

Statistisches Bundesamt (Destatis) (Ed.)

**Periodical Part**

WISTA – Wirtschaft und Statistik, Ausgabe 1/2023

WISTA - Wirtschaft und Statistik

**Provided in Cooperation with:**

Statistisches Bundesamt (Destatis), Wiesbaden

*Suggested Citation:* Statistisches Bundesamt (Destatis) (Ed.) (2023) : WISTA – Wirtschaft und Statistik, Ausgabe 1/2023, WISTA - Wirtschaft und Statistik, ISSN 1619-2907, Statistisches Bundesamt (Destatis), Wiesbaden, Vol. 75, Iss. 1

This Version is available at:

<https://hdl.handle.net/10419/268963>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

---

# WISTA

## Wirtschaft und Statistik

---

Janina Hundenborn | Tobias Hees

Neue Statistik untergebrachter wohnungsloser Personen in Deutschland

Jan Eberle

Aufbereitung von Angaben zu Geburtsorten im Ausländerzentralregister mit OpenStreetMap

Johannes Weber

Bedeutung raumstruktureller und arbeitsmarktrelevanter Faktoren bei innerdeutschen Wanderungen von Geflüchteten

Shari Stehrenberg | Johannes Volk

Trusted Smart Surveys: Fragebogen trifft neue digitale Daten

Arijana Amina Ramic

Modernisierung der Statistikproduktion durch die neue Geheimhaltungs-Informationsplattform

Göran Kauermann | Giacomo De Nicola

Übersterblichkeit durch Corona?

---

# 1 | 2023

#### ABKÜRZUNGEN

|          |                                               |
|----------|-----------------------------------------------|
| D        | Durchschnitt (bei nicht addierfähigen Größen) |
| Vj       | Vierteljahr                                   |
| Hj       | Halbjahr                                      |
| a. n. g. | anderweitig nicht genannt                     |
| o. a. S. | ohne ausgeprägten Schwerpunkt                 |
| Mill.    | Million                                       |
| Mrd.     | Milliarde                                     |

#### ZEICHENERKLÄRUNG

|          |                                                                                            |
|----------|--------------------------------------------------------------------------------------------|
| –        | nichts vorhanden                                                                           |
| 0        | weniger als die Hälfte von 1 in der letzten besetzten Stelle, jedoch mehr als nichts       |
| .        | Zahlenwert unbekannt oder geheim zu halten                                                 |
| . . .    | Angabe fällt später an                                                                     |
| X        | Tabellenfach gesperrt, weil Aussage nicht sinnvoll                                         |
| I oder — | grundsätzliche Änderung innerhalb einer Reihe, die den zeitlichen Vergleich beeinträchtigt |
| /        | keine Angaben, da Zahlenwert nicht sicher genug                                            |
| ( )      | Aussagewert eingeschränkt, da der Zahlenwert statistisch relativ unsicher ist              |
|          | Abweichungen in den Summen ergeben sich durch Runden der Zahlen.                           |
|          | Tiefer gehende Internet-Verlinkungen sind hinterlegt.                                      |

# INHALT

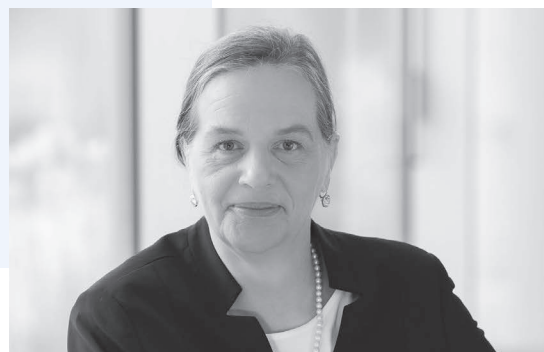
|    |                                                                                                                                                                                                                                                          |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 3  | Editorial                                                                                                                                                                                                                                                |
| 4  | Kennzahlen                                                                                                                                                                                                                                               |
| 8  | Aktuelle Informationsangebote                                                                                                                                                                                                                            |
| 10 | Kurznachrichten                                                                                                                                                                                                                                          |
| 15 | Janina Hundenborn, Tobias Hees<br><b>Neue Statistik untergebrachter wohnungsloser Personen in Deutschland</b><br><i>New statistics of homeless people accommodated in Germany</i>                                                                        |
| 27 | Jan Eberle<br><b>Aufbereitung von Angaben zu Geburtsorten im Ausländerzentralregister mit OpenStreetMap</b><br><i>Using OpenStreetMap to process data on birthplaces held in the Central Register of Foreigners</i>                                      |
| 43 | Johannes Weber<br><b>Bedeutung raumstruktureller und arbeitsmarktrelevanter Faktoren bei innerdeutschen Wanderungen von Geflüchteten</b><br><i>Importance of spatial-structural and labour market relevant factors in internal migration of refugees</i> |
| 59 | Shari Stehrenberg, Johannes Volk<br><b>Trusted Smart Surveys: Fragebogen trifft neue digitale Daten</b><br><i>Trusted Smart Surveys: questionnaire meets new digital data</i>                                                                            |

## INHALT

- |    |                                                                                                                                                                                                                                                     |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 69 | Arijana Amina Ramic<br><b>Modernisierung der Statistikproduktion durch die neue Geheimhaltungs-Informationsplattform</b><br><i>Modernisation of statistical production based on the new information platform for statistical disclosure control</i> |
| 80 | Göran Kauermann, Giacomo De Nicola<br><b>Übersterblichkeit durch Corona?</b><br><i>Excess mortality caused by Covid-19?</i>                                                                                                                         |

# EDITORIAL

Dr. Ruth Brand



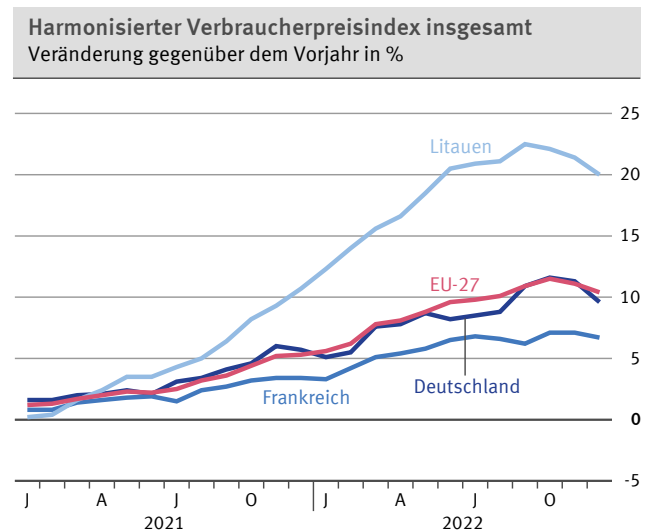
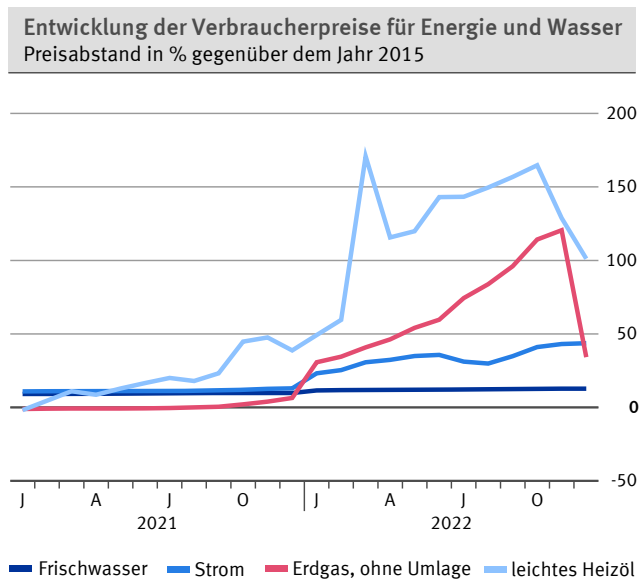
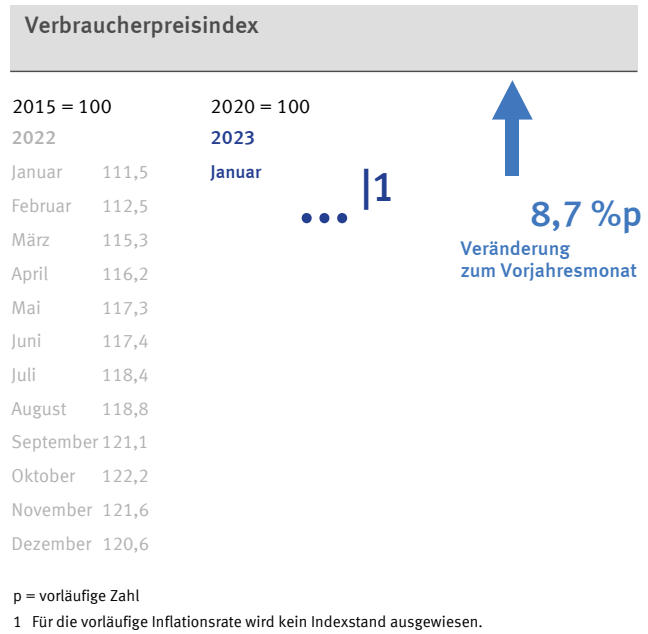
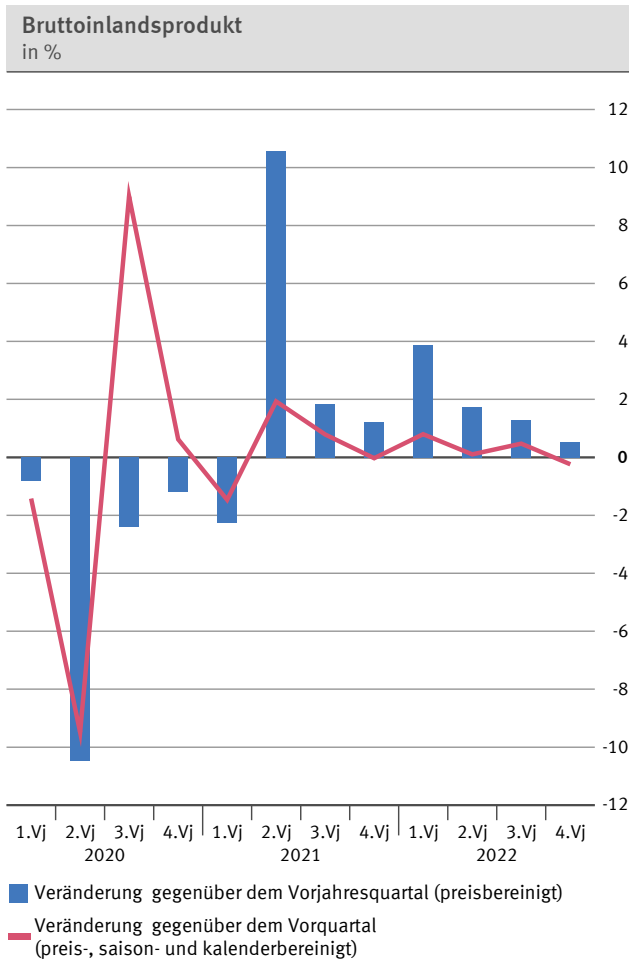
## LIEBE LESERIN, LIEBER LESER,

als neue Präsidentin des Statistischen Bundesamtes ist es mir persönlich eine große Freude, Ihnen beginnend mit dieser Ausgabe unseres Wissenschaftsmagazins regelmäßig dessen Inhalte vorzustellen und Ihr Interesse an der Lektüre zu verstärken. WISTA hat mich selbst schon lange in meinem Berufsleben begleitet: In meiner Zeit als Gruppenleiterin für die Landwirtschaftsstatistiken war ich Autorin mehrerer Artikel, als Abteilungsleiterin für die Bereiche Gesundheit, Soziales, Bildung und Private Haushalte habe ich die methodischen Beiträge in WISTA zu diesen Themen fachlich verantwortet.

Der erste Aufsatz in der vorliegenden Ausgabe knüpft direkt an einen Aspekt meines früheren Tätigkeitsbereichs an: Er informiert über die neue Statistik untergebrachter wohnungsloser Personen in Deutschland, die eine Datenlücke über Umfang und Ausmaß von Wohnungslosigkeit schließt und eine amtliche Datengrundlage für sozialpolitische Entscheidungen schafft. Welche Nutzungsmöglichkeiten Geodaten eröffnen, zeigen die nächsten beiden Artikel: zum einen eine Machbarkeitsstudie zur Aufbereitung und Geokodierung von Angaben zum Geburtsort im Ausländerzentralregister, zum anderen Analysen über den Zusammenhang zwischen raumstrukturellen sowie arbeitsmarktrelevanten Merkmalen und der Binnenmobilität von Geflüchteten mit Schutzstatus. Weitere Beiträge stellen das europäische Projekt ESSnet Smart Surveys, die im Statistischen Verbund entwickelte zentrale Geheimhaltungs-Informationenplattform sowie die im November 2022 mit dem Corona-Sonderpreis des Statistischen Bundesamtes ausgezeichneten Arbeiten vor.

Ich wünsche Ihnen eine erkenntnisreiche Lektüre!

Präsidentin des Statistischen Bundesamtes

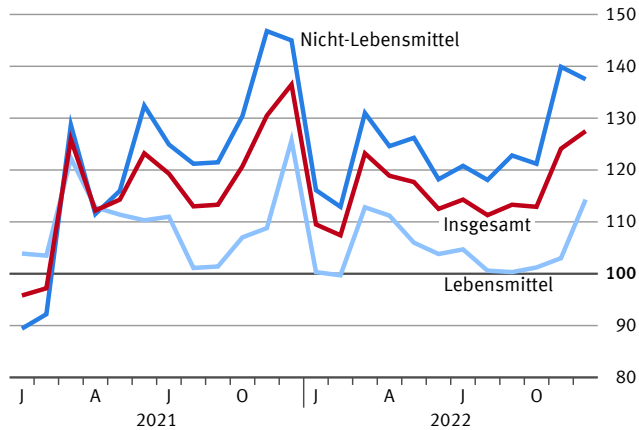


Dargestellt sind neben Deutschland und der Europäischen Union insgesamt (EU-27) die Länder mit der höchsten und der niedrigsten Veränderungsrate innerhalb der EU.

