</b>	<b>Datenaufbereitung in den FDZ.....</b>	<b>2</b>
1.1	Datenaufbereitung .....	2
1.2	Anonymisierungsmaßnahmen .....	3
1.3	Methodik der Verknüpfung.....	3
<b>2.</b>	<b>Produkt .....</b>	<b>4</b>
2.1	Merkmale und Merkmalsbeschreibung .....	4
2.1.1	Merkmalsdefinitionen .....	4
2.1.2	Datensatzbeschreibung / Schlüsselverzeichnis.....	4
2.2	Vergleichbarkeit der Merkmale über die Zeit .....	5
2.3	Eckwerte relevanter Merkmale und Merkmalskombinationen .....	6
2.4	Auswertbare regionale Ebene.....	7
2.5	Produktversionen .....	8
<b>3</b>	<b>Praktische Hinweise .....</b>	<b>9</b>
3.1	Hinweise zur Geheimhaltung .....	9
3.2	FAQ .....	11
3.3	Verfügbare Tools .....	12
<b>Anhang</b>	<b>.....</b>	<b>13</b>

# 1 Datenaufbereitung in den FDZ

## 1.1 Datenaufbereitung

Das Datenmaterial wird einer Vollzähligkeits- und Vollständigkeitskontrolle unterzogen, bei der geprüft wird, ob alle Erhebungsmerkmale und die dazugehörigen, typisierten Merkmale, die für das Mikrozensusjahr vorgesehen sind, belegt sind. Nicht belegte Merkmale werden aus dem Datenmaterial entfernt. Gleichzeitig wird darauf geachtet, dass die einzelnen Variablen die korrekten Formate aufweisen.

Im Zuge des Abgleichs mit den zentralen Veröffentlichungen<sup>1</sup> des Statistischen Bundesamts (siehe Metadatenreport Teil I, Kapitel 4 Zentrale Veröffentlichungen) werden die Hochrechnungsfaktoren so angeglichen, dass sie ohne zusätzliche Multiplikation auf die Gesamtbevölkerung hochrechnen<sup>2</sup>. Fälle, die keinen positiven Hochrechnungsfaktor besitzen, werden aus dem Datenmaterial entfernt. Dies betrifft ca. 9% der Fälle aus dem Originaldatenmaterial. In den betroffenen Fällen hat kein Interview stattgefunden. Es handelt sich beispielsweise um leerstehende Wohnungen, weggefallene Haushalte und Haushalte für die keine Auskunft vorliegt.

Die einzelnen Merkmale werden auf inhaltliche Konsistenz geprüft. Inkonsistenzen werden im Schlüsselverzeichnis (siehe auch Kapitel 2.1.2) dokumentiert und, sofern diese plausibel auflösbar sind, bereinigt. Dabei wird gewährleistet, dass die inhaltliche Aussage der Angaben nicht verändert wird.

Weiterhin wird geprüft, ob sich die Filterführung laut Fragebogen im Datenmaterial widerspiegelt. Für Personengruppen, denen gemäß Filterführung bestimmte Fragen nicht gestellt wurden, werden für die betroffenen Merkmale systemfehlende Werte vergeben (Ausprägung „leer“). Treten Inkonsistenzen bei der Filterführung auf, werden diese im Schlüsselverzeichnis (siehe auch Kapitel 2.1.2) dokumentiert und gegebenenfalls bereinigt, ohne die inhaltliche Aussage der Angaben zu verändern.

Zudem wird überprüft, ob unterschiedliche fachliche Ebenen, wie Haushalte, Familien und Lebensformen korrekt abgegrenzt und ausgewertet werden können.

Außerdem wird der in den Daten vorhandene Gebietsstand ermittelt. Die einzelnen Regionalkennungen (siehe Abschnitt 2.5 Auswertbare regionale Ebene) werden in diesem Zuge auf inhaltliche

---

<sup>1</sup> Exemplarisch finden Sie im Anhang dieses Reports Häufigkeitstabellen sowie die dazugehörige SPSS-Syntax mit Verweis auf die jeweilige Publikation des Statistischen Bundesamtes.

<sup>2</sup> Nähere Informationen zu den in diesem Produkt vorliegenden Hochrechnungsfaktoren finden Sie im Abschnitt 3.2 FAQ „Wann wird welcher Hochrechnungsfaktor verwendet?“.

Konsistenz geprüft. Treten Inkonsistenzen auf, werden diese dokumentiert und gegebenenfalls bereinigt.

## **1.2 Anonymisierungsmaßnahmen**

Direkte Identifikatoren und Hilfsmerkmale werden aus dem Datenmaterial entfernt.

Die Ordnungsnummern des Auswahlbezirks, des Haushalts im Auswahlbezirk und der Person im Haushalt werden systemfrei sortiert, so dass anhand der Position des einzelnen Falls im Datenmaterial kein Rückschluss auf einzelne Personen möglich ist.

Im Datenmaterial für die Gastwissenschaftsarbeitsplätze ist die Gemeindeebene für das Bundesland Bayern pseudonymisiert, so dass zwar die Information verfügbar ist, ob zwei Haushalte derselben Gemeinde angehören, nicht aber, um welche Gemeinde es sich handelt.

