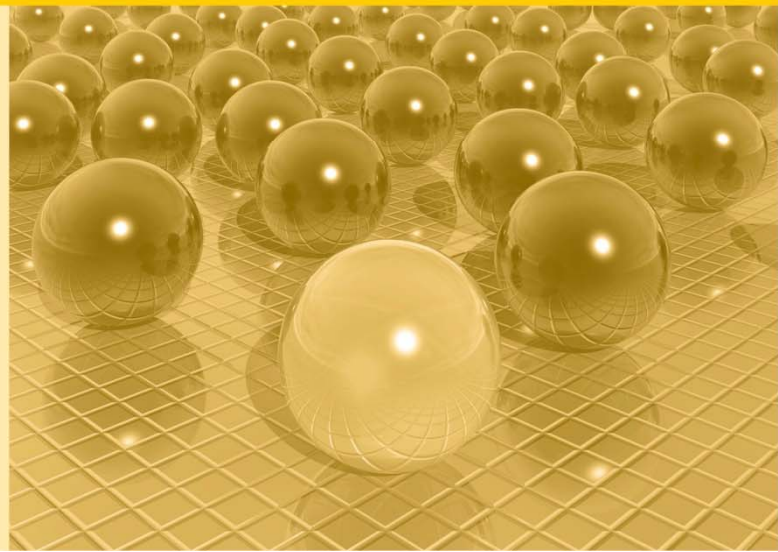


# Metadatenreport



Teil II: Produktspezifische Informationen zur Nutzung des Mikrozensus 2015  
am Gastwissenschaftsarbeitsplatz sowie per kontrollierter  
Datenfernverarbeitung (sog. On-Site-Zugangswege/-Nutzungen)

DOI: 10.21242/12211.2015.00.00.1.1.1

Version 1

## 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 9449-2883  
Fax: 0211 9449-8087  
[forschungsdatenzentrum@it.nrw.de](mailto:forschungsdatenzentrum@it.nrw.de)

Erscheinungsfolge: unregelmäßig  
Erschienen im Januar 2021

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

© Information und Technik Nordrhein-Westfalen, Düsseldorf, 2021  
(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 2015 am Gastwissenschaftsarbeitsplatz sowie per kontrollierter Datenfernverarbeitung (EVAS-Nummer: 12211, 12212, 12213). Version 1. DOI: 10.21242/12211.2015.00.00.1.1.0. Düsseldorf 2021.

# Metadatenreport

Teil II: Produktspezifische Informationen zur Nutzung des Mikrozensus 2015 am Gastwissenschaftsarbetsplatz sowie per kontrollierter Datenfernverarbeitung (sog. On-Site-Zugangswege/-Nutzungen)

DOI: 10.21242/12211.2015.00.00.1.1.1

Version 1



## Inhalt

<b>1. Datenaufbereitung in den FDZ.....</b>	<b>2</b>
<b>1.1 Datenaufbereitung .....</b>	<b>2</b>
<b>1.2 Anonymisierungsmaßnahmen.....</b>	<b>3</b>
<b>1.3 Methodik der Verknüpfung.....</b>	<b>3</b>
<b>2. Produkt .....</b>	<b>4</b>
<b>2.1 Merkmale und Merkmalsbeschreibung .....</b>	<b>4</b>
<b>2.1.1. Merkmalsdefinitionen .....</b>	<b>4</b>
<b>2.1.2 Datensatzbeschreibung / Schlüsselverzeichnis .....</b>	<b>4</b>
<b>2.2 Vergleichbarkeit der Merkmale über die Zeit .....</b>	<b>5</b>
<b>2.3 Eckwerte relevanter Merkmale und Merkmalskombinationen .....</b>	<b>6</b>
<b>2.4 Auswertbare regionale Ebene.....</b>	<b>8</b>
<b>2.5 Produktversion.....</b>	<b>8</b>
<b>3. Praktische Hinweise .....</b>	<b>10</b>
<b>3.1 Hinweise zur Geheimhaltung .....</b>	<b>10</b>
<b>3.1.1 Gesetzliche Grundlagen der statistischen Geheimhaltung.....</b>	<b>10</b>
<b>3.1.2 Geheimhaltung von Ergebnissen.....</b>	<b>10</b>
<b>3.1.3 Praktische Tipps zur Vermeidung von Geheimhaltungsfällen.</b>	<b>11</b>
<b>3.2 FAQ .....</b>	<b>11</b>
<b>3.3 Verfügbare Tools .....</b>	<b>12</b>
<b>Anhang .....</b>	<b>13</b>

# 1. Datenaufbereitung in den FDZ

## 1.1 Datenaufbereitung

Das Datenmaterial wird einer Vollzählkeits- 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 nicht in die Hochrechnung miteinfließen, 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 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 im Schlüsselverzeichnis für die betroffenen Merkmale systemfehlende Werte vergeben (Ausprägung „leer“). Treten Inkonsistenzen auf, werden diese 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 2.4 Auswertbare regionale Ebene) werden in diesem Zuge auf inhaltliche Konsistenz geprüft. Treten Inkonsistenzen auf, werden diese dokumentiert und gegebenenfalls bereinigt.