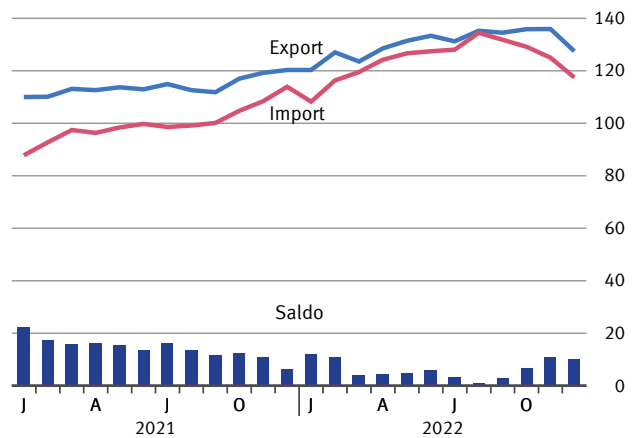
Stand: 10.02.2023

## Kennzahlen und Indikatoren

**Umsatz im Einzelhandel**  
2015 = 100

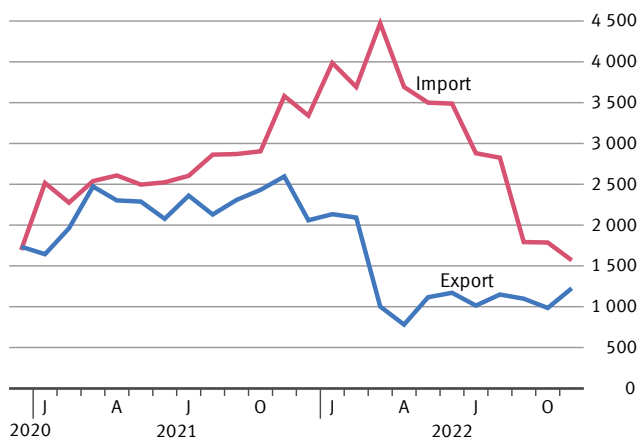


**Außenhandel**  
in Mrd. EUR

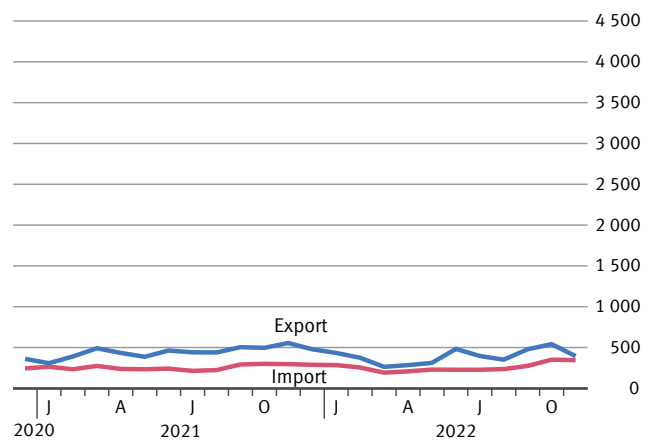


Kalender- und saisonbereinigte Werte nach dem Verfahren X13 JDemetra+. – Vorläufiges Ergebnis.

**Deutscher Außenhandel mit der Russischen Föderation**  
in Mill. EUR

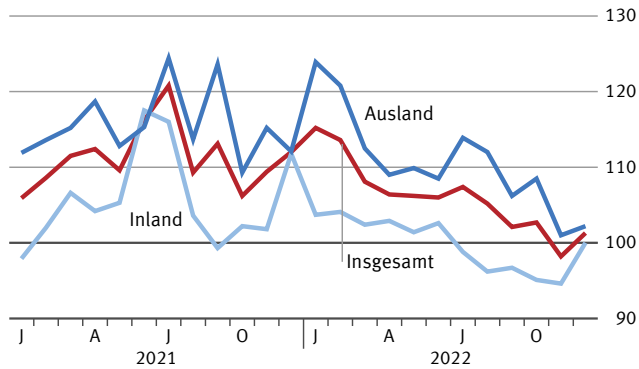


**Deutscher Außenhandel mit der Ukraine**  
in Mill. EUR



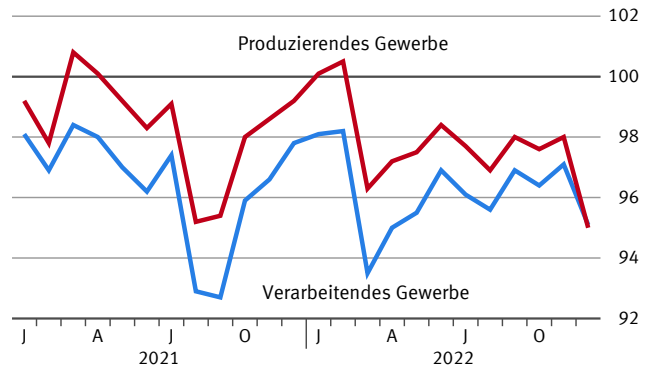
Stand: 10.02.2023

**Auftragseingang im Verarbeitenden Gewerbe**  
Volumenindex 2015 = 100



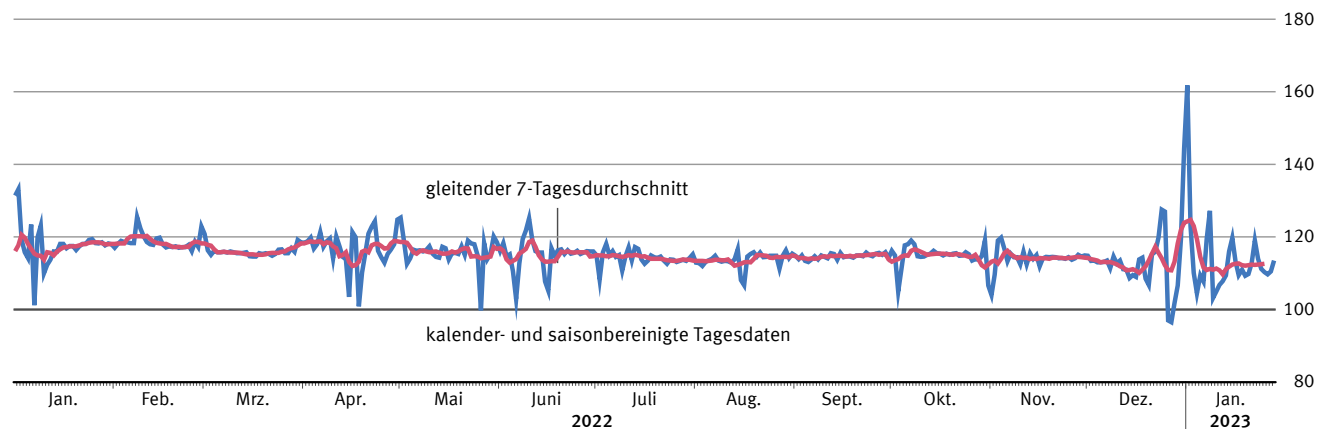
Kalender- und saisonbereinigter Wert nach dem Verfahren X13 JDemetra+. – Vorläufiges Ergebnis.

**Produktion im Produzierenden und Verarbeitenden Gewerbe**  
Index 2015 = 100



Kalender- und saisonbereinigte Werte nach dem Verfahren X13 JDemetra+. – Vorläufiges Ergebnis.

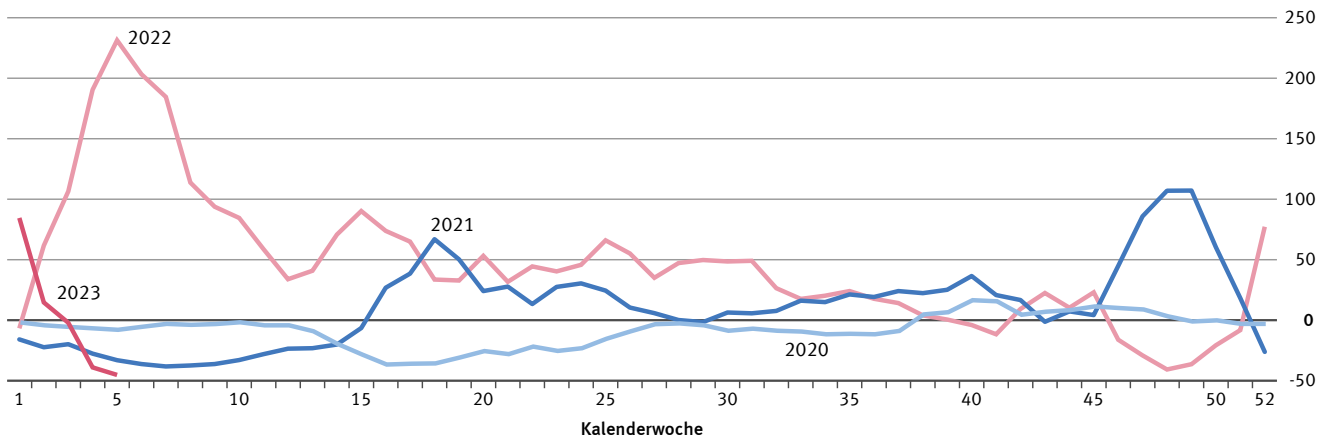
**Lkw-Maut-Fahrleistungsindex**  
2015 = 100



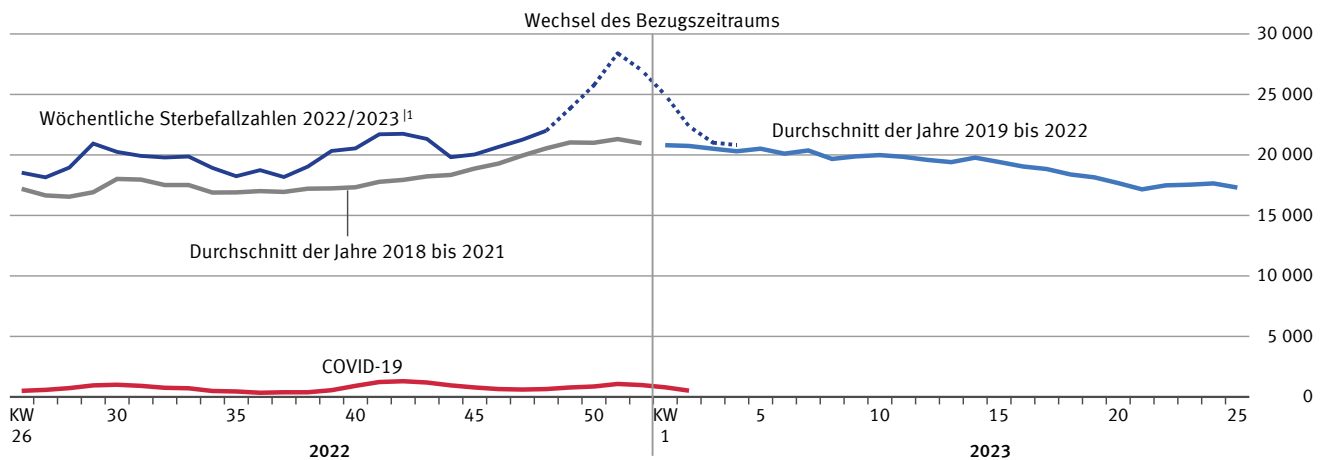
Quellen: Bundesamt für Logistik und Mobilität, Deutsche Bundesbank, Statistisches Bundesamt

### Neue Kreditverträge nach Kalenderwochen

Veränderung gegenüber der entsprechenden Vorjahreswoche in %



### Wöchentliche Sterbefallzahlen in Deutschland



Gestrichelte Werte enthalten Schätzanteil.

<sup>1</sup> Sonderauswertung der vorläufigen Sterbefallzahlen.

Quelle: Statistisches Bundesamt (Sterbefallzahlen insgesamt), Robert Koch-Institut (COVID-19-Todesfälle)

Stand: 10.02.2023



### Ukraine

Der Angriff Russlands auf die Ukraine und die damit verbundenen Sanktionen haben starke Auswirkungen auf Wirtschaft, Gesellschaft, Nahrungsversorgung und den Energiesektor. Auf einer Sonderseite zum Thema stellt das Statistische Bundesamt relevante Daten zur Verfügung. Über die Seite gelangt man auch zu Informationen und Hilfsangeboten für Geflüchtete, die vom Bundesministerium des Innern und für Heimat zusammengestellt wurden.

➤ [www.destatis.de/Im-Fokus/Ukraine](http://www.destatis.de/Im-Fokus/Ukraine)



### Dashboard Deutschland

Das vom Statistischen Bundesamt entwickelte Datenportal bietet hochaktuelle und hochfrequente Zahlen, Daten und Fakten zu den Themen Arbeitsmarkt, Bauen und Wohnen, Gesundheit, Konjunktur und Wirtschaft, Mobilität, Ukraine, Wertpapiere und Finanzen sowie Energie. Es trägt damit zu einem faktenbasierten demokratischen Diskurs der Öffentlichkeit und zur evidenzbasierten Entscheidungsfindung durch Politik und Verwaltung bei. Im Zusammenhang mit der derzeitigen Situation auf dem Energiemarkt wird das Datenangebot wöchentlich um Tagesdurchschnittspreise für Kraftstoffe erweitert.

➤ [www.dashboard-deutschland.de](http://www.dashboard-deutschland.de)



### EXDAT – Experimentelle Daten

In der Rubrik „EXDAT – Experimentelle Daten“ veröffentlicht das Statistische Bundesamt regelmäßig neue, innovative Projektergebnisse. Sie entstehen auf der Grundlage neuer Datenquellen und Methoden. Im Reifegrad unterscheiden sie sich von amtlichen Statistiken, insbesondere in Bezug auf Harmonisierung, Erfassungsbereich und Methodik. Doch die Corona-Pandemie hat gezeigt, dass experimentelle Daten gerade in Krisenzeiten eine wertvolle und unverzichtbare Ergänzung zum amtlichen Datenangebot sind, weil sie eine schnellere Berichterstattung ermöglichen.

➤ [www.destatis.de/exdat](http://www.destatis.de/exdat)

### im Fokus

#### Inflation – das statistische Angebot rund ums Thema

Die derzeit hohen Inflationsraten stehen im Fokus. Aktuelle Zahlen und Fakten sowie weiterführende Informationen stellt das Statistische Bundesamt auf der [Themenseite zur Inflationsrate auf Endverbraucherebene](#) zur Verfügung. Das Video „[Verbraucherpreisindex und Inflation kurz erklärt](#)“ bietet einen kurzen, kompakten Einstieg ins Thema. Und mithilfe des persönlichen [Inflationsrechners](#) kann ermittelt werden, wie sehr die persönliche von der amtlichen Teuerungsrate abweicht.



#### Corona-Statistiken – Auswirkungen der COVID-19-Pandemie

Die globale Ausbreitung des SARS-CoV-2-Virus und die damit verbundenen Maßnahmen zur Eindämmung der Pandemie treffen Gesellschaft und Wirtschaft stark. Die langfristigen Folgen der Corona-Pandemie sind noch nicht überschaubar. Die Sonderseite „Corona-Statistiken“ präsentiert statistische Ergebnisse, in denen sich bereits Auswirkungen zeigen oder in denen Auswirkungen zu erwarten sind. Dazu gehören auch neue Statistiken und Auswertungen, die das Statistische Bundesamt kurzfristig seit Beginn der Coronakrise bereitgestellt hat, um den Bedarf an Zahlen am aktuellen Rand zu decken:

➔ [www.destatis.de/corona](http://www.destatis.de/corona)



#### Einkommens- und Verbrauchsstichprobe 2023

Die Einkommens- und Verbrauchsstichprobe (EVS) ist die größte freiwillige Haushaltserhebung der amtlichen Statistik, sie wird alle fünf Jahre durchgeführt. Die Daten der EVS zeigen, wie viel Geld den Haushalten in Deutschland zur Verfügung steht und wofür sie es ausgeben. Die EVS zeichnet ein differenziertes Bild der Lebensbedingungen der Bevölkerung in Deutschland.

➔ [www.destatis.de](http://www.destatis.de)

---

# KURZNACHRICHTEN

---

## IN EIGENER SACHE

---

### Dr. Ruth Brand neue Präsidentin des Statistischen Bundesamtes

Zum 1. Januar 2023 wurde Dr. Ruth Brand zur Präsidentin des Statistischen Bundesamtes und zugleich zur Bundeswahlleiterin ernannt; sie folgt in beiden Ämtern Dr. Georg Thiel, der zum Jahresende 2022 in den Ruhestand gegangen ist. Die Bundesministerin des Innern und für Heimat, Nancy Faeser, hat Dr. Ruth Brand im Rahmen einer Festveranstaltung am 20. Januar 2023 in ihr Amt als Präsidentin des Statistischen Bundesamtes eingeführt.

