

Metadatenreport



Teil II: Produktspezifische Informationen zur Nutzung des Mikrozensus 2018 am
Gastwissenschaftsarbetsplatz sowie per kontrollierter Datenfernverarbeitung
(sog. On-Site-Zugangswege/-Nutzungen)

DOI: 10.21242/12211.2018.00.00.1.1.3

Version 3

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
E-Mail: 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

Informationen zum Datenangebot:

Statistisches Bundesamt
Forschungsdatenzentrum
Tel.: 0611 75-2420
Fax: 0611 75-3915
forschungsdatenzentrum@destatis.de

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

Erscheinungsfolge: unregelmäßig
Erschienen im Oktober 2021

Diese Publikation wird kostenlos als PDF-Datei zum Download unter 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 2018 am Gastwissenschaftsarbeitsplatz sowie per kontrollierter Datenfernverarbeitung (sog. On-Site-Zugangswege/-Nutzungen) (EVAS-Nummer: 12211, 12212, 12213). Version 3. DOI: 10.21242/12211.2018.00.00.1.1.3. Düsseldorf 2021.

Metadatenreport

Teil II: Produktspezifische Informationen zur Nutzung des Mikrozensus 2018 am Gastwissenschaftsarbeitsplatz sowie per kontrollierter Datenfernverarbeitung (sog. On-Site-Zugangswege/-Nutzungen)

DOI: 10.21242/12211.2018.00.00.1.1.3

Version 3

Inhalt

1. Datenaufbereitung in den FDZ.....	2
1.1 Datenaufbereitung.....	2
1.2 Anonymisierungsmaßnahmen	3
1.3 Methodik der Verknüpfung	3
2. Produkt	4
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 ..	7
2.4 Auswertbare regionale Ebene	8
2.5 Produktversionen.....	9
3. Praktische Hinweise.....	10
3.1 Hinweise zur Geheimhaltung.....	10
3.1.1 Gesetzliche Grundlagen der statistischen Geheimhaltung	10
3.1.2 Geheimhaltung von Ergebnissen	11
3.1.3 Praktische Tipps zur Vermeidung von Geheimhaltungsfällen	11
3.2 FAQ	12
3.3 Verfügbare Tools	13
Anhang	14

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 Veröffentlichungen¹ des Statistischen Bundesamts (siehe Metadatenreport Teil I, Kapitel 4 Zentrale Veröffentlichungen) werden die Hochrechnungsfaktoren so angeglichen, dass sie ohne zusätzliche Multiplikation auf die Gesamtbevölkerung hochrechnen². Fälle, die keinen positiven Hochrechnungsfaktor besitzen, werden aus dem Datenmaterial entfernt. Dies betrifft ca. 9,9% der Fälle aus dem Originaldatenmaterial. Für diese Haushalte liegen keine Auskünfte vor, weil es sich beispielsweise um leerstehende Wohnungen oder weggefallene Haushalte handelt.

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 2.4 Auswertbare regionale Ebene) werden in diesem

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

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

Zuge auf inhaltliche Konsistenz geprüft. Treten Inkonsistenzen auf, werden diese im Schlüsselerzeichnis (siehe auch Kapitel 2.1.2) 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 Merkmale EF1-EF4 und EF63, auf Individualebene die Wellen 2016-2019 miteinander zu verknüpfen.

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/> (Englisch)
- Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008)
<https://www.destatis.de/DE/Methoden/Klassifikationen/Gueter-Wirtschaftsklassifikationen/Downloads/klassifikation-wz-2008-3100100089004.pdf>
- 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)
- ISCED Bildungs- und Ausbildungsfelder 2013
<http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-fields-of-education-and-training-2013-detailed-field-descriptions-2015-en.pdf> (Englisch)
- Staats- und Gebietssystematik
https://www.destatis.de/DE/Methoden/Klassifikationen/Staat-Gebietssystematik/Staatsangehoerigkeitsgebietsschluesel_pdf.pdf
- 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 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_2018_on-site_svz.pdf

Erhebungsunterlagen

https://www.forschungsdatenzentrum.de/sites/default/files/mz_2018_eu_zusatz.pdf

2.2 Vergleichbarkeit der Merkmale über die Zeit

Gesetzlich bedingt wird in Gemeinschaftsunterkünften nur noch ein reduziertes Frageprogramm verwendet. Das Gesetz zur Neuregelung des Mikrozensus und zur Änderung weiterer Statistikgesetze ist hier einzusehen³:

https://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl116s2826.pdf

Folgende Merkmale weisen im Vergleich zum Vorjahr Veränderungen auf:

Haushalts-/Personenidentifikation

- EF34 Stellung in der Familie (traditionelles Familienkonzept): Ausprägung 2 „Ehepartner der Bezugsperson in der Familie“ textlich geändert, hier werden nun auch die gleichgeschlechtlichen Paare mitberücksichtigt, eine Vergleichbarkeit ist gegeben
- EF35 Stellung in der Familie/Lebensform (Lebensformenkonzept): Ausprägung 2 „Ehe- oder Lebenspartner/-in der Bezugsperson in der Familie/Lebensform“ textlich geändert, hier werden nun auch die gleichgeschlechtlichen Paare mitberücksichtigt, eine Vergleichbarkeit ist gegeben

Gegenwärtige (Haupt-)Erwerbstätigkeit

- EF153 Nachtarbeitsstunden (zwischen 23 und 6 Uhr): Variable entfällt

Angaben zur Arbeitsuche/zum Arbeitsplatzwechsel

- EF278 Dauer der Arbeitsuche: Ausprägungen 1 „Unter 1 Monat“ und 8 „4 Jahre oder mehr“ rein textlich verändert, eine Vergleichbarkeit ist gegeben

Gegenwärtiger Besuch von Hochschule und Schule

- EF289 Art der besuchten Schule: Ausprägung 26 „Fachhochschule (auch Hochschule (FH) für angewandte Wissenschaften), Duale Hochschule (in Baden-Württemberg und Thüringen)“ rein textlich geändert, eine Vergleichbarkeit ist gegeben

Schul-/Ausbildungsabschluss

³ Der Artikel von Hundenborn/Enderer 2019 zur Neuregelung des Mikrozensus erwähnt auch Änderungen, die bereits 2018 relevant waren: https://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2019/06/neuregelung-mikrozensus-062019.pdf?__blob=publicationFile

- EF309 Allgemeiner Schulabschluss vorhanden: Ausprägung leer „Entfällt (Personen unter 15 Jahren; Schüler an allgemeinbildenden Schulen, teilweise auch Meisterausbildung an Fachschulen)“ rein textlich geändert, eine Vergleichbarkeit ist gegeben
- EF312 Höchster beruflicher Ausbildungs- oder Hochschul-/Fachhochschulabschluss: Änderungen einiger Ausprägungen, eine Vergleichbarkeit ist gegeben
- EF313 Hauptfachrichtung: Ausprägung „Personen, vor 1954 geboren, mit Anlernausbildung“ entfällt, Bandsatzstellen wurden erweitert, eine Vergleichbarkeit ist nur bedingt gegeben
- EF313U1 Hauptfachrichtung (Stellen 1 bis 4): Ausprägung „Personen, vor 1954 geboren, mit Anlernausbildung“ entfällt, Bandsatzstellen wurden erweitert, eine Vergleichbarkeit ist nur bedingt gegeben
- EF313U2 Hauptfachrichtung (Stellen 5 bis 7): Ausprägung 999 „Ohne Angabe“ textlich geändert, eine Vergleichbarkeit ist bedingt gegeben
- EF316 Welchen höchsten weiteren beruflichen Abschluss?: Ausprägung 12 „Erzieher/-in“ textlich geändert, eine Vergleichbarkeit ist gegeben
- EF320 Höchster beruflicher Ausbildungs- oder Hochschul-/Fachhochschulabschluss mit Grad (Diplom, Bachelor, Master): Ausprägung 15 „Erzieher/-in“ textlich geändert, eine Vergleichbarkeit ist gegeben

Unterhalt, Einkommen

- EF436 Höhe des Nettoeinkommens im letzten Monat (je Haushaltsmitglied): Ausprägung 24 „18000 Euro oder mehr“ rein textlich geändert, eine Vergleichbarkeit ist gegeben
- EF442 Wie hoch ist Ihr monatlicher Nettogehalt/ monatlicher Nettolohn im Durchschnitt?: Ausprägung 24 „18000 Euro oder mehr“ rein textlich geändert, eine Vergleichbarkeit ist gegeben

Wohnsituation des Haushalts (Veränderungen im Vergleich zu 2014, Zusatzprogramm Wohnen)

In der amtlichen Statistik hat sich die Berechnung der Quadratmetermiete sowie die der Mietbelastungsquoten ab 2018 geändert.