---

<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?“.

## **1.2 Anonymisierungsmaßnahmen**

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. Da sich nach 2015 die Auswahlgrundlage der Stichprobe verändert hat, ist keine Verknüpfung mit nachfolgenden Erhebungsjahren möglich.

Genauere Informationen finden Sie unter: [https://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2016/06/die-neue-mikrozensusstichprobe-062016.pdf?\\_\\_blob=publicationFile&v=3](https://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2016/06/die-neue-mikrozensusstichprobe-062016.pdf?__blob=publicationFile&v=3)

## 2. Produkt

### 2.1 Merkmale und Merkmalsbeschreibung

#### 2.1.1. Merkmalsdefinitionen

Folgende Klassifikationen finden in dem beschriebenen Produkt Anwendung:

- Klassifikation der Berufe, Ausgabe 2010 (KldB 2010)  
<https://statistik.arbeitsagentur.de/Statischer-Content/Grundlagen/Klassifikationen/Klassifikation-der-Berufe/KldB2010/Systematik-Verzeichnisse/Generische-Publikationen/Systematisches-Verzeichnis-Berufsbenennung.xls>
- Internationale Standardklassifikation der Berufe, Ausgabe 2008 (ISCO-08)  
<http://www.ilo.org/public/english/bureau/stat/isco/isco08/> (Englisch)
- 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)
- Bildungsklassifikation International Standard Classification of Education, Ausgabe 2011 (ISCED-2011):  
<http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf> (Englisch)  
Staats- und Gebietssystematik  
[https://www.destatis.de/DE/Methoden/Klassifikationen/Staat-Gebietsystematik/Staatsangehoerigkeitsgebietsschlüssel\\_pdf.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Methoden/Klassifikationen/Staat-Gebietsystematik/Staatsangehoerigkeitsgebietsschlüssel_pdf.pdf?__blob=publicationFile)
- Amtlicher Gemeindeschlüssel (AGS)/ Regionalschlüssel (RS)  
[https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis\\_inhalt.html](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 ist dem Schlüsselverzeichnis zu entnehmen. Sofern Variablen das Ergebnis einer Frage des Fragebogens wiedergeben, ist die zugehörige Fragennummer in der Spalte „Fragebogennummer“ enthalten.

Schlüsselverzeichnis

[https://www.forschungsdatenzentrum.de/sites/default/files/mz\\_2015\\_on-site\\_svz\\_v1.pdf](https://www.forschungsdatenzentrum.de/sites/default/files/mz_2015_on-site_svz_v1.pdf)



Erhebungsunterlagen

[https://www.forschungsdatenzentrum.de/sites/default/files/mz\\_2015\\_eu.pdf](https://www.forschungsdatenzentrum.de/sites/default/files/mz_2015_eu.pdf)

Im Jahr 2015 wurde das Zusatzprogramm „Erwerbstätigkeit und Krankenversicherung“ abgefragt. Es besteht eine Vergleichbarkeit mit den Jahren 2007 und 2011. Ein Ad-Hoc-Modul gab es nicht.

## 2.2 Vergleichbarkeit der Merkmale über die Zeit

Folgende Merkmale weisen im Vergleich zum Vorjahr Veränderungen auf:

### Haushalts-/Personenidentifikation

- EF 10 (Wann wurde das Haus, in dem Sie hier wohnen, gebaut?): Die Frage nach dem Baualter der Wohnung wird wieder im Grundprogramm erfasst.