## **1.3 Methodik der Verknüpfung**

Der Mikrozensus stellt kein bereits verknüpftes Produkt dar, allerdings besteht die Möglichkeit, mit Hilfe der EF 1-4 und EF 63, auf Individualebene die Jahrgänge 2012-2015 miteinander zu verknüpfen<sup>3</sup>.

Da sich nach 2015 die Auswahlgrundlage der Stichprobe verändert hat, ist keine Verknüpfung der Jahre 2012-2015 mit nachfolgenden Erhebungsjahren möglich<sup>4</sup>..

---

<sup>3</sup> Informationen zur Verknüpfungsmethode und Mustersyntaxen sind bei Herter-Eschweiler und Schimpel-Neimanns 2018 zu finden.

<sup>4</sup> Genauere Informationen in Bihler/Zimmermann 2016 ([https://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2016/06/die-neue-mikrozensusstichprobe-062016.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2016/06/die-neue-mikrozensusstichprobe-062016.pdf?__blob=publicationFile)).

## 2. Produkt

### 2.1 Merkmale und Merkmalsbeschreibung

#### 2.1.1 Merkmalsdefinitionen

Folgende Klassifikationen finden in dem beschriebenen Produkt Anwendung:

- Klassifikation der Berufe, Ausgabe 2010 (KIdB 2010)  
<https://statistik.arbeitsagentur.de/DE/Navigation/Grundlagen/Klassifikationen/Klassifikation-der-Berufe/KIdB2010/Systematik-Verzeichnisse/Systematik-Verzeichnisse-Nav.html>
- Internationale Standardklassifikation der Berufe, Ausgabe 2008 (ISCO-08)  
<http://www.ilo.org/public/english/bureau/stat/isco/isco08/>
- Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008)  
[https://www.destatis.de/DE/Methoden/Klassifikationen/Gueter-Wirtschaftsklassifikationen/Downloads/klassifikation-wz-2008-3100100089004.pdf?\\_\\_blob=publicationFile&v=5](https://www.destatis.de/DE/Methoden/Klassifikationen/Gueter-Wirtschaftsklassifikationen/Downloads/klassifikation-wz-2008-3100100089004.pdf?__blob=publicationFile&v=5)
- Bildungsskala International Standard Classification of Education, Ausgabe 1997 (ISCED-1997)<sup>5</sup>  
<https://www.datenportal.bmbf.de/portal/de/G54.html>
- Staats- und Gebietssystematik  
[https://www.destatis.de/DE/Methoden/Klassifikationen/Staat-Gebietsystematik/Staatsangehoerigkeitsgebietsschluessel\\_pdf.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Methoden/Klassifikationen/Staat-Gebietsystematik/Staatsangehoerigkeitsgebietsschluessel_pdf.pdf?__blob=publicationFile)
- Amtlicher Gemeindeschlüssel (AGS)/ Regionalschlüssel (RS)  
<https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/inhalt.html>

#### 2.1.2 Datensatzbeschreibung / Schlüsselverzeichnis

Der Mikrozensus ist eine Haushalts- und Personenstatistik. Die vollständige Liste der im Datensatz vorhandenen Variablen finden sich im Schlüsselverzeichnis aufgelistet. Sofern Variablen das Ergebnis einer Frage des Fragebogens wiedergeben, ist die zugehörige Fragennummer in der Spalte „Fragebogennummer“ des Schlüsselverzeichnisses enthalten.

---

<sup>5</sup> Es ist nur die Ebene der ISCED-Klassifikation enthalten. Die Bildungsfelder sind entsprechend der nationalen Klassifikation „Hauptfachrichtungen“ klassifiziert.



Schlüsselverzeichnis

[http://www.forschungsdatenzentrum.de/bestand/mikrozensus/gwap\\_kdfv/2013/fdz\\_mz\\_2013\\_schlueselverzeichnis.pdf](http://www.forschungsdatenzentrum.de/bestand/mikrozensus/gwap_kdfv/2013/fdz_mz_2013_schlueselverzeichnis.pdf)

Erhebungsunterlagen

[https://www.forschungsdatenzentrum.de/sites/default/files/mz\\_2013\\_on-site\\_eu.zip](https://www.forschungsdatenzentrum.de/sites/default/files/mz_2013_on-site_eu.zip)

2013 wurde das Zusatzprogramm Gesundheit sowie betriebliche und private Altersvorsorge, Schichtarbeit“ sowie das Ad-Hoc-Modul „Arbeitsunfälle, Arbeit und Gesundheit“ abgefragt.

## 2.2 Vergleichbarkeit der Merkmale über die Zeit

Folgende Merkmale weisen im Vergleich zum Vorjahr Veränderungen auf:

Gegenwärtige (Haupt-)Erwerbstätigkeit

- EF129 (Vollzeit-/Teilzeittätigkeit (nach Korrektur)): Im Vergleich zu den Vorjahren wird nicht mehr an der 32-Stunden-Grenze, sondern an einen Übergang zwischen 25 bis 36 Stunden (normalerweise geleistete Arbeitszeit laut EF131) plausibilisiert. In dem Übergangsbereich können beide Ausprägungen zutreffend sein (die Originalangabe bleibt hier erhalten). Die Angaben werden vergleichbar mit den Vorjahren, wenn EF131 bei „31 Stunden und weniger“ = Teilzeit und „32 Stunden und mehr“ = Vollzeit dichotomisiert wird.
- EF167 (Mit wem haben Sie Ihren Ausbildungsvertrag abgeschlossen?): Die Ausprägungen wurden um die Hälfte reduziert, eine Vergleichbarkeit ist eingeschränkt gegeben.