- EF489 In welcher Art von Gebäude wohnen Sie?: Ausprägungen haben sich verändert, eine Vergleichbarkeit ist nur sehr eingeschränkt möglich
- EF490 Wie viele Wohnungen (einschließlich der leerstehenden Wohnungen) gibt es in dem Gebäude, in dem Sie wohnen?: Ausprägungen wurden reduziert (z.B. Ausprägung 1 „1 oder 2 Wohnungen“ war 2014 in zwei Ausprägungen aufgeteilt), eine Vergleichbarkeit ist eingeschränkt gegeben

- EF491 Bewohnen Sie Ihre Wohnung/Ihr Haus als...: Ausprägungen rein textlich geändert, eine Vergleichbarkeit ist gegeben
- EF492 Wie groß ist die Wohnfläche der gesamten Wohnung/des Hauses?: Schlüssel 000 nun zulässig und neue Ausprägung 999, eine Vergleichbarkeit ist gegeben
- EF493 Wann ist Ihr Haushalt in die Wohnung/das Haus eingezogen?: Variable neu
- EF494 Wann war das ursprüngliche Baujahr des Gebäudes?: textlich geändert, Ausprägungen leicht anders, eine Vergleichbarkeit ist gegeben
- EF495 Ist die von Ihnen gemietete Wohnung eine Eigentumswohnung?: Variable neu

2.3 Eckwerte relevanter Merkmale und Merkmalskombinationen

Verteilung der Bevölkerung am Hauptwohnsitz

Tab. 1: EF46NEU Geschlecht

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	männlich	40.405.761	49,5	49,5	49,5
	weiblich	41.206.824	50,5	50,5	100,0
	Gesamt	81.612.585	100,0	100,0	

Tab. 2: EF49 Familienstand

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	ledig	33.659.677	41,2	41,2	41,2
	verheiratet (inkl. eingetragener Lebenspartnerschaften)	36.892.280	45,2	45,2	86,4
	geschieden / verwitwet	11.060.629	13,6	13,6	100,0
	Gesamt	81.612.585	100,0	100,0	

Privathaushalte am Haupt- und Nebenwohnsitz

Tab. 3: EF20NEU Privathaushalte kategorisiert

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Einpersonenhaushalte	17.333.411	41,9	41,9	41,9
	Zweipersonenhaushalte	13.983.135	33,8	33,8	75,7
	Dreipersonenhaushalte	4.923.446	11,9	11,9	87,6
	Vierpersonenhaushalte	3.748.344	9,1	9,1	96,6

Fünf- und mehr Personenhaushalte	1.389.576	3,4	3,4	100,0
Gesamt	41.377.911	100,0	100,0	

Tab. 4: EF1 Bundesland

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Schleswig-Holstein	1.469.934	3,6	3,6	3,6
	Hamburg	1.003.288	2,4	2,4	6,0
	Niedersachsen	3.971.966	9,6	9,6	15,6
	Bremen	366.031	,9	,9	16,5
	Nordrhein-Westfalen	8.755.844	21,2	21,2	37,6
	Hessen	3.090.521	7,5	7,5	45,1
	Rheinland-Pfalz	1.960.746	4,7	4,7	49,8
	Baden-Württemberg	5.285.981	12,8	12,8	62,6
	Bayern	6.452.858	15,6	15,6	78,2
	Saarland	492.586	1,2	1,2	79,4
	Berlin	2.028.134	4,9	4,9	84,3
	Brandenburg	1.257.093	3,0	3,0	87,3
	Mecklenburg-Vorpommern	830.354	2,0	2,0	89,3
	Sachsen	2.156.441	5,2	5,2	94,6
	Sachsen-Anhalt	1.151.021	2,8	2,8	97,3
	Thüringen	1.103.789	2,7	2,7	100,0
	Gesamt	41.377.911	100,0	100,0	

Im Anhang befinden sich die oben aufgeführten und weitere Tabellen mit der entsprechenden Syntax für das Programm SPSS und eine Auflistung der Fundstellen in den Fachserien des Statistischen Bundesamtes.