Ministerin Faeser betonte, dass *„Statistiken [...] unverzichtbar für unsere Demokratie und für politische Entscheidungen [sind]. Die Informationen des Statistischen Bundesamtes genießen zu Recht hohes Vertrauen, weil sie unabhängig und qualitätsgesichert sind. Gerade in Krisenzeiten, wenn der Staat schnelle und weitreichende Entscheidungen treffen muss, zeigt sich, wie wichtig zuverlässige und vor allem aktuelle Informationen sind. Mit Dr. Ruth Brand gewinnt das Statistische Bundesamt eine erfahrene Statistikerin aus den eigenen Reihen zurück, die bereits fast 20 Jahre für das Haus gearbeitet hat. Sie steht für die große Verlässlichkeit, die das Statistische Bundesamt seit Jahrzehnten auszeichnet, und zugleich für weitere Innovationen und Digitalisierung.“*

Zu ihrem Amtsantritt hob Präsidentin Brand hervor, dass es *„unser Anspruch und Ansporn als Statistisches Bundesamt ist [...], der Garant für unabhängige qualitätsgesicherte Informationen in Deutschland zu sein. Wir wollen alle gesellschaftlichen Gruppen erreichen und zu einem faktenbasierten demokratischen Diskurs beitragen. Mit neuen Datenquellen, innovativen Methoden*

*und kreativen Ideen werden wir unsere Angebote in der kommenden Zeit stetig weiterentwickeln. Ich freue mich außerordentlich auf die Zusammenarbeit mit einem starken Team im Statistischen Bundesamt.“*

Die promovierte Wirtschaftswissenschaftlerin arbeitete von 2001 bis 2020 im Statistischen Bundesamt. Sie war hier in verschiedenen Leitungsfunktionen tätig, zuletzt als Abteilungsleiterin für die Statistiken im Bereich „Gesundheit, Soziales, Bildung und Private Haushalte“. Von 2020 bis 2022 leitete Dr. Ruth Brand das Beschaffungsamt des BMI.

### 75 Jahre Statistisches Bundesamt

Das Statistische Bundesamt feiert in diesem Jahr bereits seinen 75. Geburtstag: Am 21. Januar 1948 wurde mit dem Statistischen Amt des Vereinigten Wirtschaftsgebietes die Vorläuferinstitution des Statistischen Bundesamtes gegründet. Sitz des neuen Statistikamtes der amerikanischen und britischen Besatzungszonen mit zunächst 120 Mitarbeiterinnen und Mitarbeitern war Wiesbaden-Biebrich. Nach Gründung der Bundesrepublik Deutschland erhielt das Amt am 1. April 1950 den Namen Statistisches Bundesamt. Mittlerweile arbeiten mehr als 2 500 Beschäftigte an den Standorten Berlin, Bonn und Wiesbaden.

Das Statistische Bundesamt ist die erste Anlaufstelle in Deutschland, wenn es um neutrale, objektive und fachlich unabhängige Daten geht; es leistet einen wichtigen Beitrag für demokratische und faktenbasierte Entscheidungsprozesse. Dazu stellt es qualitativ hochwertige Informationen bereit und steigert weiter die Relevanz, die diese Daten für die Nutzerinnen und Nutzer haben.

Am 5. Juli 2023 wird ein gemeinsamer Festakt zum 75-jährigen Bestehen des Statistischen Bundesamtes

und zum 50-jährigen Bestehen des Bundesinstituts für Bevölkerungsforschung mit Gästen aus Politik, Ministerien, Wissenschaft sowie nationalen und internationalen Partnern stattfinden.

Fortgesetzt wird die Jubiläumsveranstaltung am 6. Juli mit einer wissenschaftlichen Fachtagung. Das Thema „Daten. Forschung. Zukunft“ wird in neun Sessions aus verschiedenen Blickwinkeln beleuchtet. Ein Auftaktvortrag nimmt insbesondere die Lehren aus der Vergangenheit auf, stellt die besondere Bedeutung der amtlichen Statistik für die Demokratie in der Vergangenheit, im Heute und in der Zukunft dar und verdeutlicht die Bedeutung unserer Daten und Ergebnisse als Grundlage für faktenbasierte Entscheidungsfindung in Politik, Wirtschaft und Gesellschaft.



Weitere Informationen zur wissenschaftlichen Fachtagung folgen in den nächsten Ausgaben dieser Zeitschrift.

## AUS DEM INLAND

---

### EVS 2023 gestartet – erstmals Teilnahme über App möglich

Unter dem Motto „Wo bleibt mein Geld?“ findet von Januar bis Dezember 2023 die Einkommens- und Verbrauchsstichprobe (EVS) statt. Die Daten dieser größten freiwilligen Haushaltserhebung der amtlichen Statistik bilden unter anderem die Grundlage für die Festsetzung von staatlichen Unterstützungsleistungen wie dem Kindergeld oder dem neuen Bürgergeld (ehemals „Hartz IV“). Darüber hinaus fließen die EVS-Daten als Basis für die Zusammensetzung des sogenannten Warenkorb in die Berechnung der Inflationsrate ein.

Die EVS findet alle fünf Jahre statt. Rund 80 000 Haushalte können sich an der Erhebung beteiligen. In den teilnehmenden Haushalten halten alle Personen ab

16 Jahre drei Monate lang ihre Ausgaben fest – von regelmäßigen festen Ausgaben (beispielsweise für Miete, Versicherungen und Mobilfunkverträge) bis hin zu täglichen Konsumausgaben (beispielsweise für Lebensmittel, Bekleidung und Freizeitaktivitäten). Zudem beantworten die Teilnehmenden Fragen unter anderem zur Haushaltszusammensetzung, zur Ausstattung mit bestimmten Gebrauchsgütern, zur Vermögenssituation sowie den Haushalts- und Personeneinkommen. Daraus gewinnt die amtliche Statistik insbesondere Daten zur Konsumstruktur der privaten Haushalte. Als Dankeschön für die Teilnahme erhalten die Haushalte eine Geldprämie von mindestens 100 Euro.

Erstmals ist die Teilnahme an der EVS über eine mobile App für Smartphones und Tablets oder eine Web-App für alle gängigen Internetbrowser möglich. Damit leistet das Statistische Bundesamt einen weiteren Beitrag zur Digitalisierung der Datenerhebung in der amtlichen Statistik. Über die App können die Teilnehmenden insbesondere die täglichen Ausgaben bequem auch von unterwegs dokumentieren. Bei aktiver Internetverbindung werden die Daten synchronisiert, sodass Eintragungen beispielsweise am Smartphone begonnen und über die Web-App am Computer weitergeführt werden können. Die mobile App ist im [Google Play Store](#) und im [Apple App Store](#) unter dem Namen „Einkommen & Verbrauch“ verfügbar. Die Web-App für alle gängigen Browser auf PC und Laptop steht unter [app-evs2023.de](http://app-evs2023.de) zur Verfügung. Zugangsdaten für die App erhalten nur Haushalte, die sich zur Teilnahme bereit erklärt haben und dafür nach dem Zufallsprinzip ausgewählt wurden. Eine Teilnahme über Papierfragebogen ist auf Wunsch ebenfalls möglich.

Um für die EVS 2023 neue Zielgruppen zu erreichen, die Teilnahmequote zu erhöhen und die App als neues Erhebungsinstrument zu etablieren, wird sie auch digital beworben. Motto und Motiv finden sich auf digitalen Werbebannern wieder, die sich insbesondere für Werbemaßnahmen über Social Media eignen. Weiterhin werden ein [Werbevideo](#) sowie ein [Erklärvideo](#) eingesetzt. Auch werden Influencer über Instagram für die Teilnahme an der EVS 2023 werben, aufbauend auf den bei der Zeitverwendungserhebung 2022 damit gesammelten Erfahrungen. Influencer erreichen über ihre eigenen Social-Media-Kanäle je nach Bekanntheit Tausende von Menschen aus unterschiedlichen Zielgruppen. So sollen insbesondere schwer zu erreichende Zielgruppen

für eine Teilnahme an der EVS 2023 gewonnen werden, beispielsweise einkommensschwache Haushalte mit Kindern. Die ausgewählten Influencer sollen Beiträge zur EVS 2023 schalten, um ihre Follower auf die EVS aufmerksam zu machen und zum Mitmachen zu bewegen. Dabei werden die Follower auf die EVS-Landingpage ([evs2023.de](https://evs2023.de)) gelenkt, mit weiteren Informationen und der Möglichkeit, sich zur Teilnahme anzumelden.

### Fachausschuss „Volkswirtschaftliche Gesamtrechnungen“

Ende November 2022 tagte der Fachausschuss „Volkswirtschaftliche Gesamtrechnungen (VGR)“ in Wiesbaden, erstmals wurde die Veranstaltung hybrid durchgeführt. Etwa 70 Personen aus verschiedenen Ministerien, der Deutschen Bundesbank, Wirtschaftsforschungsinstituten, Universitäten, Verbänden, den Statistischen Ämtern der Länder sowie Österreichs, der Schweiz und der Europäischen Union (Eurostat) nutzten die Gelegenheit, sich über die neuesten Entwicklungen in den Volkswirtschaftlichen Gesamtrechnungen zu informieren.

Wichtige Themen, über die ein Austausch erfolgte, waren die Vorbereitungen auf die nächste große Methodenrevision des System of National Accounts (SNA) 2025 sowie die Arbeit der Large Cases Unit (LCU). Daneben befassten sich Beiträge mit den regionalen und europäischen Volkswirtschaftlichen Gesamtrechnungen sowie mit nationalen Weiterentwicklungen im Bereich der Veröffentlichungen.

Der nächste VGR-Fachausschuss wird im Nachgang der nächsten VGR-Generalrevision 2024 stattfinden.

### Fachausschuss „Verkehrsstatistiken“

Ende Januar 2023 tagte der Fachausschuss „Verkehrsstatistiken“ in Wiesbaden, erstmals im hybriden Format mit insgesamt etwa 50 Teilnehmenden aus Politik, Wissenschaft, Wirtschaft und Statistik.

Externe Vortragende aus dem Bundesministerium für Digitales und Verkehr, dem Kraftfahrt-Bundesamt, dem Bundesamt für Logistik und Mobilität sowie der Bun-

desanstalt für Straßenwesen berichteten über die neuesten verkehrsstatistischen Projekte und Daten ihrer Institutionen.

Daneben lag der inhaltliche Schwerpunkt des Fachausschusses auf den anstehenden methodischen Weiterentwicklungen im Personennahverkehr mit Bussen und Bahnen, einer neuen Pilotstudie der Europäischen Union zur Personenbeförderung in der Binnenschifffahrt und aktuellen Herausforderungen bei der Verkehrsunfallstatistik. Zusätzlich wurden Erkenntnisse dargestellt, die aus den experimentellen Mobilitätsanalysen mit Mobilfunkdaten gewonnen werden konnten.

Weitere Kurzbeiträge befassten sich mit der Statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft (NACE Rev. 2.1), dem Sektorenwechsel des öffentlichen Personennahverkehrs (ÖPNV) zum Sektor Staat in den Volkswirtschaftlichen Gesamtrechnungen als Folge der Einführung des Deutschlandtickets sowie dem beim Digitalgipfel der Bundesregierung vorgestellten Mobilitätskompass.

### Projekt „Harmonisierung von Bodennutzungs\_codes“

Landwirtinnen und Landwirte liefern über den Antrag auf flächenbezogene Direktzahlungen im Rahmen der EU-Agrarförderung jedes Jahr zahlreiche Angaben zu ihren bewirtschafteten Flächen. Die Abwicklung der EU-Agrarförderung verläuft in Deutschland elektronisch über das Integrierte Verwaltungs- und Kontrollsystem (InVeKoS) der 13 Zahlstellen der Länder. Über dieses System werden sowohl die Anträge als auch die Zahlungen an die landwirtschaftlichen Betriebe verwaltet und kontrolliert. Für alle in Deutschland geförderten Kulturarten sind im InVeKoS Nutzcodes hinterlegt, die sich aufgrund von landesindividuellen Zuordnungen unterscheiden können. Letzteres führt bei der jährlichen statistischen Erhebung zur Flächennutzung (Bodennutzungshaupterhebung) dazu, dass eine rein auf Verwaltungsdaten basierende sekundärstatistische Erhebung zurzeit noch nicht möglich ist.

Das Projekt „Harmonisierung von Bodennutzungs\_codes“ hatte zum Ziel, die Nutzung von Verwaltungsdaten zu verbessern und eine bürokratische Entlastung sowohl für landwirtschaftliche Betriebe als auch für die Statis-

tischen Ämter des Bundes und der Länder zu erreichen. Das im Statistischen Bundesamt angesiedelte Dienstleistungszentrum der Bundesregierung für Bessere Rechtsetzung führte die Analysen durch und übernahm als neutraler Ansprechpartner die Aufgabe, einen für alle Stakeholder positiven Lösungsprozess anzustoßen und konstruktiv zu begleiten. Das Projekt wurde in Kooperation mit dem Bundeskanzleramt und dem Bundesministerium für Ernährung und Landwirtschaft durchgeführt.

Über eine speziell konzipierte Onlineerhebung wurden die für InVeKoS zuständigen Zahlstellen der Länder befragt. Um die notwendigen Anpassungen an den Nutzcodes seitens der Statistik umzusetzen, fand ein gemeinsamer Workshop statt. Ergebnis dieses Austauschs ist ein gemeinsamer, von allen Ländern getragener Beschlussvorschlag, welcher verschiedene Prüfungsaufträge und Maßnahmen für die Zahlstellen und eine Umsetzung ab dem Jahr 2023 vorsieht. Hierzu gehört auch die Informationsbeschaffung über eine zentrale Anlaufstelle (Single Point of Contact), was die Belastung der Landwirtschaftsbetriebe weiter reduzieren soll. Ein weiteres positives Ergebnis des Projektes ist die Fortsetzung des engen fachlichen Austauschs des Statistischen Bundesamtes mit den Zahlstellen der Länder über künftige Datenbedarfe in der amtlichen Statistik.

Der Projektbericht mit allen Ergebnissen steht zum Download zur Verfügung unter [www.destatis.de](https://www.destatis.de)

## NEUE INFORMATIONENANGEBOTE

---

### Pulsmesser für die Wirtschaft

Der Pulsmesser für die Wirtschaft erweitert das Dashboard Deutschland: Er bildet die Entwicklung aktueller Konjunkturindikatoren in Echtzeit ab und ermöglicht somit eine schnelle Einordnung der wirtschaftlichen Situation auf Basis amtlicher Daten des Statistischen Bundesamtes, ergänzt um schnell verfügbare nicht amtliche Daten.

➤ [www.dashboard-deutschland.de](https://www.dashboard-deutschland.de)

### Podcast „Arbeitskräftemangel: Welche Potenziale bietet eine alternde Gesellschaft?“

Der Arbeits- und Fachkräftemangel beschäftigt Wirtschaft, Politik und Gesellschaft mit zunehmender Intensität. Für den Mangel gibt es mehrere Ursachen. Dabei spielen auch demografische Entwicklungen eine wichtige Rolle, sowohl was entstehende Lücken betrifft, als auch wenn es um die Aktivierung möglicher gesellschaftlicher und demografischer Potenziale für den Arbeitsmarkt geht. Wer fehlt eigentlich – und warum? Wie entwickeln sich Arbeitszeit und Renteneintrittsalter? Welche Rolle spielen Geschlechterfragen oder die Zuwanderung? Darüber sprechen in der neuen Folge von „StatGespräch“, dem Podcast des Statistischen Bundesamtes, Prof. Dr. Norbert F. Schneider, Präsident der Deutschen Gesellschaft für Demographie und früherer Direktor des Bundesinstituts für Bevölkerungsforschung, und Frank Schüller, Leiter des Referats „Arbeitsmarkt“ im Statistischen Bundesamt.