Altersvorsorge (im Zusatzprogramm; Vergleich zu 2009)

- EF394 (Art der zusätzlichen betrieblichen Altersversorgung): Die Ausprägungen haben sich komplett verändert.

Fragen zur Gesundheit (im Zusatzprogramm; Vergleich zu 2009)

- EF465 (Ist für Sie eine Behinderung durch amtlichen Bescheid festgestellt worden? (freiwillige Beantwortung)): Unterscheidung nur noch zwischen „Ja“ und „Nein“, und nicht wie zuvor, auch nach der Herkunft des Bescheides bzw. dem Beantragungsstatus.

Regionalkennungen

- EF560UG2 (Untergruppe 2: Kreis): Die Kreisreform für Mecklenburg-Vorpommern aus dem Jahr 2011 ist in den Daten umgesetzt. Diese Änderung wirkt sich auch auf weitere Regionalkennungen aus, eine Vergleichbarkeit ist somit nur eingeschränkt gegeben.

## 2.3 Eckwerte relevanter Merkmale und Merkmalskombinationen

### Verteilung der Bevölkerung am Hauptwohnsitz

**Tab. 1: EF46NEU Geschlecht**

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	männlich	39.454.210	48,9	48,9	48,9
	weiblich	41.156.827	51,1	51,1	100,0
Gesamt		80.611.037	100,0	100,0	

**Tab. 2: EF49NEU Familienstand**

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	ledig	32.535.891	40,4	40,4	40,4
	verheiratet (inkl. eingetragener Lebenspartnerschaften)	36.748.979	45,6	45,6	85,9
	geschieden	11.326.167	14,1	14,1	100,0
	Gesamt	80.611.037	100,0	100,0	

### Privathaushalte am Haupt- und Nebenwohnsitz

**Tab. 3: EF20NEU Privathaushalte kategorisiert**

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Einpersonenhaushalte	16.176.318	40,5	40,5	40,5
	Zweipersonenhaushalte	13.747.639	34,4	34,4	74,9
	Dreipersonenhaushalte	4.988.642	12,5	12,5	87,4
	Vierpersonenhaushalte	3.687.954	9,2	9,2	96,7
	Fünf- und mehr Personenhaushalte	1.332.281	3,3	3,3	100,0
Gesamt		39.932.835	100,0	100,0	

**Tab. 4: EF1 Bundesland**

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Pro- zente
Gültig	Schleswig-Holstein	1.403.433	3,5	3,5	3,5
	Hamburg	965.099	2,4	2,4	5,9
	Niedersachsen	3.827.891	9,6	9,6	15,5
	Bremen	353.222	,9	,9	16,4
	Nordrhein-Westfalen	8.518.603	21,3	21,3	37,7
	Hessen	2.937.352	7,4	7,4	45,1
	Rheinland-Pfalz	1.887.461	4,7	4,7	49,8
	Baden-Württemberg	5.014.338	12,6	12,6	62,4
	Bayern	6.123.056	15,3	15,3	77,7
	Saarland	492.915	1,2	1,2	78,9
	Berlin	1.933.548	4,8	4,8	83,8
	Brandenburg	1.238.367	3,1	3,1	86,9
	Mecklenburg-Vorpommern	827.932	2,1	2,1	89,0
	Sachsen	2.146.093	5,4	5,4	94,3
	Sachsen-Anhalt	1.154.193	2,9	2,9	97,2
	Thüringen	1.109.333	2,8	2,8	100,0
	Gesamt	39.932.835	100,0	100,0	

Im Anhang finden Sie diese und weitere relevante Tabellen mit der Syntax für das Programm SPSS sowie allen Fundstellen in den Fachserien des Statistischen Bundesamtes.

## 2.4 Auswertbare regionale Ebene

- Bundesebene (NUTS 0)
- Landesebene (NUTS 1)
- Regierungsbezirksebene (NUTS 2)
- Kreisebene (ggf. eingeschränkte Aussagekraft in kleinen Landkreisen) (NUTS 3)
- Gemeindeebene (nur zu methodischen Zwecken, nicht zur Ergebnisdarstellung empfohlen<sup>6</sup>) (LAU)

<sup>6</sup> Es steht kein geeigneter Hochrechnungsfaktor zur Verfügung. Kleine Gemeinden können vollständig entfallen, da kein Auswahlbezirk in die Stichprobe gelangt ist oder der Klumpeneffekt hat besonders starke Auswirkungen.

- Anpassungsschichten (regionale Schicht und regionale Untergruppe)<sup>7</sup>
- Raumordnungsregionen
- Arbeitsmarktregionen
- Nielsegebiete
- Arbeitsagenturbezirke (ggf. eingeschränkte Aussagekraft durch geringe Zellbesetzungen)

## 2.5 Produktversionen

### 1.0.0

- Erstveröffentlichung

### 1.0.1

- Variablenlabel von EF6 wurde angepasst: Angabe „(Nur wenn kein Jahresüberhang)“ wurde gestrichen.