### Angaben zur Person

- EF67 (Haben Sie irgendeinen Nebenjob oder eine Tätigkeit, mit der Sie Geld verdienen?): besitzt eine textlich veränderte Ausprägung (3), welche nun definiert wird als „Ja, und zwar mit einer kurzfristigen Beschäftigung (höchstens 2 Monate oder 50 Arbeitstage im Jahr)“
- EF160 (Ist Ihre Tätigkeit eine geringfügige Beschäftigung?): besitzt eine textlich veränderte Ausprägung (3):“ Ja, eine kurzfristige Tätigkeit (höchstens 2 Monate oder 50 Arbeitstage im Jahr)“

### Zusatzprogramm (Vergleich zu 2011)

- EF172 (Welche Tätigkeit üben Sie schwerpunktmäßig aus?): Die vorherige Antwortkategorie „Forschen, Entwerfen, Konstruieren, Gestalten von Produkten, Plänen, Programmen“ wurde ausdifferenziert in „Forschen und Entwickeln“ (21) sowie „Konstruieren, Gestalten von Produkten, Plänen, Programmen“(9).  
Beide Antwortkategorien sind Teilaspekte des Bereichs „Büro, Technisches Büro, EDV, Forschen“.
- EF175 (Was trifft auf Ihre gegenwärtige Tätigkeit zu?): erfragt die Stellung im Betrieb und weist Änderungen bei den Merkmalsausprägungen der Auszubildende auf. Es wird bei den Ausprägungen differenziert, ob es sich um Beamtenanwärter im einfachen, gehobenen usw.

Dienst handelt. Im Jahr 2011 wurden Auszubildende in zwei Kategorien (11,12) erfasst. Es werden im Jahr 2015 hingegen insgesamt sechs Ausprägungen verwendet (22 bis 27).

## 2.3 Eckwerte relevanter Merkmale und Merkmalskombinationen

### Verteilung der Bevölkerung am Hauptwohnsitz

EF46 Geschlecht					
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Pro- zente
Gültig	männlich	39.971.576	49,1	49,1	49,1
	weiblich	41.432.612	50,9	50,9	100,0
Gesamt		81.404.188	100,0	100,0	

### Privathaushalte am Haupt- und Nebenwohnsitz

Tab. 3: 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 Mehrpersonenhaushalte	1.332.281	3,3	3,3	100,0
Gesamt		39.932.835	100,0	100,0	

**Tab. 4: EF1 Bundesland**

		<b>EF1 Land der Bundesrepublik</b>			
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Pro- zente
Gültig	Schleswig-Holstein	14.394.75	3,5	3,5	3,5
	Hamburg	986.986	2,4	2,4	6,0
	Niedersachsen	3.901.563	9,6	9,6	15,5
	Bremen	364.783	,9	,9	16,4
	Nordrhein-Westfalen	8.681.082	21,3	21,3	37,7
	Hessen	2.991.356	7,3	7,3	45,0
	Rheinland-Pfalz	1.918.915	4,7	4,7	49,7
	Baden-Württemberg	5.161.196	12,7	12,7	62,4
	Bayern	6.304.780	15,5	15,5	77,9
	Saarland	490.137	1,2	1,2	79,1
	Berlin	2.000.063	4,9	4,9	84,0
	Brandenburg	1.244.558	3,1	3,1	87,0
	Mecklenburg-Vorpommern	836.975	2,1	2,1	89,1
	Sachsen	2.171.520	5,3	5,3	94,4
	Sachsen-Anhalt	1.167.378	2,9	2,9	97,3
	Thüringen	1.113.507	2,7	2,7	100,0
	<b>Gesamt</b>	<b>40.774.274</b>	<b>100,0</b>	<b>100,0</b>	

Im Anhang befinden sich die oben aufgeführten und weitere Tabellen mit der entsprechenden Syntax für das Programm SPSS und eine Aufführung der 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 durch geringe Zellbesetzungen) (NUTS 3)
- Gemeindeebene (ggf. eingeschränkte Aussagekraft durch geringe Zellbesetzungen) (LAU)
- Anpassungsschichten (regionale Schicht und regionale Untergruppe)<sup>3</sup>
- Raumordnungsregionen
- Arbeitsmarktregionen
- Nielseengebiete
- Arbeitsagenturbezirke

## 2.5 Produktversion

### 1.0.0

- Erstveröffentlichung

### 1.1.0

- Der Hochrechnungsfaktor.EF951 wurde mit 1000 multipliziert. Dies geschieht einheitlich zu den anderen HFR um auf die Bevölkerung hochrechnen zu können.  $EF951 = Ef951 * 1000$ .

### 1.1.2

- Ersterstellung Englischübersetzung
- Änderung EF564: Anpassung der Stellenzahl vorher 2 jetzt 4
- Korrektur der Bezeichnung der Anpassungsschicht

### 1.2.2 (DOI:10.21242/12211.2015.00.00.1.1.0)

- AGS des Mikrozensus an den AGS des Gemeindeverzeichnisses angepasst. Betroffen sind Berlin, Hamburg, Aachen
- EF313 EF313U1 und EF313U2 in den Datensatz aufgenommen
- Eineindeutige Benennung von Regionaler Anpassungsschicht (EF564), Regionaler Schicht (EF566) und Regionaler Untergruppe (EF568)

---

<sup>3</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.