➤ [www.destatis.de](https://www.destatis.de)

### Themenseite Bildungsindikatoren

Einen Überblick über die Bildungslandschaft in Deutschland bietet die neue Themenseite Bildungsindikatoren im Internetangebot des Statistischen Bundesamtes. Sie enthält ausgewählte statistische Ergebnisse, die anhand von interaktiven Grafiken und Tabellen zu zentralen Bildungsindikatoren aus den Bereichen Bildungsstand und Bildungsbeteiligung, Schule, Berufsbildung sowie Hochschule vorgestellt werden.

➤ [www.destatis.de](https://www.destatis.de)

### NEUERSCHEINUNGEN

---

#### Kulturfinanzbericht 2022

Die Gemeinschaftsveröffentlichung der Statistischen Ämter des Bundes und der Länder enthält neben Daten zu den öffentlichen Kulturausgaben von Bund, Ländern und Gemeinden auch Informationen zur Kulturförderung der Europäischen Union, zur privaten Kulturförderung und zur Künstlersozialkasse. Der Kulturfinanzbericht wird alle zwei Jahre im Auftrag der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Kultusministerkonferenz) und der Beauftragten der Bundesregierung für Kultur und Medien im Einvernehmen mit dem Deutschen Städtetag erstellt.

↳ [www.destatis.de](http://www.destatis.de)

#### Kulturindikatoren auf einen Blick. Ein Ländervergleich

Der Bericht „Kulturindikatoren auf einen Blick“ ist eine Gemeinschaftsveröffentlichung der Statistischen Ämter des Bundes und der Länder. Die Ausgabe 2022 enthält zahlreiche Kennzahlen und Indikatoren zum Kulturbetrieb in Deutschland im Jahr 2020, etwa zu den Erwerbstätigen in Kulturberufen, den Lernenden in Musikschulen oder der Ausstattung der privaten Haushalte mit Geräten der Unterhaltungselektronik.

↳ [www.destatis.de](http://www.destatis.de)

#### Energy Technology Perspectives 2023

Die COVID-19-Pandemie und der Angriff Russlands auf die Ukraine haben zu erheblichen Störungen der globalen Energie- und Technologieversorgungsketten geführt. Steigende Energie- und Materialpreise sowie Engpässe bei wichtigen Mineralien, Halbleitern und anderen Komponenten sind zum Risiko für die Energiewende geworden.

Die Veröffentlichung „Energy Technology Perspectives 2023“ der Internationalen Energieagentur (International Energy Agency – IEA) analysiert die Perspektiven für die Entwicklung und den Ausbau nachhaltiger und wider-

standsfähiger Energie- und Technologieversorgungsketten in den kommenden Jahren.

↳ [iea.blob.core.windows.net](https://iea.blob.core.windows.net)

# NEUE STATISTIK UNTERGEBRACHTER WOHNUNGSLOSER PERSONEN IN DEUTSCHLAND

Methode, Inhalt und erste Ergebnisse für 2022

Janina Hundenborn, Tobias Hees

🔗 **Schlüsselwörter:** Wohnungslosigkeit – Armuts- und Reichtumsberichterstattung – Übernachtungsgelegenheiten – Unterkünfte für Wohnungslose – Unterbringungsangebote

## ZUSAMMENFASSUNG

Die neue Statistik untergebrachter wohnungsloser Personen schließt eine Datenlücke über Umfang und Ausmaß von Wohnungslosigkeit und schafft eine amtliche Datengrundlage für sozialpolitische Entscheidungen. Die erstmalige Durchführung der Statistik ist erfolgreich verlaufen, dennoch werden erst weitere Erhebungen Auskunft über die Entwicklung von Wohnungslosigkeit in Deutschland geben und Verbesserungspotenziale in der Wohnungslosenberichterstattung aufzeigen können. Die Ergebnisse der neuen Statistik zeigen, dass es zum Stichtag 31. Januar 2022 in Deutschland etwa 178 000 untergebrachte wohnungslose Menschen gab.

🔗 **Keywords:** homelessness – reports on poverty and wealth – night shelters – living quarters for homeless people – accommodation offers

## ABSTRACT

*The new statistics of homeless people accommodated fill a gap in the data available on the extent and dimension of homelessness and create a basis of official data for social policy decisions. The first set of statistics has been compiled successfully. Further surveys will however be required to examine the development of homelessness in Germany and to identify possible improvements in the reports on homelessness. The results of the new statistics show that, as of 31 January 2022, about 178,000 homeless people were accommodated in Germany.*

### Dr. Janina Hundenborn

ist Volkswirtin und promovierte an der Universität Kapstadt (Südafrika) zur Evolution von Einkommensungleichheiten seit dem Ende der Apartheid. Ihre Aufgabenschwerpunkte als Referentin im Referat „Leistungen der Sozialhilfe“ des Statistischen Bundesamtes sind neben der Statistik untergebrachter wohnungsloser Personen die Wohnungsgeld- und Asylbewerberleistungsstatistiken.

### Tobias Hees

ist Soziologe und als wissenschaftlicher Mitarbeiter ebenfalls im Referat „Leistungen der Sozialhilfe“ des Statistischen Bundesamtes mit dem Schwerpunkt Konzeption und Aufbau der Statistik untergebrachter wohnungsloser Personen tätig.

## 1

### Einleitung

Wohnungslosigkeit ist eine gravierende Form von Armut und ein vielschichtiges Phänomen. Die Bundesregierung betrachtet das Ausmaß von Wohnungslosigkeit<sup>1</sup> regelmäßig in ihrem Armuts- und Reichtumsbericht, der Auskunft über die Lebenslagen in Deutschland gibt. Bislang mussten dafür Schätzungen genutzt werden, da auf Bundesebene keine belastbaren Daten zu Umfang und Struktur von Wohnungslosigkeit vorlagen. Um diese Datenlücke zu schließen, hat der Bundestag am 4. März 2020 das Wohnungslosenberichterstattungsgesetz (WoBerichtsG) verabschiedet und somit die rechtliche Grundlage für eine bundesweite Erfassung von Wohnungslosigkeit geschaffen.

Das Wohnungslosenberichterstattungsgesetz ordnet jährlich zum Stichtag 31. Januar die Durchführung einer Statistik über untergebrachte wohnungslose Personen durch das Statistische Bundesamt an, sowie alle zwei Jahre eine erweiterte Berichterstattung. Gemeinsam sollen die zentral durchgeführte Statistik und die durch das zuständige Ministerium<sup>2</sup> vorzulegende Berichterstattung sozialpolitisch fundierte Entscheidungen ermöglichen. Die neu geschaffene Datengrundlage kann somit einen wichtigen Beitrag zur Armuts- und Reichtumsberichterstattung des Bundes leisten.

Die ersten Ergebnisse der neuen Statistik zeigen, dass zum 31. Januar 2022 rund 178 000 wohnungslosen Personen Wohnräume oder Übernachtungsgelegenheiten durch Gemeinden und Gemeindeverbände oder mit Kostenerstattung durch andere Träger zur Verfügung gestellt worden sind. Zusätzlich hat die Statistik erstmals Informationen über die Art und Dauer der Unterbringung sowie über ausgewählte demografische Merkmale untergebrachter wohnungsloser Personen erhoben.

---

1 Die Begriffe Obdachlosigkeit und Wohnungslosigkeit wurden in Deutschland von offiziellen Stellen bis zum Jahr 2005 synonym verwendet. Im Jahr 2005 hat der Forschungsverbund „Wohnungslosigkeit und Hilfen in Wohnungsnotfällen“ die Definition aktualisiert, seitdem nutzen offizielle Stellen nur noch den Begriff Wohnungslosigkeit (Henke, 2020, hier: Seite 9 f.).

2 Zum Zeitpunkt der ersten Durchführung der Statistik war das Bundesministerium für Arbeit und Soziales zuständig, seit dem 1. Januar 2023 liegt die Zuständigkeit beim Bundesministerium für Wohnen, Stadtentwicklung und Bauwesen.

Die folgenden Kapitel ordnen die Statistik untergebrachter wohnungsloser Personen in die existierenden Studien zu Wohnungslosigkeit in Deutschland ein und geben einen Überblick über deren verschiedene Formen (Kapitel 2). Kapitel 3 erläutert, welche Stellen über die untergebrachten wohnungslosen Personen Auskunft geben und welche Merkmale erhoben werden. Kapitel 4 präsentiert die Ergebnisse zum Stichtag 31. Januar 2022, Kapitel 5 stellt zentrale Erkenntnisse aus der ersten Durchführung der Statistik vor. Der Beitrag schließt in Kapitel 6 mit einem Fazit und einem Ausblick auf die Zukunft der Wohnungslosenberichterstattung in Deutschland.

## 2

### Literaturübersicht und Hintergründe zur Statistik

Bis zur Einführung der Statistik untergebrachter wohnungsloser Personen waren vor allem Studien und Schätzungen der Gesellschaft für innovative Sozialforschung und Sozialplanung e.V. (GISS) und der Bundesarbeitsgemeinschaft Wohnungslosenhilfe e.V. (BAG Wohnungslosenhilfe) die Informationsgrundlage für die Bundesregierung zum Thema Wohnungslosigkeit (BMAS, 2021, hier: Seite 334 f.). Nach den letzten Schätzungen dieser Institutionen waren im Jahr 2018 zwischen 340 000 und 540 000 Menschen wohnungslos. Ursache für diese gravierenden Unterschiede sind vor allem die verschiedenen Ansätze für die Berechnungen: Die Schätzung der BAG Wohnungslosenhilfe (etwa 340 000 Menschen) basiert auf Daten des Landes Nordrhein-Westfalen sowie eigenen Annahmen und weist neben einer Jahresgesamtzahl seit 2017 auch einen Stichtagswert zum Stichtag 30. Juni eines Jahres aus (BAG Wohnungslosenhilfe, 2019). Der Schätzung der GISS (540 000 Menschen) für den Stichtag 31. Mai 2018 liegen Befragungsergebnisse aus 67 Kommunen verschiedener Größenklassen aus allen Bundesländern zugrunde (BMAS, 2021, hier: Seite 496). Somit betrachten diese Studien nicht nur unterschiedliche Stichtage, die Schätzung der BAG Wohnungslosenhilfe berücksichtigt zusätzlich auch Schätzungen über den Zeitraum eines Jahres.

Teilweise haben die bisherigen Studien der GISS oder der BAG Wohnungslosenhilfe folgende Personengruppen berücksichtigt (BMAS, 2021, hier: Seite 495), die allerdings nicht in der Statistik untergebrachter wohnungsloser Personen einbezogen sind:

- › Personen, die Beratungsangebote zum Thema Wohnungslosigkeit in Anspruch nehmen, aber am Stichtag nicht untergebracht sind,
- › Personen, die aufgrund einer angedrohten Zwangsräumung von Wohnungslosigkeit bedroht, aber (noch) nicht betroffen sind und
- › Personen, die Wohnungslosigkeit überwunden haben, aber noch Unterstützung durch das Hilfesystem erhalten.

Obwohl die Zahlen dieser Studien weit auseinanderliegen und auf Schätzungen beruhen, liefern die Untersuchungen wertvolle Einblicke in die vielschichtigen Aspekte von Wohnungslosigkeit. Da sich die neue Statistik auf untergebrachte wohnungslose Personen bezieht, ist sie dagegen auf diese Erscheinungsform beschränkt. Um dennoch die bestehende Datenlücke zu schließen, hat der Gesetzgeber eine erweiterte Berichterstattung vorgesehen. Ziel ist, Informationen und Analysen über Ausmaß und Struktur der Formen von Wohnungslosigkeit zu gewinnen, die über den Umfang der Erhebung hinausgehen (§ 8 Absatz 1 WoBerichtsG).

Im Auftrag des Bundesministeriums für Arbeit und Soziales (BMAS) hat die GISS deshalb in Zusammenarbeit mit Kantar Public Anfang Februar 2022 parallel zum Stichtag der Statistik Untersuchungen zu zusätzlichen Facetten von Wohnungslosigkeit durchgeführt. Entsprechend der Gesetzgebung standen wohnungslose Menschen ohne Unterkunft und verdeckt Wohnungslose, die bei Freunden oder Bekannten unterkommen, im Fokus dieser Studie. Die Ergebnisse der Befragungen und Hochrechnungen durch GISS/Kantar Public zeigen, dass knapp 87 000 Personen verdeckt oder ohne Unterkunft wohnungslos waren (BMAS, 2022a).

Die Ergebnisse der Studie der GISS/Kantar Public sowie aus der Statistik untergebrachter wohnungsloser Personen sind gemeinsam in den Wohnungslosenbericht eingeflossen, den das Bundesministerium für Arbeit und Soziales am 8. Dezember 2022 vorgestellt hat

(BMAS, 2022b).<sup>13</sup> Danach waren insgesamt knapp 263 000 Personen zum 31. Januar 2022 wohnungslos. Der Wohnungslosenbericht analysiert strukturelle Merkmale derjenigen Personen, die entweder verdeckt oder untergebracht wohnungslos waren beziehungsweise ohne Unterkunft wohnungslos waren; er diskutiert Empfehlungen zur Weiterentwicklung der Berichterstattung sowie künftige politische Handlungsansätze.

Die Wohnungslosenberichterstattung einschließlich der Statistik untergebrachter wohnungsloser Personen ist in Art und Umfang beispielgebend in Europa (siehe etwa Henke, 2022). Die Auswahl der Personengruppen, die im Wohnungslosenbericht, also in der Statistik und der erweiterten Berichterstattung, berücksichtigt werden, richtet sich auch nach der europäischen Typologie für Wohnungslosigkeit (ETHOS Light; FEANTAS, 2005). ETHOS Light ordnet die unterschiedlichen Erscheinungsformen von Wohnungslosigkeit sechs Gruppen zu. Von den sechs verschiedenen Kategorien werden zwei durch die erweiterte Berichterstattung erfasst (Kategorie I: Wohnungslose Menschen ohne Unterkunft sowie Kategorie VI: Verdeckt Wohnungslose). Zwei weitere Kategorien (Kategorie II: Menschen in Notunterkünften; Kategorie III ohne III.6: Menschen in speziellen Unterkünften für Wohnungslose) sind durch die Statistik abgedeckt. Nach der derzeitigen Gesetzgebung sind die Kategorien III.6, IV und V weder für die Statistik noch für die erweiterte Berichterstattung vorgesehen.<sup>14</sup> [↗ Übersicht 1](#)

Im Gesetzgebungsprozess wurde für die Entscheidung, welche Personengruppen in die Statistik und die erweiterte Berichterstattung einzubeziehen sind, die Expertise der freien Wohlfahrtsverbände genutzt, ebenso die von Expertinnen und Experten aus Wissenschaft und Praxis. Zudem hat ein wissenschaftlicher Beirat die erste Durchführung der Statistik begleitet.

In diesem Beirat waren neben Sachverständigen der GISS und der BAG Wohnungslosenhilfe das Bundesministerium für Arbeit und Soziales und das Statistische Bundesamt vertreten sowie verschiedene Verbände der freien Träger, verschiedene Landesministerien sowie der

3 Den nächsten Wohnungslosenbericht im Jahr 2024 wird das Bundesministerium für Wohnen, Stadtentwicklung und Bauwesen veröffentlichen.

4 Für diese Bereiche hat die GISS im Auftrag des Bundesministeriums für Arbeit und Soziales gemäß § 8 Absatz 4 WoBerichtsG eine Machbarkeitsstudie durchgeführt (BMAS, 2022c).