### 1.1.1

- Die Korrektur von EF94 bis EF94UG4 ist in der Aufbereitungssyntax der vorherigen Versionen nicht vollständig durchgelaufen. In der Anmerkungsliste war die Korrektur bereits beschrieben. Die Syntax wurde so angepasst, dass die Korrekturen von EF94 bis EF94UG4 wirksam werden.

### 1.2.1

- Personen mit EF117<4 fielen in EF442 komplett in die Kategorie 99 ("ohne Angabe"). Die betroffenen Fälle wurden als systemföhlend rekodiert.

### 1.3.2 (DOI 10-21242-12211-2013-00-00-1-1-0)

- AGS des Mikrozensus an den AGS angepasst (betroffen sind: Berlin, Hamburg, Aachen)
- EF313 EF313U1 und EF313U2 in den Datensatz aufgenommen
- Eineindeutige Benennung von Regionalen Anpassungsschichten (EF564), Regionalen Schichten (EF566) und Regionalen Untergruppen (EF568)
- FDZ1-Variable gebildet
- Hinweise auf Bayern-Pseudonymisierung im SV aufgenommen (EF560UG1)
- Standard-File-Hinweis entfernt

### 1.4.0 (DOI 10-21242-12211-2013-00-00-1-1-1)

- Ertüchtigung zur Längsschnittverknüpfbarkeit
- EF3, EF4 und EF63 zur Bildung einer jahresübergreifenden Personen-ID aufgenommen

---

<sup>7</sup> In Berlin und Hamburg ist eine Auswertung auf Bezirksebene möglich. In München kann eine Aufteilung nach innerstädtischen Bezirken und nach Bezirken in der Randlage der Stadt München vorgenommen werden. Die regionalen Untergruppen entsprechen häufig den Kreisen.

## **3 Praktische Hinweise**

### **3.1 Hinweise zur Geheimhaltung**

#### **3.1.1 Gesetzliche Grundlagen der statistischen Geheimhaltung**

Unter Geheimhaltung versteht man das Sicherstellen der absoluten Anonymität der Ergebnisse statistischer Analysen. Konkret bedeutet das, dass im Rahmen der Geheimhaltung sichergestellt wird, dass mit den veröffentlichten Ergebnissen keine Rückschlüsse auf einen Einzelfall (z. B. Person, Haushalt, Betrieb, Einrichtung) gezogen werden können. Statistische Geheimhaltung wird überall dort angewendet, wo statistische Ergebnisse oder Einzeldaten die geschützten Räume der amtlichen Statistik verlassen.

Die Geheimhaltung in der amtlichen Statistik ist in § 16 Bundesstatistikgesetz (BStatG) geregelt und beinhaltet, dass Einzelangaben über persönliche und sachliche Verhältnisse, die für eine Bundesstatistik angegeben werden, von den jeweils durchführenden statistischen Stellen geheim zu halten sind, soweit es keine anderslautenden Bestimmungen gibt. Dies wird auch als Statistikgeheimnis bezeichnet. Das Statistikgeheimnis verpflichtet die amtliche Statistik, die erhaltenen Informationen zu schützen, d. h. sie in einer Form zu anonymisieren, die keine Rückschlüsse mehr auf die betreffende Person und den dargelegten Sachverhalt enthält. Die Geheimhaltung ist auch im Hinblick auf die informationelle Selbstbestimmung von besonderem Interesse: Viele Erhebungen der amtlichen Statistik unterliegen der Auskunftspflicht, somit steht es den Befragten nicht frei selbst zu entscheiden, ob sie eine Information weitergeben möchten. Die amtliche Statistik muss deshalb sicherstellen, dass die erhobenen Daten keinem Befragten zugeordnet werden können.

Das BStatG sieht jedoch auch Fälle vor, in denen das Statistikgeheimnis nicht gilt. In § 16 BStatG sind die Ausnahmen von der Geheimhaltungspflicht dargelegt. Unter anderem wird dort festgelegt, unter welchen Umständen die Daten der amtlichen Statistik für die Wissenschaft zugänglich gemacht werden dürfen und welche Regeln dabei einzuhalten sind.

#### **3.1.2 Geheimhaltung von Ergebnissen**

Um die gesetzlich vorgeschriebene Geheimhaltung von Einzelfällen in den Daten sicherzustellen, müssen alle Ergebnisse, die am Gastwissenschaftsarbetsplatz oder per Kontrollierter Datenfernverarbeitung erzeugt werden, vor ihrer Freigabe an den Nutzer von den FDZ einer Geheimhaltungsprüfung unterzogen werden. Dabei stellen die FDZ sicher, dass die Ergebnisse absolut anonym sind

und eine Reidentifikation einzelner Befragter nach menschlichem Ermessen ausgeschlossen werden kann. Entsprechend handeln auch die Fachabteilungen der Statistischen Ämter vor der Veröffentlichung von Ergebnissen.

Zur Sicherstellung der Geheimhaltung wenden die FDZ verschiedene Geheimhaltungsregeln an, die individuell auf die jeweilige Statistik zugeschnitten sind. In der Broschüre „Regelungen zur Auswertung von Mikrodaten in den Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder“ werden die gebräuchlichsten Regeln zur primären Geheimhaltung dargestellt. Diese Regeln werden in den FDZ im Grunde auf alle Statistiken angewendet. Die Anlage dieser Broschüre enthält Informationen darüber, welche Geheimhaltungsregeln auf welche Statistiken anzuwenden sind.