2.4 Auswertbare regionale Ebene

- Bundesebene (NUTS 0)
- Landesebene (NUTS 1)

- Regierungsbezirke/Statistische Bezirke (NUTS 2)
- Kreisebene (ggf. eingeschränkte Aussagekraft in kleinen Landkreisen) (NUTS 3)
- Gemeindeebene (nur zu methodischen Zwecken, nicht zur Ergebnisdarstellung empfohlen⁴) (LAU)
- Anpassungsschichten (regionale Schicht und regionale Untergruppe)⁵
- Raumordnungsregionen
- Planungsregionen
- Arbeitsmarktregionen
- Nielsegebiete
- Arbeitsagenturbezirke

2.5 Produktversionen

1.0.0 (DOI: 10.21242/12211.2018.00.00.1.1.0)

- Erstveröffentlichung

1.1.0 (DOI: 10.21242/12211.2018.00.00.1.1.1)

- EF63 zu Verknüpfungszwecken aufgenommen
- EF5 durch EF5U1 ersetzt
- EF5U2=MISSING (kein Überhang) durch EF5U2=0 ersetzt
- Anpassungsschichten (EF564) der Region Hannover eindeutig benannt

1.2.0 (DOI: 10.21242/12211.2018.00.00.1.1.2)

- Missings kodiert
- EF313 und EF328 neue Labels
- Umschlüsselung der Hauptfachrichtungen auf HFR 2003 (EF313A und EF328A)
- EF376 bis EF378 aufgenommen
- EF637 bis EF652 aufgenommen

⁴ 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.

⁵ 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.

- Informationen aus EF2817 und EF2819 auf alle Haushaltsmitglieder übertragen
- Fälle ohne Befragung aufgenommen

1.3.0 (DOI: 10.21242/12211.2018.00.00.1.1.3)

- EF154 (F68 Schichtarbeit in den letzten 4 Wochen) 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

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 den Haupteinkommensbeziehern im Haushalt (EF662=1-4) vorgenommen werden. Alternativ dazu ist es möglich eine eindeutige Haushalts-ID über die Aneinanderreihung der Eingabefelder Bundesland (EF1), Auswahlbezirk (EF3), Haushaltsnummer (EF4) und Kennzeichnung der Überhänge (EF5U2) zu generieren und das Datenmaterial anhand dieser ID über alle Personen des Haushalts zu aggregieren. EF5U2 beinhaltet eine 0 bei Nicht-Überhanginterviews.

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

Die Verwendung von EF5U2 ist notwendig, da es Befragungen als Überhang aus der Stichprobe des Vorjahres Haushalte gibt, die zweimal in einem Kalenderjahr befragt wurden. Diese müssen, sofern sie in die Analyse eingehen sollen als separate Haushalte betrachtet werden, da sonst die Zuordnung von Familienmitgliedern zu einer bestimmten Befragung nicht möglich wäre.

Über die Eingabefelder EF53 und EF55 besteht die Möglichkeit die Personnummer der Mutter bzw. des Vaters zu ermitteln. Die zugehörigen Eltern im Haushalt können nun über EF5U1 im Haushalt identifiziert 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?

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. Ein tiefregionalisierter Hochrechnungsfaktor steht für die Quartale nicht zur Verfügung. Eine ausführliche Erläuterung der Modelle für die einzelnen Faktoren ist dem Metadatenreport Teil I in Abschnitt 2.6 zu entnehmen.

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

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)⁶:

```
**** 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 Prozente
Gültig	männlich	40.405.761	49,5	49,5	49,5
	weiblich	41.206.824	50,5	50,5	100,0
Gesamt		81.612.585	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.
```

⁶ Vgl. Zahlen in DESTATIS Fachserie 1, Reihe 4.1, 2018
https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEHeft_derivate_00055421/2010410187004_korr24032020.pdf Tabellen 1.2, 4.1 und 4.3

VARIABLE LABELS
 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 Prozente
Gültig	ledig	33.659.677	41,2	41,2	41,2
	verheiratet (inkl. eingetragener Lebenspartnerschaften)	36.892.280	45,2	45,2	86,4
	geschieden / verwitwet	11.060.629	13,6	13,6	100,0
	Gesamt	81.612.585	100,0	100,0	

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

RECODE EF401 (7=2) (5 THRU 9=5) (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 Prozente
Gültig	Erwerbs-/Berufstätigkeit	37.963.676	46,5	46,5	46,5
	Arbeitslosengeld/-hilfe	3.257.067	4,0	4,0	50,5
	Rente, Pension	17.947.713	22,0	22,0	72,5
	Unterhalt durch Eltern, Ehemann/- frau oder andere Angehörige	19.690.347	24,1	24,1	96,6
	eigenes Vermögen, Vermietung, Zinsen, Altenteil	2.753.782	3,4	3,4	100,0
	Gesamt	81.612.585	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 Prozente
Gültig	Erwerbstätige	41.894.706	51,3	51,3	51,3
	Erwerbslose	1.465.313	1,8	1,8	53,1
	Nichterwerbspersonen	38.252.566	46,9	46,9	100,0
	Gesamt	81.612.585	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) (8=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 Beamten, Soldatinnen und Soldaten, Richterinnen und Richter,