## Übersicht 1

### Kategorien nach der europäischen Typologie für Wohnungslosigkeit (ETHOS Light)

| Operative Kategorie                                                                                            | Wohnsituation                                                                   | Definition                                                                                                                                                    |
|----------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------|
| I Obdachlose Menschen                                                                                          | 1 Im öffentlichen Raum, in Verschlägen, unter Brücken usw.                      | Auf der Straße lebend, an öffentlichen Plätzen wohnend, ohne eine Unterkunft, die als solche bezeichnet werden kann                                           |
| II Menschen in Notunterkünften                                                                                 | 2 Notschlafstellen, Wärmestuben                                                 | Menschen ohne festen Wohnsitz, die in häufig wechselnden Notschlafstellen und niederschweligen Einrichtungen übernachten                                      |
| III Menschen, die in Wohnungsloseneinrichtungen wohnen                                                         | 3 Übergangswohnheime                                                            | Einrichtungen, bei denen die Aufenthaltsdauer begrenzt ist und keine Dauerwohnplätze zur Verfügung stehen                                                     |
|                                                                                                                | 4 Asyle und Herbergen                                                           |                                                                                                                                                               |
|                                                                                                                | 5 Übergangswohnungen                                                            |                                                                                                                                                               |
|                                                                                                                | 6 Frauenhäuser, Aufnahmeeinrichtungen für Flüchtlinge und andere Auffangstellen |                                                                                                                                                               |
| IV Menschen, die von Institutionen entlassen werden                                                            | 7 Medizinische Einrichtungen, Psychiatrie, Reha-Einrichtungen usw.              | Bleiben weiter hospitalisiert, weil kein Wohnplatz zur Verfügung steht                                                                                        |
|                                                                                                                | 8 Gefängnisse, Strafanstalten                                                   | Nach Haftentlassung kein ordentlicher Wohnsitz vorhanden                                                                                                      |
| V Menschen, die in Wohnprovisorien hausen                                                                      | 9 Wohnwagen                                                                     | Wohnen in Behausungen, die für konventionelles Wohnen nicht gedacht sind, weil kein eigener Wohnraum verfügbar ist, und ohne einen Hauptwohnsitz zu begründen |
|                                                                                                                | 10 Behausungen, die für konventionelles Wohnen nicht gedacht sind               |                                                                                                                                                               |
|                                                                                                                | 11 Zelte, vorübergehende Behausungen                                            |                                                                                                                                                               |
| VI Menschen, die in konventionellem Wohnraum wohnen als temporäre Unterkunft bei Freunden/Bekannten/Verwandten | 12 Regulärer Wohnraum, aber ohne einen Hauptwohnsitz zu begründen               | Wohnen in regulärem Wohnraum, aber ohne einen Hauptwohnsitz zu begründen und nur als vorübergehender Unterschlupf, weil kein eigener Wohnraum verfügbar ist   |

Quelle: ETHOS – European Typology on Homelessness and Housing Exclusion (entwickelt von der Nichtregierungsorganisation [NGO] FEANTSA, der European Federation of National Organisations Working with the Homeless); verfügbar unter: [www.feantsa.org](http://www.feantsa.org)

Deutsche Landkreistag und der Deutsche Städtetag.<sup>15</sup> Der Beirat unterstützte durch die Expertise seiner Mitglieder das Statistische Bundesamt und das Bundesministerium für Arbeit und Soziales außerdem in inhaltlichen Fragen (zum Beispiel zur Ausgestaltung einzelner Erhebungsmerkmale) und wurde auch in Bezug auf die Schlussfolgerungen aus der ersten Erhebung einbezogen.

<sup>15</sup> Im Einzelnen gehörten folgende Institutionen, Verwaltungen und Verbände dem Beirat an: Ministerium für Arbeit, Gesundheit und Soziales des Landes Nordrhein-Westfalen; Ministerium für Soziales, Arbeit, Gesundheit und Demografie des Landes Rheinland-Pfalz; die Stadt Bonn; die Zentrale Beratungsstelle Niedersachsen; der Deutsche Landkreistag; der Deutsche Städtetag; das Armutsnetzwerk e.V.; der AWO Bundesverband e.V.; die Bundesarbeitsgemeinschaft Wohnungslosenhilfe e.V.; der Deutsche Caritasverband e.V.; der Deutsche Verein für öffentliche und private Fürsorge e.V.; der Deutsche Paritätische Wohlfahrtsverband – Gesamtverband e.V.; die Diakonie Deutschland – Evangelisches Werk für Diakonie und Entwicklung e.V.; die Selbstvertretung wohnungsloser Menschen e.V.; der Sozialverband Deutschland e.V.; die Alice Salomon Hochschule Berlin; die Gesellschaft für innovative Sozialforschung und Sozialplanung e.V. (GISS); das Bundesministerium für Arbeit und Soziales; das Statistische Bundesamt.

## 3

### Inhalte der Statistik: Auskunftspflicht und Erhebungsmerkmale

Die gesellschaftliche Aufgabe, Wohnungslosigkeit zu vermeiden, liegt in Deutschland grundsätzlich bei den Kommunen, die nach den Ordnungs- und Polizeigesetzen der Länder verpflichtet sind, unfreiwillig wohnungslose Menschen unterzubringen (Deutscher Bundestag, 2021). Zur Statistik sind daher diejenigen Stellen auskunftspflichtig, die nach jeweiligem Landesrecht im örtlichen Zuständigkeitsbereich für die polizei- und ordnungsrechtliche Unterbringung von wohnungslosen Personen verantwortlich sind (§ 6 WoBerichtsG). In der Regel sind dies Ordnungsämter und Verwaltungen. Die Bundesregierung unterstützt dabei die Länder und Kommunen durch geeignete Maßnahmen basierend auf dem

Zweiten beziehungsweise Zwölften Buch Sozialgesetzbuch (SGB II beziehungsweise SGB XII). Daraus folgt, dass eine Vielzahl unterschiedlicher Stellen aus dem öffentlichen und privaten Bereich mit der Unterbringung von wohnungslosen Personen betraut ist. Die auskunftspflichtigen Stellen der Kommunen können daher entweder selbst die Daten aus allen relevanten Einrichtungen an das Statistische Bundesamt übermitteln oder die Auskunftspflicht an andere Stellen beziehungsweise Einrichtungen delegieren. Die Entscheidung darüber, welche Stellen Daten an das Statistische Bundesamt übermitteln, obliegt ausschließlich den Kommunen und Einrichtungen vor Ort.

Der Vorteil einer derart strukturierten Auskunftspflicht ist, dass die in erster Instanz auskunftspflichtigen Gemeinden einfach zu recherchieren sind und den Kommunen vor Ort in der Regel bekannt ist, welche Einrichtungen die dort ansässigen wohnungslosen Personen unterbringen. Eine amtliche Übersicht zu allen Stellen und Einrichtungen in Deutschland, die wohnungslose Personen unterbringen, existiert dagegen nicht. Somit erachtete es der Gesetzgeber für sinnvoll, den Kommunen die Entscheidung darüber zu überlassen, ob sie die Daten zu den am Stichtag untergebrachten wohnungslosen Personen selbst übermitteln oder die Auskunftspflicht delegieren.

Um ein Bild von Anzahl und Eigenschaften wohnungsloser Menschen in Deutschland zu bekommen, werden bestimmte Merkmale für jede untergebrachte wohnungslose Person erhoben. Zunächst wird über den Amtlichen Gemeindeschlüssel festgehalten, in welcher Gemeinde die Person untergebracht ist. Zusätzlich werden demografische Merkmale, wie das Lebensalter zum Stichtag der Erhebung, das Geschlecht oder die Staatsangehörigkeit, erfragt. Außerdem wird festgehalten, in welcher Haushaltskonstellation die Person untergebracht ist (Haushaltstyp und Haushaltgröße) und seit wann sich die Person in der aktuellen Unterbringung befindet. Diese Angaben sind für jede Person eines untergebrachten Haushalts individuell zu erfassen. Zusätzlich werden für jede Person Angaben zur Unterbringung erhoben, beispielsweise die Art der Überlassung von Räumen zu Wohnzwecken sowie der Anbieter von Unterbringungsangeboten.

Die beschriebenen Merkmale werden für alle Personen festgehalten, denen Wohnraum vorübergehend überlassen wird, ohne dass dadurch die Wohnungslosigkeit

beendet wird (§ 3 Absatz 2 WoBerichtsG). Die Überlassung von Wohnraum erfolgt, wenn Personen zum Stichtag ordnungsrechtlich untergebracht sind, im Rahmen von Maßnahmen zur Überwindung besonderer sozialer Schwierigkeiten nach den §§ 67 ff. SGB XII oder durch Maßnahmen der Gemeinden oder Gemeindeverbänden mithilfe von anderen Finanzierungsinstrumenten – wie etwa zuwendungsrechtliche Förderungen von Kältehilfen. Dies umfasst Personen, die in Not- und Gemeinschaftsunterkünften oder gegebenenfalls auch gewerblichen Unterkünften und Normalwohnraum untergebracht sind.

Personen, die zwar in einer Einrichtung untergebracht sind, das Ziel dieser Unterbringung aber nicht die Abwendung von Wohnungs- oder Obdachlosigkeit ist, sind nicht in der Statistik untergebrachter wohnungsloser Personen nachzuweisen. Hierzu zählen Bewohnerinnen und Bewohner von Pflegeeinrichtungen, Heimen für Menschen mit Behinderung, Frauenhäusern, Suchtkliniken oder Einrichtungen des betreuten Wohnens der Jugendhilfe. Außerdem werden geflüchtete Personen nicht in der Statistik erfasst, wenn sie als Schutzsuchende über das Asylbewerberleistungsgesetz untergebracht sind. Dies betrifft Personen, deren Asylverfahren noch nicht oder mit negativem Ergebnis abgeschlossen wurde, und geflüchtete Personen, die aus anderen Gründen außerhalb des Wohnungslosenhilfesystems untergebracht sind.

Geflüchtete Personen mit einem positiven Abschluss des Asylverfahrens verfügen prinzipiell über das Recht, sich in Deutschland Wohnraum zu suchen. Dennoch verbleiben auch anerkannte Geflüchtete oft in einer Aufnahmeeinrichtung, wenn sie keine Wohnung finden. Werden also anerkannte Geflüchtete vorübergehend zur Vermeidung von Wohnungslosigkeit weiterhin in einer Gemeinschaftsunterkunft oder in einer Aufnahmeeinrichtung untergebracht, sind sie in der Statistik zu erfassen, obwohl das primäre Ziel dieser Einrichtungen nicht die Abwendung von Wohnungslosigkeit ist.

Wie für alle anderen Personengruppen gilt auch für Geflüchtete, dass sie in der Statistik zu berücksichtigen sind, wenn sie durch das „Wohnungsnotfallhilfesystem“ untergebracht sind. Das bedeutet, dass Geflüchtete, die durch die oben beschriebenen Maßnahmen des Polizei- und Ordnungsrechts, durch Angebote nach §§ 67 ff. SGB XII oder in Einrichtungen der Wohnungsnotfallhilfe untergebracht sind, in der Statistik zu erfassen sind.

## 4

## Ergebnisse zum 31. Januar 2022

Die erste Erhebung der Statistik untergebrachter wohnungsloser Personen hat ergeben, dass in Deutschland zum Stichtag 31. Januar 2022 insgesamt 178 145 Personen wegen Wohnungslosigkeit untergebracht waren.

Der Vergleich nach Bundesländern zeigt, dass mit jeweils knapp 36 000 Personen die meisten Menschen in Baden-Württemberg und Nordrhein-Westfalen wegen Wohnungslosigkeit untergebracht waren, gefolgt von Berlin mit rund 26 000 Personen. Die Bundesländer Sachsen-Anhalt (365 Personen), Mecklenburg-Vorpommern (405 Personen) und Bremen (790 Personen) haben die wenigsten untergebrachten wohnungslosen Personen gemeldet. [Tabelle 1](#)

Auf Bundesebene waren 61,9 % der untergebrachten Personen männlich und 36,8 % weiblich. In 1,3 % der Fälle wurde das Geschlecht mit „unbekannt“ angegeben, weil den Berichtsstellen hierzu keine Informationen vorlagen.<sup>6</sup> Die meisten Fälle, in denen das Geschlecht

6 Davon zu unterscheiden sind Personen, die nach § 22 Absatz 3 Personenstandsgesetz die Geschlechtsangabe „divers“ und „ohne Angabe“ tragen. Diese wurden aus Gründen der statistischen Geheimhaltung nach dem Zufallsprinzip dem männlichen oder weiblichen Geschlecht zugeordnet.

unbekannt war, traten in Berlin auf. In anderen Bundesländern kam dies kaum bis gar nicht vor.

Auf die zehn größten Städte Deutschlands entfielen etwa 44 % aller untergebrachten wohnungslosen Personen. Berlin hat zwar die meisten untergebrachten wohnungslosen Personen gemeldet, jedoch wies Hamburg mit 1,02 % den höchsten Anteil untergebrachter wohnungsloser Personen an der Bevölkerung auf. Dies ist auch darauf zurückzuführen, dass ein großer Teil der Schutzsuchenden in Hamburg über das Ordnungsrecht untergebracht wird. Durchschnittlich lag der Anteil von untergebrachten wohnungslosen Menschen an der Bevölkerung mit 0,66 % in den Großstädten deutlich über dem Bundesdurchschnitt von 0,21 %. [Tabelle 2](#)

Zum Stichtag 31. Januar 2022 besaßen 31 % der gemeldeten Personen die deutsche Staatsangehörigkeit, bei 69 % lag eine andere Staatsangehörigkeit vor oder die Angaben zu diesem Merkmal waren nicht bekannt.<sup>7</sup> Aus dem europäischen Ausland stammten 16,6 % aller erfassten Personen, darunter 9,0 % aus der Europäischen Union (EU). Rund 40 % der untergebrachten wohnungslosen Personen wiesen die Staatsangehörigkeiten von Ländern auf, aus denen ein großer Teil der Asyl-

7 Personen, deren Staatsangehörigkeit mit „unbekannt“, „ungeklärt“ oder „staatenlos“ signiert wurde.

**Tabelle 1**

Untergebrachte wohnungslose Personen am 31. Januar 2022

|                        | Insgesamt |      | Männlich <sup>1</sup> | Weiblich <sup>1</sup> |
|------------------------|-----------|------|-----------------------|-----------------------|
|                        | Anzahl    | %    |                       |                       |
| Insgesamt              | 178 145   | 100  | 61,9                  | 36,8                  |
| Schleswig-Holstein     | 8 555     | 4,8  | 61,6                  | 38,2                  |
| Hamburg                | 18 915    | 10,6 | 58,5                  | 41,5                  |
| Niedersachsen          | 10 860    | 6,1  | 62,5                  | 37,5                  |
| Bremen                 | 790       | 0,4  | 78,2                  | 21,8                  |
| Nordrhein-Westfalen    | 35 815    | 20,1 | 62,7                  | 36,2                  |
| Hessen                 | 12 110    | 6,8  | 61,3                  | 38,6                  |
| Rheinland-Pfalz        | 5 830     | 3,3  | 61,7                  | 38,2                  |
| Baden-Württemberg      | 35 905    | 20,2 | 61,3                  | 38,6                  |
| Bayern                 | 17 910    | 10,1 | 64,9                  | 34,8                  |
| Saarland               | 815       | 0,5  | 59,7                  | 40,3                  |
| Berlin                 | 25 975    | 14,6 | 60,3                  | 32,6                  |
| Brandenburg            | 1 295     | 0,7  | 68,8                  | 31,2                  |
| Mecklenburg-Vorpommern | 405       | 0,2  | 79,2                  | 20,8                  |
| Sachsen                | 1 665     | 0,9  | 71,9                  | 28,1                  |
| Sachsen-Anhalt         | 365       | 0,2  | 75,3                  | 24,7                  |
| Thüringen              | 940       | 0,5  | 64,3                  | 35,7                  |

1 Personen mit den Angaben „divers“ und „ohne Angabe“ (nach § 22 Absatz 3 Personenstandsgesetz) werden aus Gründen der statistischen Geheimhaltung per Zufallsprinzip dem männlichen oder weiblichen Geschlecht zugeordnet.