FDZ-Broschüre „Regelung zur Auswertung von Mikrodaten“:

[https://www.forschungsdatenzentrum.de/sites/default/files/fdz\\_broschuere\\_regelungen\\_2019.pdf](https://www.forschungsdatenzentrum.de/sites/default/files/fdz_broschuere_regelungen_2019.pdf)

Die amtliche Statistik veröffentlicht aus dem Mikrozensus keine Ergebnisse, die hochgerechnet weniger als 5000 Fälle umfassen, daher erübrigt sich dort eine klassische Geheimhaltung.

### **3.1.3 Praktische Tipps zur Vermeidung von Geheimhaltungsfällen**

Treten in den erstellten Analysen Geheimhaltungsfälle auf, werden diese Werte von den FDZ zur Sicherstellung der Geheimhaltung durch ein Sperrmuster ersetzt. Gerade in Kreuztabellen entstehen so durch die notwendige Sekundärsperrung schnell viele Lücken in den Auswertungen. Da eine einmal zur Sekundärsperrung herangezogene Tabellenzelle auch in allen folgenden Analysen gesperrt werden muss (tabellenübergreifende Geheimhaltung) – auch, wenn es in der neu erstellten Tabelle nicht nötig wäre – ist es sinnvoll, bei jeder Ergebniserstellung darauf zu achten, dass möglichst keine Geheimhaltungsfälle erzeugt werden. Treten in einem Output Geheimhaltungsfälle auf, steht es dem betreuenden FDZ frei, die Prüfung und Freigabe des Outputs abzulehnen.

Um Geheimhaltungsfälle in den Analysen zu vermeiden, sollte immer darauf geachtet werden, dass die erstellten Analysen auf ausreichend großen Fallzahlen beruhen. Bei geringen Fallzahlen empfiehlt es sich, Variablenausprägungen zusammenzufassen und damit größere Fallzahlen zu erzielen.

## 3.2 FAQ

### **Wie werden in den Mikrozensusdaten Haushalte selektiert?**

Für die Auswertung auf Haushaltsebene kann eine Selektion nach allen Privathaushalten (EF31=1) und der Bezugsperson im Haushalt (EF37=1) vorgenommen werden. Alternativ dazu ist es möglich eine eindeutige Haushalts-ID über die Aneinanderreihung der Eingabefelder Bundesland (EF1), Auswahlbezirk (EF3) und Haushaltsnummer (EF4) zu generieren und das Datenmaterial anhand dieser ID über alle Personen des Haushalts zu aggregieren.

### **Wie können Mütter und Väter ausgewertet werden?**

Über die Eingabefelder EF53 und EF55 besteht die Möglichkeit die Personennummer der Mutter bzw. des Vaters zu ermitteln. Da diese Personennummer lediglich den Kindern und nicht den Eltern selbst zugeordnet ist, muss zunächst eine neue Personennummer über Bundesland (EF1), Auswahlbezirk (EF3s), Haushaltsnummer (EF4s) und die Personennummer der Mutter (EF53) bzw. des Vaters (EF55) zusammengesetzt werden. Mit Hilfe dieser Nummer können nun die Eltern gekennzeichnet werden.

### **Wieso sind in den Daten doppelte Personennummern enthalten?**

Die Dopplung entsteht durch die in den Daten enthaltenen Jahresüberhänge aus dem Vorjahr. Über das Merkmal EF5U2 können diese von den Fällen abgegrenzt werden, die zur Stichprobe des aktuellen Erhebungsjahres gehören.

### **Wann wird welcher Hochrechnungsfaktor verwendet?**

Der Hochrechnungsfaktor für tiefregionalisierte Ergebnisse (EF953) wird dann verwendet, wenn Ergebnisse auf Ebene der regionalen Untergruppen ausgewertet werden. Für Ergebnisse auf Ebene der Anpassungsschichten und darüber wird der Standardhochrechnungsfaktor (EF952) verwendet. Bei der quartalsweisen statistischen Analyse (die Ergebnisse basieren also nur auf der Anzahl der Personen, die in dem jeweiligen Quartal befragt wurden) findet der Standardhochrechnungsfaktor für die Quartale (EF951) Anwendung.

Beinhaltet eine Analyse auf Bundesebene nur Personen der zehnprozentigen EU-Unterstichprobe des Ad-Hoc-Modules, ist der Hochrechnungsfaktor EF956 zu verwenden.

Im Mikrozensus 2013 gibt es einen Hochrechnungsfaktor speziell für Analysen der Gruppe der Menschen mit Behinderung (EF957) auf Bundesebene. Im Rahmen des Zusatzprogramms wurde auf freiwilliger Basis erfragt, ob eine amtlich festgestellte Behinderung vorliegt (EF465). Die Angaben zu den schwerbehinderten Menschen<sup>8</sup> werden dabei anhand der Ergebnisse der Schwerbehindertenstatistik, geschichtet nach Alter und Geschlecht, hochgerechnet. Zusätzlich werden die Antwortausfälle, ebenfalls geschichtet nach Alter und Geschlecht, bei Menschen mit leichten Behinderungen eingeschätzt. Eine regionalisierte Auswertung ist mit diesem Hochrechnungsfaktor nicht möglich.