Beamtenanwärterinnen und -anwärter'

5 'Angestellte (inkl. Beschäftigte mit Mini-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 Prozente
Gültig	1 Selbständige mit und ohne Beschäftigte	4.008.741	4,9	9,9	9,9
	3 unbezahlt mithelfende Familienangehörige	138.556	,2	,3	10,3
	4 Beamtinnen und Beamte, Soldatinnen und Soldaten, Richterinnen und Richter, Beamtenanwärterinnen und - anwärter	2.007.203	2,5	5,0	15,2
	5 Angestellte (inkl. Beschäftigte mit Mini-Job, ohne Auszubildende)	27.289.399	33,4	67,6	82,8
	6 Arbeiter/-innen	5.450.627	6,7	13,5	96,3
	13 'Auszubildende, Volontärinnen und Volontäre, Trainees, bezahlte Praktikantinnen und Praktikanten'	1.500.090	1,8	3,7	100,0
	Gesamt	40.394.617	49,5	100,0	
Fehlend	System	41.217.969	50,5		
Gesamt		81.612.585	100,0		

Privathaushalte am Haupt- und Nebenwohnsitz⁷:

**** 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 Prozente
Gültig	Einpersonenhaushalte	17.333.411	41,9	41,9	41,9
	Zweipersonenhaushalte	13.983.135	33,8	33,8	75,7
	Dreipersonenhaushalte	4.923.446	11,9	11,9	87,6
	Vierpersonenhaushalte	3.748.344	9,1	9,1	96,6
	Fünf- und Mehrpersonenhaushalte	1.389.576	3,4	3,4	100,0
	Gesamt	41.377.911	100,0	100,0	

⁷ Vgl. Zahlen in DESTATIS Fachserie 1, Reihe 3, 2018
https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEHeft_derivate_00054484/2010300187004.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 Prozente
Gültig	Schleswig-Holstein	1.469.934	3,6	3,6	3,6
	Hamburg	1.003.288	2,4	2,4	6,0
	Niedersachsen	3.973.289	9,6	9,6	15,6
	Bremen	366.031	,9	,9	16,5
	Nordrhein-Westfalen	8.755.844	21,2	21,2	37,6
	Hessen	3.090.521	7,5	7,5	45,1
	Rheinland-Pfalz	1.960.746	4,7	4,7	49,8
	Baden-Württemberg	5.285.981	12,8	12,8	62,6
	Bayern	6.452.858	15,6	15,6	78,2
	Saarland	492.586	1,2	1,2	79,4
	Berlin	2.028.134	4,9	4,9	84,3
	Brandenburg	1.257.093	3,0	3,0	87,3
	Mecklenburg-Vorpommern	830.354	2,0	2,0	89,3
	Sachsen	2.156.441	5,2	5,2	94,6
	Sachsen-Anhalt	1.151.021	2,8	2,8	97,3
	Thüringen	1.103.789	2,7	2,7	100,0
	Gesamt	41.377.911	100,0	100,0	


```

**** Umkodieren von EF735 entsprechend der Publikation umkodiert.
*****
RECODE EF735
(6=2) (2=2) (8=4) (9=5) (3=2) (7=2) (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'
    4 'geschieden/ Lebenspartnerschaft aufgehoben'
    5 'verwitwet/ Lebenspartner verstorben'.

EXECUTE.

FREQUENCIES
VARIABLES= EF735NEU
/ORDER= ANALYSIS.

```

EF735NEU Familienstand des Haupteinkommensbeziehers

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	ledig	12.794.917	30,9	30,9	30,9
	verheiratet	18.880.505	45,6	45,6	76,6
	geschieden/ Lebenspartnerschaft aufgehoben	4.884.339	11,8	11,8	88,4
	verwitwet/ Lebenspartner verstorben	4.818.151	11,6	11,6	100,0
	Gesamt	41.377.911	100,0	100,0	

Statistische Ämter des Bundes und der Länder,
Metadatenreport – Teil II: Produktspezifische Informationen zur Nutzung des Mikrozensus 2018 am
Gastwissenschaftsarbetsplatz sowie per kontrollierter Datenfernverarbeitung (sog. On-Site-Zugangswege/-
Nutzungen)

Fotorechte Umschlag: ©artSILENCEcom – Fotolia.com