## Neue Statistik untergebrachter wohnungsloser Personen in Deutschland

**Tabelle 2**

### Untergebrachte wohnungslose Personen in den zehn größten Städten Deutschlands

|                      | Einwohnerzahl<br>31. Dezember 2021 | Untergebrachte Wohnungslose<br>31. Januar 2022 | Anteil an der Bevölkerung<br>in % |
|----------------------|------------------------------------|------------------------------------------------|-----------------------------------|
| Berlin               | 3 677 472                          | 25 975                                         | 0,71                              |
| Hamburg              | 1 853 935                          | 18 915                                         | 1,02                              |
| München              | 1 487 708                          | 8 975                                          | 0,60                              |
| Köln                 | 1 073 096                          | 7 235                                          | 0,67                              |
| Frankfurt am Main    | 759 224                            | 5 675                                          | 0,75                              |
| Stuttgart            | 626 275                            | 5 350                                          | 0,85                              |
| Düsseldorf           | 619 477                            | 3 335                                          | 0,54                              |
| Leipzig              | 601 866                            | 550                                            | 0,09                              |
| Dortmund             | 586 852                            | 1 375                                          | 0,23                              |
| Essen                | 579 432                            | 765                                            | 0,13                              |
| Großstädte insgesamt | 11 865 337                         | 78 150                                         | 0,66                              |
| Deutschland          | 83 237 124                         | 178 145                                        | 0,21                              |

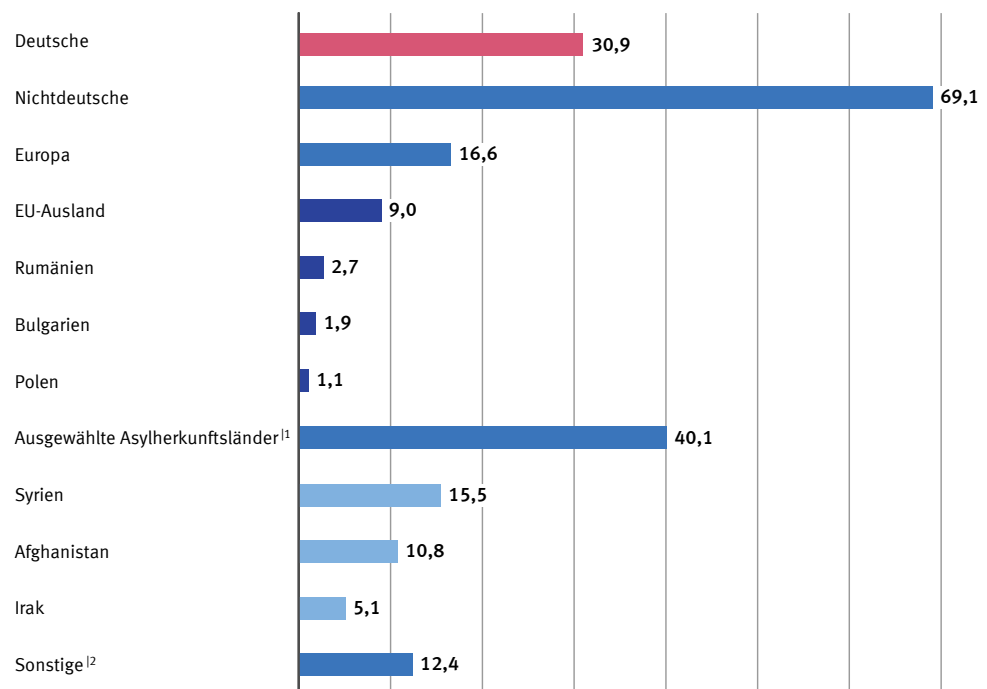
suchenden in Deutschland kommt.<sup>18</sup> So hatten 15,5 % aller gemeldeten untergebrachten wohnungslosen Per-

sonen die syrische und 10,8 % die afghanische Staatsangehörigkeit. [↗ Grafik 1](#)

8 Afghanistan, Eritrea, Irak, Iran, Nigeria, Pakistan, Somalia, Syrien.

**Grafik 1**

### Untergebrachte wohnungslose Personen nach Staatsangehörigkeiten in %



Stichtag: 31. Januar 2022.

<sup>1</sup> Afghanistan, Eritrea, Irak, Iran, Nigeria, Pakistan, Somalia, Syrien.

<sup>2</sup> Einschließlich Personen, deren Staatsangehörigkeit mit "unbekannt", "ungeklärt" oder "staatenlos" angegeben wurde.

2023 - 005

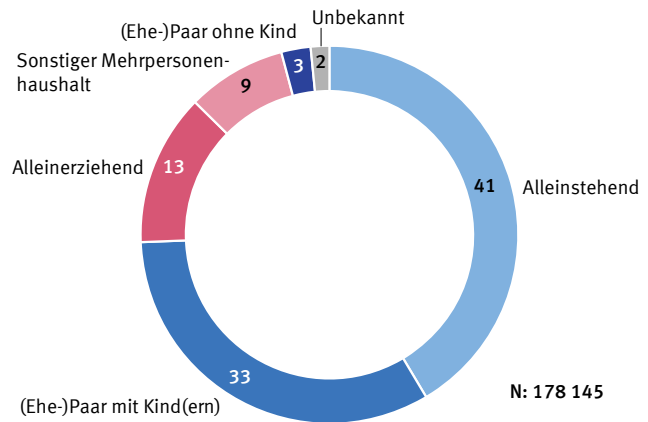
Bisherige Studien zur Wohnungslosigkeit in Deutschland belegten häufig, dass die Mehrheit der Betroffenen männlich war. Dies bestätigen auch die Ergebnisse der Statistik untergebrachter wohnungsloser Personen: Mehr als die Hälfte der Untergebrachten ist männlich (siehe Tabelle 1). Jedoch zeigt die neue Statistik überdies, dass ein großer Teil der untergebrachten Personen auch wohnungslose Frauen und Kinder sind. Ungefähr ein Viertel (26,5%) der Personen war jünger als 18 Jahre. Dieser relativ große Anteil der unter 18-Jährigen ist auf die Unterbringung von Kindern in Familienkonstellationen zurückzuführen. Das Durchschnittsalter der am 31. Januar 2022 untergebrachten wohnungslosen Personen betrug 32 Jahre.

Die Statistik erhebt zum Stichtag auch Angaben zu den Haushaltstypen, in denen die untergebrachten wohnungslosen Personen leben. Rund 46% der erfassten Personen waren als Alleinerziehende oder in Paarhaushalten mit Kind(ern) untergebracht und 41% als Alleinstehende. In sonstigen Mehrpersonenhaushalten lebten 8,5% der untergebrachten Personen, beispielsweise mit Geschwistern oder in anderen Familienkonstellationen. Wohn- oder Zweckgemeinschaften, bei denen Personen zwar gemeinsam untergebracht waren, die jedoch nicht gemeinsam wirtschafteten, sind jeweils als Alleinstehende erfasst worden. Bei 1,6% der gemeldeten Personen war der Haushaltstyp unbekannt. [↗ Grafik 2](#)

[↗ Tabelle 3](#) differenziert die Haushaltstypen zusätzlich nach Haushaltsgröße. In Bezug auf die Haushaltsgröße entspricht der Anteil von 41,4% Einpersonenhaushalten dem Anteil der als alleinstehend gemeldeten Personen. Mehr als ein Viertel der gemeldeten Personen

**Grafik 2**

**Untergebrachte wohnungslose Personen nach Haushaltstyp in %**



Stichtag: 31. Januar 2022.

2023 - 006

(29,7%) gehörte hingegen einem Haushalt mit 5 oder mehr Personen an. Hierzu zählen vor allem Paare mit Kind(ern).

Die Art der Überlassung von Räumen zu Wohnzwecken unterscheidet zwischen kurzfristigem Hilfeangebot (dazu zählen Notunterkünfte oder Übernachtungsstellen), (teil-)stationärem Wohnraum und sonstigen Angeboten, die keiner anderen Kategorie zugeordnet werden können. Am häufigsten wurden sonstige Angebote gemeldet (57%), gefolgt von kurzfristigen Angeboten (27%), stationären Angeboten (13%) und teilstationären Angeboten (2,8%). Der große Anteil der sonstigen Angebote könnte darauf zurückzuführen sein, dass häufig Personen in durch die Kommunen zur Verfügung

**Tabelle 3**

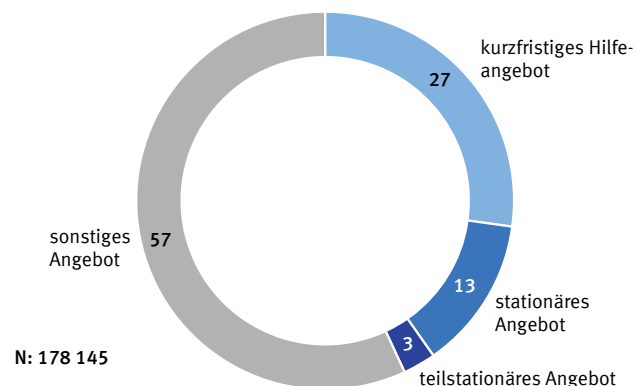
**Untergebrachte wohnungslose Personen am 31. Januar 2022 nach Haushaltstyp und Haushaltsgröße**

|                          | Haushalte mit |            |            |            |                     | Haushaltsgröße unbekannt | Insgesamt <sup>1</sup> |
|--------------------------|---------------|------------|------------|------------|---------------------|--------------------------|------------------------|
|                          | 1 Person      | 2 Personen | 3 Personen | 4 Personen | 5 und mehr Personen |                          |                        |
|                          | %             |            |            |            |                     |                          |                        |
| Alleinstehende           | 41,4          | 0          | 0          | 0          | 0                   | 0                        | 41,4                   |
| (Ehe-)Paar ohne Kind     | 0             | 2,5        | 0          | 0          | 0                   | 0                        | 2,5                    |
| Alleinerziehend          | 0             | 3,7        | 3,4        | 2,6        | 3,3                 | 0                        | 13,0                   |
| (Ehe-)Paar mit Kind(ern) | 0             | 0          | 3,2        | 7,2        | 22,2                | 0,2                      | 32,9                   |
| Sonstige Haushaltstypen  | 0             | 1,5        | 1,2        | 1,3        | 4,2                 | 0,3                      | 8,5                    |
| Haushaltstyp unbekannt   | 0             | 0          | 0          | 0          | 0                   | 1,6                      | 1,6                    |
| Insgesamt <sup>1</sup>   | 41,4          | 7,7        | 7,8        | 11,1       | 29,7                | 2,3                      | 100                    |

<sup>1</sup> Abweichungen bei der Summierung sind auf geheimhaltungsbedingte Rundungen zurückzuführen.

**Grafik 3**

Untergebrachte wohnungslose Personen nach Art der Überlassung von Räumen zu Wohnzwecken in %



Stichtag: 31. Januar 2022.

2023 - 007

gestelltem Wohnraum untergebracht wurden, oder es sich um Personen mit Fluchthintergrund handelt. Für

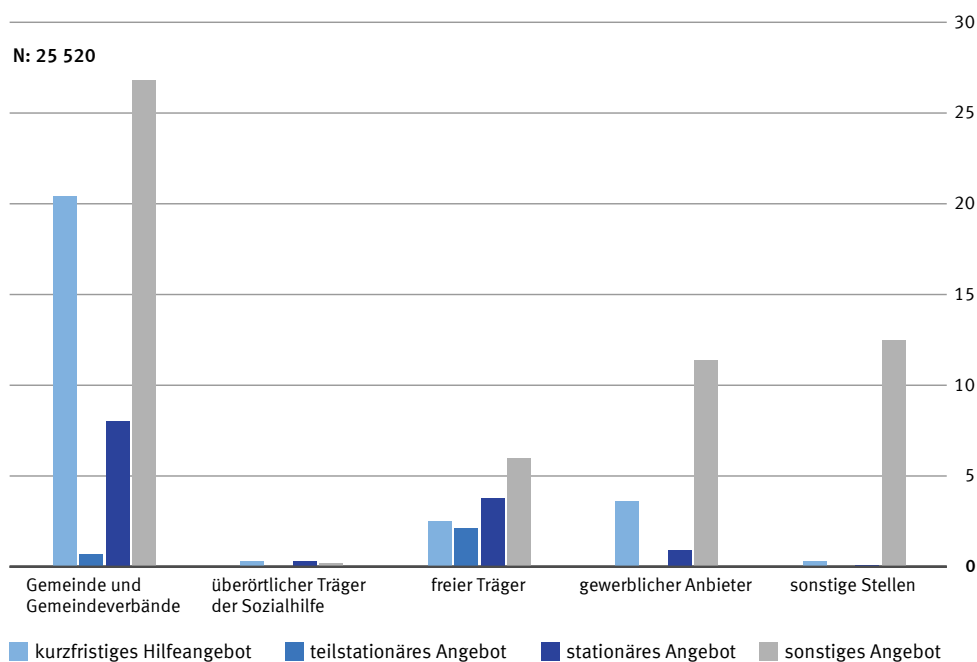
diese Art von Überlassung sieht das Gesetz derzeit keine gesonderten Ausprägungen vor. [➤ Grafik 3](#)

Gemeinden und Gemeindeverbände haben am Stichtag 31. Januar 2022 insgesamt fast 56 % der Unterbringungsmöglichkeiten zur Verfügung gestellt. Davon waren 20,4 % kurzfristige Hilfeangebote, 26,8 % sonstige Angebote sowie 8,0 % stationäre und 0,7 % teilstationäre Angebote. [➤ Grafik 4](#)

Gewerbliche Anbieter wie Hotels und Pensionen brachten rund 16 % der wohnungslosen Personen unter. Von den etwas über 14 % der Unterbringungsmöglichkeiten von Einrichtungen der freien Wohlfahrtspflege entfielen 2,5 % auf kurzfristige Angebote, 2,1 % auf teilstationäre, 3,8 % auf stationäre und 6,0 % auf sonstige Angebote. Knapp 13 % der wohnungslosen Personen waren durch Angebote von sonstigen Stellen untergebracht. Weniger als ein Prozent der Unterbringungsmöglichkeiten wurde durch überörtliche Träger der Sozialhilfe zur Verfügung gestellt.

**Grafik 4**

Untergebrachte wohnungslose Personen nach Anbietern der Unterbringungsmöglichkeiten und Angeboten in %

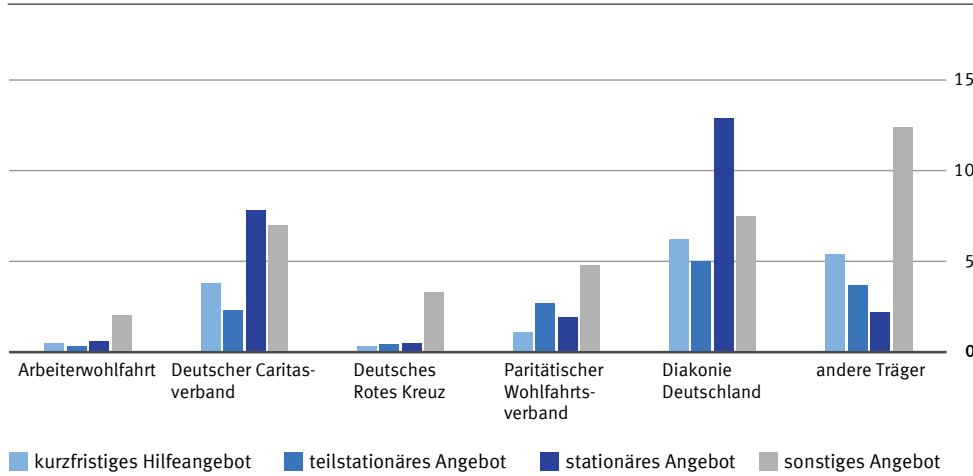


Stichtag: 31. Januar 2022.