### **3.3 Verfügbare Tools**

Für dieses Produkt werden seitens der FDZ keine weiterführenden Tools angeboten.

---

<sup>8</sup> gemäß SGB IX. § 2 Abs. 2, wenn der Grad der Behinderung mindestens 50% beträgt; EF466: Merkmalsausprägungen 04 bis 09



## Anhang

### Ausgewählte Häufigkeitstabellen und SPSS-Syntax (Anhang zu Abschnitt 2.3):

#### Bevölkerung (Personen am Hauptwohnsitz)<sup>9</sup>:

\*\*\*\* Bevölkerung am Hauptwohnsitz auswählen.

\*\*\*\*\*

IF (EF30=1 OR EF30=2) Hauptwohnsitz=1.

FILTER BY Hauptwohnsitz.

EXECUTE.

\*\*\* Hochrechnung: Hochrechnungsfaktor (Personenfaktor) in der regionalen.

\*\*\*\* Anpassungsschicht (EF952).

\*\*\*\*\*

WEIGHT BY EF952.

EXECUTE.

VARIABLE LABELS

EF46 'EF46NEU Geschlecht'.

EXECUTE.

ADD VALUE LABELS

EF46

1 'männlich'

2 'weiblich'.

EXECUTE.

\*\*\*\* Häufigkeiten für Geschlecht (EF46).

FREQUENCIES

VARIABLES= EF46

/ORDER= ANALYSIS.

**EF46NEU Geschlecht**

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Pro- zente
Gültig	männlich	39.454.210	48,9	48,9	48,9
	weiblich	41.156.827	51,1	51,1	100,0
	Gesamt	80.611.037	100,0	100,0	

\*\*\*\* Familienstand entsprechend der Publikationen zusammenfassen.

\*\*\*\*\*

RECODE EF49 (4=3) (5=2) (6=3) (7=3) (ELSE=COPY) INTO EF49NEU.

EXECUTE.

FORMATS EF49NEU (F1.0).

EXECUTE.

VARIABLE LABELS

<sup>9</sup> Vgl. Zahlen in DESTATIS Fachserie 1, Reihe 4.1.1, 2013 [https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/DEHeft\\_derivate\\_00014977/2010411137004\\_korr.pdf](https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/DEHeft_derivate_00014977/2010411137004_korr.pdf) Tabellen 1.2, 4.1 und 4.3

EF49NEU 'EF49NEU Familienstand'.  
EXECUTE.

ADD VALUE LABELS  
EF49NEU

- 1 'ledig'
- 2 'verheiratet (inkl. eingetragener Lebenspartnerschaften)'
- 3 geschieden/ 'verwitwet (inkl. aufgehobener Lebenspartnerschaften und Personen mit verstorbenem Lebenspartner).'

EXECUTE.

\*\*\*\* Häufigkeiten für Familienstand (EF49)

FREQUENCIES  
VARIABLES= EF49NEU  
/ORDER= ANALYSIS.

### EF49NEU Familienstand

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Pro- zente
Gültig	ledig	32.535.891	40,4	40,4	40,4
	verheiratet (inkl. eingetragener Le- benspartnerschaften)	36.748.979	45,6	45,6	85,9
	geschieden	11.326.167	14,1	14,1	100,0
	Gesamt	80.611.037	100,0	100,0	

\*\*\*\* Überwiegender Lebensunterhalt EF401. Wird entsprechend der Publikation umkodiert.

\*\*\*\*\*

RECODE EF401 (7=2) (5 THRU 9=3) (ELSE=COPY) INTO EF401NEU.  
EXECUTE.

FORMATS EF401NEU (F1.0).  
EXECUTE.

\*\*\*\* Labeln der Eingabefelder, die zur Prüfung der Daten benötigt werden.

\*\*\*\*\*

VARIABLE LABELS  
EF401NEU 'EF401NEU Überwiegender Lebensunterhalt'.  
EXECUTE.

ADD VALUE LABELS  
EF401NEU

- 1 'Erwerbs-/Berufstätigkeit'
- 2 'Arbeitslosengeld/-hilfe'
- 3 'Rente, Pension'
- 4 'Unterhalt durch Eltern, Ehemann/-frau oder andere Angehörige'
- 5 'Eigenes Vermögen, Vermietung, Zinsen, Altenteil'
- 6 'Sozialhilfe'
- 8 'Sonstige Unterstützungen (z.B. BAföG)'
- 9 'Elterngeld/Erziehungsgeld'.

EXECUTE.

\*\*\*\* Häufigkeiten für

\*\*\*\* Überwiegender Lebensunterhalt (EF401).

FREQUENCIES  
VARIABLES= EF401NEU  
/ORDER= ANALYSIS.