- Hinweise auf Bayern Pseudonymisierung in das Schlüsselverzeichnis aufgenommen (EF560UG1)
- Standard-File-Hinweis entfernt

1.3.0 (DOI: 10.21242/12211.2012.00.00.1.1.1)

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

## **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, das heißt 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 Gastwissenschaftsarbeitsplatz 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.

Die Broschüre finden Sie hier:

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

### **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?**

In den Veröffentlichungen der amtlichen Statistik wird eine Selektion nach allen Privathaushalten (EF31=1) und der Bezugsperson im Haushalt (EF37=1) vorgenommen. Alternativ dazu ist es möglich eine eindeutige Haushalts-ID über die Aneinanderreihung der Eingabefelder Bundesland (EF1), Auswahlbezirk (EF3), Haushaltsnummer (EF4) und Kennzeichnung von Überhanginterviews (EF5U2) zu generieren und das Datenmaterial anhand dieser ID über alle Personen des Haushalts zu aggregieren. Sofern es sich nicht um ein Überhanginterview handelt (EF5U2=missing), empfiehlt es sich der Haushalts-ID eine 0 anzuhängen, um Dubletten aufgrund der Verschiebung von Stellen zu vermeiden.

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

Zunächst muss eine Haushaltsidentifikationsnummer erstellt werden. Hierfür werden folgende Variablen verkettet:

- EF1 Bundesland
- EF3 Auswahlbezirk
- EF4 Haushaltsnummer
- EF5U2 Kennzeichnung der Jahresüberhänge

Sofern EF5U2 missing ist, handelt es sich nicht um ein Überhang Interview. Um Dubletten durch die Verschiebung von Stellen zu vermeiden, ist in diesen Fällen eine 0 anzuhängen. Über die Eingabefelder EF53 und EF55 besteht die Möglichkeit die Personnummer der Mutter bzw. des Vaters zu ermitteln. Bei Überhanginterviews ist diesen Nummern EF5U2 anzuhängen. Über EF5 können diese Nummern Personen nun im Haushalt identifiziert werden. EF5 ist die EF5U2 bereits angehängt.

### **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 zur Jahrestichprobe gehörenden Fällen abgegrenzt werden.

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

Für Ergebnisse auf Ebene der Anpassungsschichten und höher (zum Beispiel Bundesländer) wird der Standardhochrechnungsfaktor (EF952) verwendet. Für Ergebnisse unterhalb der Anpassungsschichten (zum Beispiel Kreisebene) bis hinunter zur regionalen Untergruppe wird der Hochrechnungsfaktor für tiefregionalisierte Ergebnisse (EF953) empfohlen. 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. Eine ausführliche Erläuterung der Modelle für die einzelnen Faktoren ist dem Metadatenreport Teil I in Abschnitt 2.6 zu entnehmen

## **3.3 Verfügbare Tools**

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



# Anhang

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

### Bevölkerung (Personen am Hauptwohnsitz)<sup>4</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	39.971.576	49,1	49,1	49,1
	weiblich	41.432.612	50,9	50,9	100,0
Gesamt		81.404.188	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
EF49NEU 'EF49NEU Familienstand'.
```

<sup>4</sup> Vgl. Zahlen in DESTATIS Fachserie 1, Reihe 4.1.1, 2015 [https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEHeft\\_derivate\\_00029145/2010411157004\\_Korr21112016.pdf](https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEHeft_derivate_00029145/2010411157004_Korr21112016.pdf) Tabellen 1.2, 4.1 und 4.3

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	33.037.757	40,6	40,6	40,6
	verheiratet (inkl. eingetragener Lebenspartnerschaften)	36.715.060	45,1	45,1	85,7
	geschieden/ verwitwet (inkl. aufgehobener Lebenspartnerschaften und Personen mit verstorbenem Lebenspartner)	11.651.371	14,3	14,3	100,0
	Gesamt	81.404.188	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 'eigene 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	eigene Erwerbs-/Berufstätig- keit	36.209.098	44,5	44,5	44,5
	Arbeitslosengeld/-hilfe	3.699.206	4,5	4,5	49,0
	Rente, Pension	21.349.876	26,2	26,2	75,3
	Unterhalt durch Eltern, Ehe- mann/-frau oder andere An- gehörige	20.146.008	24,7	24,7	100,0
	Gesamt	81.404.188	100,0	100,0	

\*\*\*\* 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	40.278.885	49,5	49,5	49,5
	Erwerbslose	1.948.948	2,4	2,4	51,9
	Nichterwerbspersonen	39.176.355	48,1	48,1	100,0
	Gesamt	81.404.188	100,0	100,0	