2023 - 008

**Grafik 5**

**Untergebrachte wohnungslose Personen nach einzelnen Verbänden der freien Wohlfahrtspflege und Angeboten am 31. Januar 2022**  
in %



Stichtag: 31. Januar 2022.

2023 - 009

➤ Grafik 5 zeigt die Aufteilung der Angebote von einzelnen Verbänden der freien Wohlfahrtspflege. Am 31. Januar 2022 waren insgesamt 25 520 wohnungslose Personen bei freien Trägern untergebracht: in Unterbringungsangeboten der Diakonie Deutschland (31,6%), in Angeboten des deutschen Caritasverbands (20,9%), in Angeboten des Paritätischen Wohlfahrtsverbands (10,4%), in Angeboten des Deutschen Roten Kreuzes (4,5 %) und in Angeboten der Arbeiterwohlfahrt (AWO; 3,4 %). Andere Träger waren für knapp 30 % der Unterbringungen zuständig. Im Vergleich zu den restlichen Wohlfahrtsverbänden spielte bei der Diakonie Deutschland und dem deutschen Caritasverband das stationäre Angebot eine wesentlich größere Rolle.

## 5

### Zentrale Erkenntnisse der ersten Durchführung der Statistik

Die erstmals zum Stichtag 31. Januar 2022 durchgeführte Erhebung zu untergebrachten wohnungslosen Personen in Deutschland lieferte wertvolle Informationen zu diesem Personenkreis. Allerdings ist die Einführung einer neuen Statistik sowohl für die Auskunft gebenden Stellen als auch für das Statistische Bundesamt stets mit Herausforderungen verbunden.

Eine der Herausforderungen ist in diesem Fall, dass anerkannte Geflüchtete, die als Wohnungslose untergebracht werden, in der Statistik zu erfassen sind, die gesetzliche Grundlage zur Auskunftspflicht jedoch Ausländerbehörden bisher nicht explizit berücksichtigt. Diese Lücke in der gesetzlichen Auskunftspflicht führte dazu, dass einige Berichtsstellen aufgrund von Datenschutzrichtlinien keine Informationen über den Personenkreis der Geflüchteten von den zuständigen Behörden erhalten konnten. In anderen Fällen gaben die Kommunen an, dass auch die hohe Auslastung bei den Ausländerbehörden eine Zuarbeit verhindert habe. Dadurch wurden in einigen Fällen keine Geflüchteten im örtlichen Zuständigkeitsbereich gemeldet. In anderen Fällen führte es dazu, dass alle Geflüchteten gemeldet wurden, weil keine differenzierten Angaben über den Status einer Person vorlagen, die nötig waren, um festzustellen, ob diese wohnungslos untergebracht war.

Einige der Verbände, welche die erste Durchführung der Statistik im Beirat begleitet haben, meldeten zurück, dass eventuelle Überlastungen durch die Corona- und die Ukraine-Krise die Kapazitäten in den Kommunen erschöpft hätten. Dies führte dazu, dass nicht alle unterbringenden Einrichtungen zum Stichtag Daten abgefragt haben. Daher unterstützt das Statistische Bundesamt vor der zweiten Durchführung der Statistik die Kommunen, indem es mithilfe der Beiratsmitglieder zusätzlich sämtliche Einrichtungen in Deutschland recherchiert, die

wohnungslose Personen unterbringen. Dieses Vorgehen soll sicherstellen, dass alle relevanten Einrichtungen in der Statistik berücksichtigt werden, einer eventuellen Untererfassung entgegenwirken und den Kommunen die Arbeit an der Statistik erleichtern. Die Untererfassung wird zwar allgemein als gering eingeschätzt, sie lässt sich jedoch nicht vollständig ausschließen.

Für die neue Statistik mussten sich die Kommunen mit neuen Inhalten auseinandersetzen und die benötigten Informationen zusammentragen. So waren zum Beispiel alle zu einem Haushalt gehörenden Personen zu melden. Jedoch lagen in manchen Kommunen lediglich Informationen zum Haushaltsvorstand vor, was dazu führte, dass nicht alle untergebrachten wohnungslosen Personen in diesen Gemeinden gemeldet werden konnten. Mit vermehrten Informationsangeboten und erweiterten Kontakten zu diesen Berichtsstellen will das Statistische Bundesamt die Gemeinden in diesem Punkt noch besser auf die nächste Erhebung vorbereiten.

Zudem gelten Kältehilfen, Notunterkünfte und Übernachtungsstellen bei dieser Statistik als „kurzfristige Hilfeangebote“. Wohnraum, der durch die Kommune mittel- oder längerfristig zur Verfügung gestellt wird, gilt als „sonstiges Angebot“. Die Unterbringungsangebote der Kommunen sollen häufig zwar nur kurze Zeiträume überbrücken, bestehen in der Realität aber oftmals mehrere Jahre. Daher ist es möglich, dass die ordnungsrechtliche Unterbringung von Wohnungslosen in der Statistik als kurzfristiges Hilfeangebot angegeben wurde, obwohl eine Signierung als sonstiges Angebot korrekt gewesen wäre.

Diese Fälle traten nur vereinzelt auf, daher schränken die genannten Herausforderungen die Aussagekraft der Statistik nicht wesentlich ein. Jedoch werden erst die Ergebnisse weiterer Erhebungen das Ausmaß dieser Einflüsse genauer belegen können. Mit den nächsten Erhebungen werden auch die Kenntnisse zu den Inhalten der Statistik weiter steigen und die Zusammenarbeit zwischen dem Statistischen Bundesamt und den Kommunen werden die Meldungen verbessern.

## 6

### Fazit und Ausblick

---

Mit der Einführung der Statistik hat der Gesetzgeber zum ersten Mal eine bundesweite amtliche Datengrundlage über Wohnungslosigkeit in Deutschland geschaffen, die sich allerdings ausschließlich auf untergebrachte wohnungslose Personen bezieht. Straßenobdachlosigkeit und Personen, die bei Freunden oder Bekannten unterkommen, werden in der alle zwei Jahre durchgeführten Berichterstattung des Bundesministeriums für Wohnen, Stadtentwicklung und Bauwesen betrachtet.

Die erstmalige Durchführung der Statistik ist erfolgreich verlaufen. Um die Qualität der Statistik weiter zu steigern und in Zukunft die Lage von wohnungslosen Personen in Deutschland noch besser abzubilden, bieten sich trotzdem einige Verbesserungsmöglichkeiten an. Beispielsweise ist eine Ausweitung der Auskunftspflicht auf die Ausländerbehörden zu prüfen. Zudem sollten beim Erhebungsmerkmal zur Art der Überlassung von Wohnräumen weitere Merkmalsausprägungen ergänzt werden, um den hohen Anteil der Signierung „sonstiges Angebot“ besser zu erklären oder zu verringern.

Zusätzlich sollte der regelmäßige Kontakt zwischen dem Statistischen Bundesamt und den Berichtsstellen dafür sorgen, die beteiligten Stellen noch umfassender mit Informationen zum Thema Wohnungslosigkeit im Sinne der Statistik zu versehen. Damit sollte sich auch schrittweise das allgemeine Verständnis für die Erhebung verbessern.

Die regelmäßige jährliche Erhebung ermöglicht es, die Dunkelziffer von wohnungslosen Personen in Deutschland nicht nur einmalig zu beleuchten, sondern künftig die Entwicklung von Wohnungslosigkeit in Deutschland genauer zu beschreiben. Die neue Statistik leistet einen wichtigen Beitrag, um den verantwortlichen Stellen auf kommunaler sowie Landes- und Bundesebene Grundlagen für politische Entscheidungen bereitzustellen und detaillierte Informationen zur Armuts- und Reichtumsberichterstattung des Bundes beizusteuern.

Weiterführende Informationen zur Statistik untergebrachter wohnungsloser Personen stehen unter [www.destatis.de/Wohnungslosigkeit](http://www.destatis.de/Wohnungslosigkeit) und in der [Datenbank GENESIS-Online](#) (Tabelle [22971](#)) zur Verfügung. 

## LITERATURVERZEICHNIS

---

BAG Wohnungslosenhilfe e.V. *Wohnungslosigkeit: Kein Ende in Sicht*. Pressemitteilung vom 11. November 2019. [Zugriff am 17. Januar 2023]. Verfügbar unter: [www.bagw.de](http://www.bagw.de)

BMAS (Bundesministerium für Arbeit und Soziales). *Empirische Untersuchung zum Gegenstand nach § 8 Absatz 2 und 3 Wohnungslosenberichterstattungsgesetz*. Forschungsbericht 605. 2022a. [Zugriff am 19. Januar 2023]. Verfügbar unter: [www.bmas.de](http://www.bmas.de)

BMAS (Bundesministerium für Arbeit und Soziales). *Ausmaß und Struktur von Wohnungslosigkeit. Der Wohnungslosenbericht 2022 des Bundesministeriums für Arbeit und Soziales*. 2022b. [Zugriff am 17. Januar 2023]. Verfügbar unter: [www.bmas.de](http://www.bmas.de)

BMAS (Bundesministerium für Arbeit und Soziales). *Machbarkeitsstudie für eine regelmäßige Berichterstattung gemäß § 8 Absatz 4 Wohnungslosenberichterstattungsgesetz*. Forschungsbericht 606. 2022c. [Zugriff am 19. Januar 2023]. Verfügbar unter: [www.bmas.de](http://www.bmas.de)

BMAS (Bundesministerium für Arbeit und Soziales). *Lebenslagen in Deutschland. Der Sechste Armuts- und Reichtumsbericht der Bundesregierung*. 2021. [Zugriff am 17. Januar 2023]. Verfügbar unter: [www.armuts-und-reichtumsbericht.de](http://www.armuts-und-reichtumsbericht.de)

Deutscher Bundestag. *Wohnungslosigkeit in Deutschland, Maßnahmen des Bundes*. In: Wissenschaftliche Dienste. Aktenzeichen: WD 6 - 3000 - 087/21. 2021.

FEANTSA. *ETHOS – European Typology on Homelessness and Housing Exclusion*. 2005. [Zugriff am 16. Januar 2023]. Verfügbar unter: [www.feantsa.org](http://www.feantsa.org)

Gesellschaft für innovative Sozialplanung und Sozialforschung e.V. (GISS)/Kantar Public. *Empirische Untersuchung zum Gegenstand nach § 8 Absatz 2 und 3 Wohnungslosenberichterstattungsgesetz*. Erstellt im Auftrag des Bundesministeriums für Arbeit und Soziales. 2022. [Zugriff am 17. Januar 2023]. Verfügbar unter: [www.bmas.de](http://www.bmas.de)

Henke, Jutta. *Wie lässt sich Wohnungslosigkeit verhindern? Soziale Arbeit kontrovers*, Ausgabe 23. Berlin 2020.

Henke, Jutta. *Mehr Wissen – Auftrag zum Handeln. Lebenslagen von Menschen ohne Unterkunft und in verdeckter Wohnungslosigkeit*. In: *Wohnungslos*. 64. Jahrgang. Ausgabe 3-4/2022, Seite 91 ff.

## RECHTSGRUNDLAGEN

---

Wohnungslosenberichterstattungsgesetz vom 4. März 2020 (BGBl. I Seite 437), das durch Artikel 2 des Gesetzes vom 5. Dezember 2022 (BGBl. I Seite 2160) geändert worden ist.

Personenstandsgesetz (PStG) vom 19. Februar 2007 (BGBl. I Seite 122), das zuletzt durch Artikel 1 des Gesetzes vom 19. Oktober 2022 (BGBl. I Seite 1744) geändert worden ist.

# AUFBEREITUNG VON ANGABEN ZU GEBURTSORTEN IM AUSLÄNDERZENTRALREGISTER MIT OPENSTREETMAP

Jan Eberle

📌 **Schlüsselwörter:** Ausländerzentralregister – Schutzsuchende – Geburtsort – Geokodierung – OpenStreetMap

## ZUSAMMENFASSUNG

Angaben zu Geburtsorten von Ausländerinnen und Ausländern ermöglichen eine Vielzahl neuer Auswertungsmöglichkeiten für amtliche Statistiken. Kommen somalische Schutzsuchende vermehrt aus Regionen, die durch Dürre und Nahrungsmittelknappheit besonders stark vom Klimawandel betroffen sind? In welchen US-amerikanischen und indischen Regionen gelingt das Anwerben von Fachkräften für den heimischen Arbeitsmarkt? Für die Beantwortung dieser und ähnlicher Fragen kann die Angabe zum Geburtsort als Indikator für die regionale Herkunft einer Person innerhalb eines Staates herangezogen werden. Der Artikel liefert eine Machbarkeitsstudie zur Aufbereitung und Geokodierung von Angaben zum Geburtsort im Ausländerzentralregister mithilfe der Geodaten von OpenStreetMap. Die Ergebnisse werden nachträglich als interaktives Kartenangebot zur Verfügung gestellt.

📌 **Keywords:** Central Register of Foreigners – people seeking protection – birthplace – geocoding – OpenStreetMap

## ABSTRACT

*Data on the birthplaces of foreigners create a multitude of new options for analysis in official statistics. Do Somali people seeking protection mostly come from regions that are particularly affected by climate change because of droughts and food shortages? From which U.S. and Indian regions can Germany successfully attract skilled workers for the domestic labour market? To answer these and similar questions, information on the birthplace is useful as an indicator of a person's regional origin within a country. This article presents a feasibility study on the processing and geocoding of birthplace data held in the Central Register of Foreigners, using OpenStreetMap geodata. The results will later also be available as an interactive map.*



Jan Eberle

ist Volkswirt und angehender Data Scientist. Im Statistischen Bundesamt arbeitet er als Referent im Bereich der laufenden Bevölkerungsstatistiken. Zu seinen Arbeitsschwerpunkten gehören dabei auch Auswertungen des Ausländerzentralregisters zur ausländischen Bevölkerung sowie zu Schutzsuchenden.

## 1

### Einleitung

Die Nutzung von Verwaltungsdaten und die damit einhergehende Entlastung der Auskunftspflichtigen ist ein Kernziel der Digitalen Agenda des Statistischen Bundesamtes (Statistisches Bundesamt, 2019) und auch in seiner Hausstrategie fest verankert. Die Nutzung des Ausländerzentralregisters (AZR) hat dabei eine lange Tradition in der Bevölkerungsstatistik. Bereits seit den 1970er-Jahren werden dessen Daten aufbereitet, um Informationen zur ausländischen Bevölkerung in Deutschland zu gewinnen (Fleischer, 1989).

Das AZR ist eines der größten Verwaltungsregister in Deutschland. Es enthält Informationen zu allen ausländischen Personen, die sich nicht nur vorübergehend in Deutschland aufhalten. In der Regel sind das Personen, deren Aufenthalt länger als drei Monate andauert.<sup>1</sup> Das AZR führt die Datenbestände aller Behörden zusammen, die mit asyl- und aufenthaltsrechtlichen Aufgaben betraut sind, und dient diesen als zentrale Informationsplattform.

Wie bei Verwaltungsdaten üblich, ist die statistische Verwertbarkeit der Daten damit nur einer von vielen Verwendungszwecken. Entsprechend unterscheidet sich die Datenqualität auch von jener von Daten aus Primärerhebungen. Vor einer Verwertung der Daten für statistische Zwecke steht daher eine sorgfältige Qualitätsprüfung und Aufbereitung.