**EF401NEU Überwiegender Lebensunterhalt**

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Pro- zente
Gültig	Erwerbs-/Berufstätigkeit	35.424.311	43,9	43,9	43,9
	Arbeitslosengeld/-hilfe	3.860.493	4,8	4,8	48,7
	Rente, Pension	20.629.452	25,6	25,6	74,3
	Unterhalt durch Eltern, Ehemann/- frau oder andere Angehörige	20.696.781	25,7	25,7	100,0
	<b>Gesamt</b>	<b>80.611.037</b>	<b>100,0</b>	<b>100,0</b>	

\*\*\*\* Um Erwerbsstatus zu erhalten, wird EF29 (Nichterwerbspersonen) in Anlehnung an die  
\*\*\*\* Publikation umkodiert.

\*\*\*\*\*  
RECODE EF29 (3=4) (ELSE=COPY) INTO EF29NEU.  
EXECUTE.

FORMATS EF29NEU (F1.0).  
EXECUTE.

\*\*\*\* Labeln der Eingabefelder, die zur Prüfung der Daten benötigt werden.

\*\*\*\*\*  
VARIABLE LABELS  
EF29NEU 'EF29NEU Erwerbstyp'.  
EXECUTE.

ADD VALUE LABELS  
EF29NEU

1 'Erwerbstätige'  
2 'Erwerbslose'  
4 'Nichterwerbspersonen'.

EXECUTE.

\*\*\*\* Häufigkeiten für Erwerbstyp (EF29)

FREQUENCIES  
VARIABLES= EF29NEU  
/ORDER= ANALYSIS.

**EF29NEU Erwerbstyp**

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Pro- zente
Gültig	Erwerbstätige	39.617.920	49,1	49,1	49,1
	Erwerbslose	2.180.916	2,7	2,7	51,9
	Nichterwerbspersonen	38.812.201	48,1	48,1	100,0
	<b>Gesamt</b>	<b>80.611.037</b>	<b>100,0</b>	<b>100,0</b>	

\*\*\*\* Um die Stellung im Beruf nur von Erwerbstätigen zu.

\*\*\*\* erhalten, wird hier EF117 (Stellung im Beruf) umkodiert in EF117\_NEU.

\*\*\*\*\*  
DO IF (EF29=1).  
RECODE EF117 (1=1) (2=1) (3=3) (4=4) (5=5) (6=6) (9=4) (12=5) (13=13) (14=4) (15=13) (16=4) (17=5)

INTO EF117NEU.  
 END IF.  
 EXECUTE.  
 FORMATS EF117NEU (F2.0).  
 EXECUTE.

\*\*\*\* Labeln der Eingabefelder, die zur Prüfung der Daten benötigt werden.

\*\*\*\*\*

VARIABLE LABELS

EF117NEU 'EF117NEU Stellung im Beruf für Erwerbstätige'

EXECUTE.

ADD VALUE LABELS

EF117NEU

1 'Selbständige mit und ohne Beschäftigte'

3 'unbezahlt Mithelfende Familienangehörige'

4 'Beamte, Beamtinnen, Soldatinnen und Soldaten, Richterinnen und Richter, Beamtenanwärterinnen und -anwärter'

5 'Angestellter (inkl. Beschäftigte mit kleinem Job, ohne Auszubildende)'

6 'Arbeiter/-innen'

13 'Auszubildende, Volontärinnen und Volontäre, Trainees, bezahlte Praktikantinnen und Praktikanten'

EXECUTE.

\*\*\*\* \*\* Stellung im Beruf für Erwerbstätige (EF117\_NEU).

FREQUENCIES

VARIABLES= EF117NEU

/ORDER= ANALYSIS.

### EF117NEU Stellung im Beruf für Erwerbstätige

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Pro- zente
Gültig	1 Selbständige mit und ohne Beschäftigte	4.239.292	5,3	10,7	10,7
	3 unbezahlt mithelfende Familienangehörige	190.402	,2	,5	11,2
	4 Beamte, Beamtinnen, Soldatinnen und Soldaten, Richterinnen und Richter, Beamtenanwärterinnen und -anwärter	2.005.712	2,5	5,1	16,2
	5 Angestellter (inkl. Beschäftigte mit kleinem Job, ohne Auszubildende)	23.496.676	29,1	59,3	75,6
	6 Arbeiter/-innen	8.176.110	10,1	20,6	96,2
	13 Auszubildende, Volontärinnen und Volontäre, Trainees, bezahlte Praktikantinnen und Praktikanten	1.509.728	1,9	3,8	100,0
	Gesamt	39.617.920	49,1	100,0	
Fehlend	System	40.993.117	50,9		
Gesamt		80.611.037	100,0		

**Privathaushalte am Haupt- und Nebenwohnsitz<sup>10</sup>:**

\*\*\*\* Alle Fälle auswählen, die Bezugsperson in Privathaushalten sind, um die.  
 \*\*\*\* Privathaushalte auszählen zu können.  
 IF (EF31=1 AND EF37=1) Privathaushalte=1.  
 FILTER BY Privathaushalte.  
 EXECUTE.  
 WEIGHT BY EF952.  
 EXECUTE.

\*\*\*\* Umkodieren von EF20 (Anzahl der Personen in Privathaushalten) in.  
 \*\*\*\* EF20NEU, um die Haushaltsgröße zu kategorisieren.  
 RECODE EF20  
 (1=1) (2=2) (3=3) (4=4) (0=0) (ELSE=5) INTO EF20NEU.  
 EXECUTE.  
 FORMATS EF20NEU (F1.0).  
 EXECUTE.