```

**** 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 'Beamtinnen und Beamte, Soldatinnen und Soldaten, Richterinnen und Richter, Beamtenanwärterinnen und -an-
wärter'
    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	Selbständige mit und ohne Beschäftigte	4.160.737	5,1	10,3	10,3
	unbezahlt mithelfende Familienangehörige	160.455	,2	,4	10,7
	Beamtinnen und Beamte, Soldatinnen und Soldaten, Richterinnen und Richter, Beamtenanwärterinnen und -anwärter	1.990.905	2,4	4,9	15,7
	Angestellter (inkl. Beschäftigte mit kleinem Job, ohne Auszubildende)	23.862.289	29,3	59,2	74,9
	Arbeiter/-innen	8.622.910	10,6	21,4	96,3
	Auszubildende, Volontärinnen und Volontäre, Trainees, bezahlte Praktikantinnen und Praktikanten	1.481.589	1,8	3,7	100,0
	Gesamt	40.278.885	49,5	100,0	
Fehlend	System	41.125.303	50,5		
Gesamt		81.404.188	100,0		

## Privathaushalte am Haupt- und Nebenwohnsitz<sup>5</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 Mehrpersonenhaushalte'.

EXECUTE.

FREQUENCIES

VARIABLES= EF20NEU

/ORDER= ANALYSIS.

### EF20NEU Privathaushalte kategorisiert

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Pro- zente
Gültig	Einpersonenhaushalte	16.875.489	41,4	41,4	41,4
	Zweipersonenhaushalte	13.956.079	34,2	34,2	75,6
	Dreipersonenhaushalte	4.940.162	12,1	12,1	87,7
	Vierpersonenhaushalte	3.679.426	9,0	9,0	96,8
	Fünf- und Mehrpersonenhaushalte	1.323.119	3,2	3,2	100,0
	Gesamt	40.774.274	100,0	100,0	

<sup>5</sup> Vgl. Zahlen in Destatis Fachserie 1, Reihe 3, 2015 [https://www.statistischebibliothek.de/mir/servlets/MCR-FileNodeServlet/DEHeft\\_derivate\\_00026228/2010300157004.pdf](https://www.statistischebibliothek.de/mir/servlets/MCR-FileNodeServlet/DEHeft_derivate_00026228/2010300157004.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	1.439.475	3,5	3,5	3,5
	Hamburg	986.986	2,4	2,4	6,0
	Niedersachsen	3.901.563	9,6	9,6	15,5
	Bremen	364.783	,9	,9	16,4
	Nordrhein-Westfalen	8.681.082	21,3	21,3	37,7
	Hessen	2.991.356	7,3	7,3	45,0
	Rheinland-Pfalz	1.918.915	4,7	4,7	49,7
	Baden-Württemberg	5.161.196	12,7	12,7	62,4
	Bayern	6.304.780	15,5	15,5	77,9
	Saarland	490.137	1,2	1,2	79,1
	Berlin	2.000.063	4,9	4,9	84,0
	Brandenburg	1.244.558	3,1	3,1	87,0
	Mecklenburg-Vorpommern	836.975	2,1	2,1	89,1
	Sachsen	2.171.520	5,3	5,3	94,4
	Sachsen-Anhalt	1.167.378	2,9	2,9	97,3
	Thüringen	1.113.507	2,7	2,7	100,0
	<b>Gesamt</b>	<b>40.774.274</b>	<b>100,0</b>	<b>100,0</b>	



\*\*\*\* 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	12.263.614	30,1	30,1	30,1
	verheiratet zusammenlebend (inkl. eingetragene Lebens- partnerschaft)	17.230.208	42,3	42,3	72,3
	verheiratet getrennt lebend (inkl. eingetragene Lebens- partnerschaft)	1.440.096	3,5	3,5	75,9
	geschieden/ Lebenspartner- schaft aufgehoben	4.854.376	11,9	11,9	87,8
	verwitwet/ Lebenspartner verstorben	4.985.980	12,2	12,2	100,0
	Gesamt	40.774.274	100,0	100,0	

Statistische Ämter des Bundes und der Länder,  
Metadatenreport – Teil II: Produktspezifische Informationen zur Nutzung des Mikrozensus 2015 am  
Gastwissenschaftsarbetsplatz sowie per kontrollierter Datenfernverarbeitung (sog. On-Site-Zugangswege/-  
Nutzungen)

Fotorechte Umschlag: ©artSILENCEcom – Fotolia.com