\*\*\*\* Labeln und Häufigkeiten.  
 \*\*\*\*\*

VARIABLE LABELS  
 EF20NEU 'EF20NEU Privathaushalte kategorisiert'.

ADD VALUE LABELS  
 EF20NEU  
           0 'kein Privathaushalt'  
           1 'Einpersonenhaushalte'  
           2 'Zweipersonenhaushalte'  
           3 'Dreipersonenhaushalte'  
           4 'Vierpersonenhaushalte'  
           5 'Fünf- und mehr Personenhaushalte'.

EXECUTE.

FREQUENCIES  
 VARIABLES= EF20NEU  
 /ORDER= ANALYSIS.

**EF20NEU Privathaushalte kategorisiert**

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Pro- zente
Gültig	Einpersonenhaushalte	16.176.318	40,5	40,5	40,5
	Zweipersonenhaushalte	13.747.639	34,4	34,4	74,9
	Dreipersonenhaushalte	4.988.642	12,5	12,5	87,4
	Vierpersonenhaushalte	3.687.954	9,2	9,2	96,7
	Fünf- und mehr Personenhaushalte	1.332.281	3,3	3,3	100,0
	Gesamt	39.932.835	100,0	100,0	

<sup>10</sup> Vgl. Zahlen in Destatis Fachserie 1, Reihe 3, 2013 [https://www.destatis.de/GPStatistik/servlets/MCR-FileNodeServlet/DEHeft\\_derivate\\_00014754/2010300137004.pdf](https://www.destatis.de/GPStatistik/servlets/MCR-FileNodeServlet/DEHeft_derivate_00014754/2010300137004.pdf) Tab. 1.1 und 1.4.

\*\*\*\* Labeln und Häufigkeiten.

\*\*\*\*\*

VARIABLE LABELS  
EF1 'EF1 Bundesland'.

FREQUENCIES  
VARIABLES= EF1  
/ORDER= ANALYSIS.

**EF1 Bundesland**

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Pro- zente
Gültig	Schleswig-Holstein	1.403.433	3,5	3,5	3,5
	Hamburg	965.099	2,4	2,4	5,9
	Niedersachsen	3.827.891	9,6	9,6	15,5
	Bremen	353.222	,9	,9	16,4
	Nordrhein-Westfalen	8.518.603	21,3	21,3	37,7
	Hessen	2.937.352	7,4	7,4	45,1
	Rheinland-Pfalz	1.887.461	4,7	4,7	49,8
	Baden-Württemberg	5.014.338	12,6	12,6	62,4
	Bayern	6.123.056	15,3	15,3	77,7
	Saarland	492.915	1,2	1,2	78,9
	Berlin	1.933.548	4,8	4,8	83,8
	Brandenburg	1.238.367	3,1	3,1	86,9
	Mecklenburg-Vorpommern	827.932	2,1	2,1	89,0
	Sachsen	2.146.093	5,4	5,4	94,3
	Sachsen-Anhalt	1.154.193	2,9	2,9	97,2
	Thüringen	1.109.333	2,8	2,8	100,0
	Gesamt	39.932.835	100,0	100,0	

\*\*\*\* Umkodieren von EF735 entsprechend der Publikation umkodiert.

\*\*\*\*\*

RECODE EF735  
(6=2) (8=4) (9=5) (7=3) (ELSE=COPY) INTO EF735NEU.  
EXECUTE.

FORMATS EF735NEU (F1.0).  
EXECUTE.

\*\*\*\* Labeln und Häufigkeiten.

\*\*\*\*\*

VARIABLE LABELS  
EF735NEU 'EF735NEU Familienstand des Haupteinkommensbeziehers'.

ADD VALUE LABELS  
EF735NEU

- 1 'ledig'
- 2 'verheiratet zusammenlebend (inkl. eingetragene Lebenspartnerschaft)'
- 3 'verheiratet getrennt lebend (inkl. eingetragene Lebenspartnerschaft)'
- 4 'geschieden/ Lebenspartnerschaft aufgehoben'
- 5 'verwitwet/ Lebenspartner verstorben'.

EXECUTE.

FREQUENCIES  
VARIABLES= EF735NEU  
/ORDER= ANALYSIS.

**EF735NEU Familienstand des Haupteinkommensbeziehers**

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Pro- zente
Gültig	ledig	11.744.225	29,4	29,4	29,4
	verheiratet zusammenlebend (inkl. eingetragene Lebenspartner- schaft)	17.288.713	43,3	43,3	72,7
	verheiratet getrennt lebend (inkl. eingetragene Lebenspartner- schaft)	1.374.123	3,4	3,4	76,1
	geschieden/ Lebenspartnerschaft aufgehoben	4.594.631	11,5	11,5	87,7
	verwitwet/ Lebenspartner verstor- ben	4.931.142	12,3	12,3	100,0
	Gesamt	39.932.835	100,0	100,0	

Statistische Ämter des Bundes und der Länder,  
Metadatenreport – Teil II: Produktspezifische Informationen zur Nutzung des Mikrozensus 2013 am  
Gastwissenschaftsarbetsplatz sowie per kontrollierter Datenfernverarbeitung (sog. On-Site-Zugangswege/-  
Nutzungen)

Fotorechte Umschlag: ©artSILENCEcom – Fotolia.com