Infolge der Fluchtmigration der Jahre 2015 und 2016 hat der Gesetzgeber das AZR kontinuierlich weiterentwickelt und dabei auch die Nutzungsmöglichkeiten für die amtliche Statistik weiter verbessert.<sup>2</sup> Seit Anfang 2020 erhält das Statistische Bundesamt zusätzlich Angaben zu Geburtsstaaten und Geburtsorten aus dem Ausländerzentralregister für die Verwendung in der amtlichen Statistik (§ 23 AZR-Gesetz). Die beiden Merkmale werden im Verwaltungskontext von Ausländerbehörden, Erstaufnahmeeinrichtungen oder Polizeibehörden ohne Plausibilitätskontrollen erfasst. Die Erfassung des

Geburtsstaats erfolgt strukturiert als ISO-Code<sup>3</sup>, ist als freiwillige Angabe jedoch in erheblichem Umfang unvollständig. Angaben zum Geburtsort werden in dem Register als Pflichtangabe weitgehend vollständig erfasst, allerdings erfolgt die Erfassung unstrukturiert als Freitextangabe.

Vor diesem Hintergrund hat das Statistische Bundesamt zunächst ein Verfahren zur Aufbereitung der Angaben zum Geburtsstaat entwickelt (Canan/Eberle, 2022). Der vorliegende Aufsatz stellt darauf aufbauend ein Konzept zur Aufbereitung und Geokodierung der Angaben zum Geburtsort vor.

Lösungsansätze einer vergleichbaren Problemstellung liefern Šimbera und andere (2021) für administrative Daten aus Tschechien sowie Conti und Cimbelli (2018) für Italien. Šimbera und andere (2021) entwickeln einen eigenen Matching-Algorithmus mit aufwendigem Preprocessing zur Standardisierung der Ortsangaben. Conti und Cimbelli (2018) hingegen nutzen mit OpenRefine ein bestehendes Open-Source-Werkzeug für den gesamten Aufbereitungsprozess. Hierbei verbleiben gewisse Details, wie der Umgang mit mehrdeutigen Ortsnamen, in der Black Box. Das in diesem Artikel vorgeschlagene Aufbereitungskonzept versucht, einen Mittelweg zwischen Aufwand und Transparenz zu gehen.

Als konkretes Anwendungsbeispiel sind die Geburtsorte von Ende 2020 im AZR registrierten Schutzsuchenden im Zuge einer Machbarkeitsstudie aufbereitet und geokodiert worden.<sup>4</sup> Eine zentrale Rolle bei der Aufbereitung spielen die Geodaten von OpenStreetMap (OSM), anhand derer die unstrukturierten Angaben in den administrativen Daten abgeglichen, standardisiert und geokodiert werden. Damit liefert dieser Beitrag auch ein Beispiel dafür, wie Open Web Data effektiv in der amtlichen Statistik genutzt werden kann.

1 Personen mit deutscher und zusätzlich einer ausländischen Staatsangehörigkeit sind nicht im AZR registriert.

2 Durch das Datenaustauschverbesserungsgesetz 2016, das Zweite Datenaustauschverbesserungsgesetz 2019 sowie das Gesetz zur Weiterentwicklung des Ausländerzentralregisters 2021.

3 Ländercodes der International Organization for Standardization: [www.iso.org](http://www.iso.org)

4 Schutzsuchende sind Ausländerinnen und Ausländer, die sich nach den Angaben im AZR aufenthaltsrechtlich unter Berufung auf humanitäre Gründe in Deutschland aufhalten (Eberle, 2019).

## 2

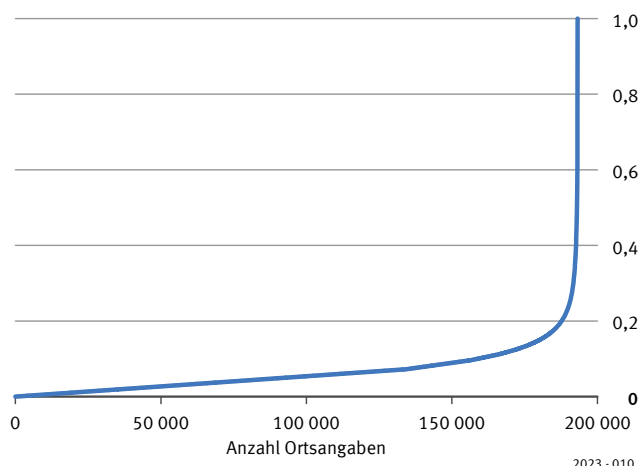
### Daten zum Geburtsort von Schutzsuchenden aus dem AZR

Der Geburtsort ist standardmäßig in vielen internationalen Ausweisdokumenten eingetragen und Pflichtangabe bei der Registrierung im AZR. Die Erfassung erfolgt durch Ausländerbehörden, Erstaufnahmeeinrichtungen sowie Polizeibehörden im Zuge der Erstregistrierung von Ausländerinnen und Ausländern.

Ende 2020 waren 1 857 000 Schutzsuchende im AZR registriert. Insgesamt finden sich in den Daten 193 000 unterschiedliche Angaben zu Geburtsorten in 176 Staaten. Die Häufigkeit der Orte in den Daten folgt einer stark linksschiefen Verteilung. Knapp 70 % der Ortsangaben kommen genau einmal vor. Dies steht auch in Zusammenhang mit der unstrukturierten Erfassung der Geburtsorte. Die Herkunftsländer von Schutzsuchenden liegen zum großen Teil im arabischen Sprachraum und oft können Schutzsuchende bei ihrer Registrierung keine offiziellen Dokumente vorlegen. Daraus ergibt sich eine Vielzahl von unterschiedlichen Schreibweisen und Schreibfehlern in den Daten, die im Aufbereitungsprozess aufzulösen sind. Die 70 % einzigartigen Ortsangaben repräsentieren allerdings nur rund 7 % aller Schutzsuchenden. Gleichzeitig stehen die zehn häufigsten

#### Grafik 1

Kumulative Häufigkeitsverteilung der Ortsangaben im Ausländerzentralregister vor der Aufbereitung  
Anteil an den Schutzsuchenden in %



Ortsangaben für 22 % aller Schutzsuchenden. Fehler bei der Geokodierung dieser Ortsangaben fallen demnach besonders ins Gewicht. Positiv formuliert ist es möglich, einem Großteil der Schutzsuchenden den richtigen Geburtsort zuzuweisen, wenn nur ein geringer Anteil der Ortsangaben korrekt aufbereitet wird. [↗ Grafik 1](#)

In einigen Fällen werden höhere administrative Verwaltungseinheiten, wie Provinzen oder Distrikte, als Geburtsorte erfasst. Diese ermöglichen immerhin eine Eingrenzung des Geburtsortes und werden daher weiterverarbeitet.<sup>15</sup> Darüber hinaus finden sich Angaben ohne verwertbaren Inhalt. Hierzu zählen Angaben von Staatennamen, ISO-Codes und eine Vielzahl unterschiedlicher Einträge, die darauf verweisen, dass der Geburtsort nicht bekannt ist. Diese nicht verwertbaren Angaben werden nicht weiter aufbereitet.

## 3

### Geodaten von OpenStreetMap

Das OSM-Projekt sammelt seit Mitte der 2000er-Jahre weltweit raumbezogene Informationen (Geodaten) und stellt diese als Open Data unter der [Open Database License](#) zur Verfügung. Ebenso wie in der freien Enzyklopädie Wikipedia werden die Informationen nach dem Prinzip des Crowdsourcing gesammelt, das heißt die Geodaten werden von Freiwilligen und ohne kommerzielle Interessen gesammelt und laufend aktualisiert. Im Jahr 2021 zählte diese OSM-Community rund 1,6 Millionen aktive Editorinnen und Editoren und 7,8 Millionen registrierte Nutzerinnen und Nutzer. Editoren sammeln und validieren Geodaten aus öffentlich zugänglichen Datenquellen und durch manuelle Eingaben. Mit insgesamt 7 Milliarden Einträgen und rund 4,5 Millionen Aktualisierungen am Tag ist OSM eine dynamische Geodatenbank. Mitte 2022 zählte OSM weltweit rund 4,7 Millionen Einträge, die als bewohnte Siedlungen (englisch: populated settlements) annotiert waren.<sup>16</sup>

[↗ Tabelle 1](#)

5 In Staaten, in denen dies vermehrt vorkommt, sollte eine Auswertung nicht unterhalb der Ebene dieser Verwaltungseinheiten erfolgen (siehe Kapitel 5).

6 <https://taginfo.openstreetmap.org>

**Tabelle 1****Siedlungsgebiete (populated settlements) in OpenStreetMap**

| Tag (Schlüssel)             | Anzahl    | Beschreibung                                                                                                                                                                                                                 |
|-----------------------------|-----------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| city (Hauptstadt/Großstadt) | 10 978    | Die größte städtische/urbane Siedlung innerhalb eines Verwaltungsgebiets.                                                                                                                                                    |
| town (Stadt/Kleinstadt)     | 106 369   | Eine städtische Siedlung von lokaler bis regionaler Bedeutung mit üblicherweise mehr als 10 000 Einwohnerinnen und Einwohnern.                                                                                               |
| village (Dorf)              | 1 417 498 | Eine kleine ländliche Siedlung mit wenigen zentralörtlichen Einrichtungen und mit weniger als 10 000 Einwohnerinnen und Einwohnern.                                                                                          |
| hamlet (Weiler)             | 1 700 548 | Eine kleinere, ländliche Ansammlung von Haushalten, typischerweise mit weniger als hundert Einwohnerinnen und Einwohnern und minimaler Infrastruktur, meist gänzlich gewerbelos und ohne Anbindung öffentlichen Nahverkehrs. |
| Sonstige                    | 1 486 399 |                                                                                                                                                                                                                              |
| Insgesamt                   | 4 721 792 |                                                                                                                                                                                                                              |

Quelle: [wiki.openstreetmap.org/wiki/Key:place](https://wiki.openstreetmap.org/wiki/Key:place) [Zugriff am 11. Januar 2023].

Die Datenqualität ist regional unterschiedlich und hängt stark davon ab, wie aktiv die jeweiligen Editoren sind. Für den Zeitraum 2013 bis 2019 identifizieren Seto und andere (2020) die aktivsten Communitys in Deutschland, in der Russischen Föderation und in den Vereinigten Staaten. Auch Herfort und andere (2021) attestieren, dass die meisten Einträge in entwickelten Ländern mit einem sehr hohen Human Development Index beigetragen werden. Nach dem verheerenden Erdbeben in Haiti im Jahr 2010 sind diese Unterschiede im Hinblick auf den Zugang zu Geodaten vermehrt in den Fokus gerückt. Seitdem schließt die OSM-Community, darunter zahlreiche Hilfsorganisationen wie Rotes Kreuz und Roter Halbmond, durch humanitäre Aktionen gezielt Lücken in den Geodaten in weniger entwickelten Regionen der Welt. Beispielsweise wurden mit dem [Missing Maps Project](#) seit 2014 rund 70 Millionen Geodaten vornehmlich in Subsahara-Afrika, Südostasien sowie Mittel- und Südamerika gesammelt. Herfort und andere (2021) berechneten, dass zwischen Januar 2008 und Mai 2020 rund 14 % aller in OSM hinzugefügten Gebäude auf humanitäre Mapping-Aktivitäten zurückzuführen sind. Die weltweite geografische Abdeckung am Beispiel der Annotationen Stadt und Dorf (englisch: city beziehungsweise village) zeigt [Grafik 2](#).

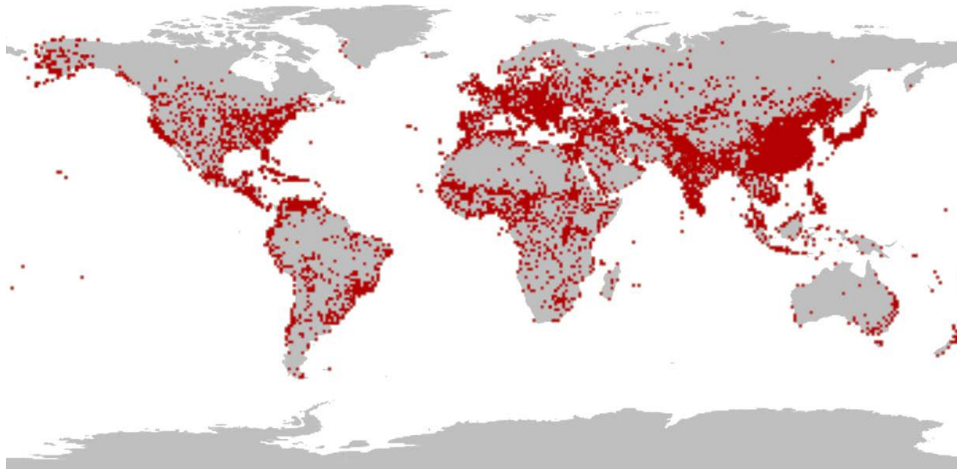
Lizenzrechtlich ist jegliche Art der Nutzung von OSM-Daten zulässig, sofern die Nutzungsbedingungen der Open Database Licence eingehalten werden. Die Nutzungsbedingungen verlangen, dass durch die Nutzung entstandene abgeleitete Datensätze ebenso frei zugänglich als Open Data zur Verfügung gestellt

werden. Ein abgeleiteter Datensatz entsteht, wenn OSM-Daten mit externen Daten weiterverarbeitet werden oder die gesamten OSM-Daten beziehungsweise ein substanzieller Teil der OSM-Datenbasis mit einer anderen Datenbasis zusammengeführt werden. Den OSM Licence Guidelines (2017) folgend entsteht durch Geokodierung, also durch die Ergänzung bestehender Datensätze mit Namen, Adressen und Geokoordinaten, kein abgeleiteter Datensatz, solange nicht systematisch OSM-Datenbestände in ihrer Gesamtheit extrahiert werden.

## Grafik 2

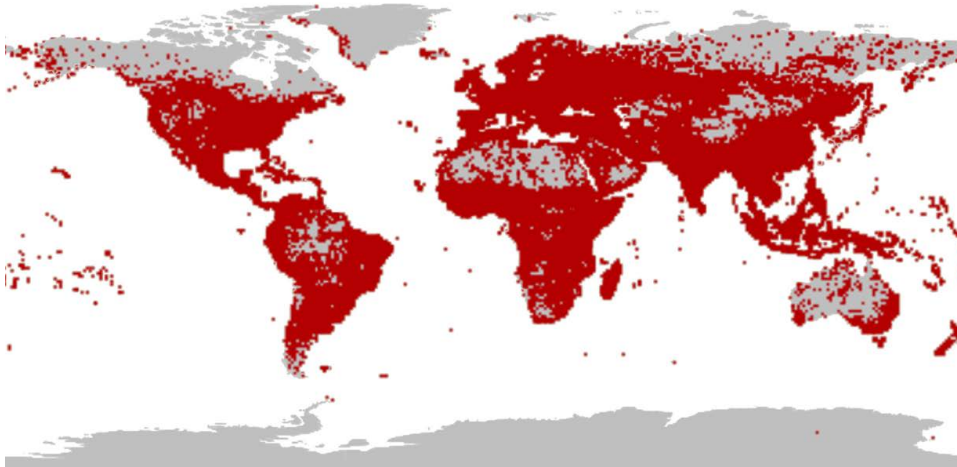
Weltweite Abdeckung der OSM-Geodaten

OSM Tag Place=city (Stadt)



Quelle: <https://taginfo.openstreetmap.org/tags/place=city#map>

OSM Tag Place=village (Dorf)



Quelle: <https://taginfo.openstreetmap.org/tags/place=village#map>

2023 - 0011

## 4

### Aufbereitung des Geburtsortes mit OSM-Geodaten

Die Aufbereitung der Angaben zu den Geburtsorten von Schutzsuchenden aus dem AZR anhand der Geodaten von OpenStreetMap erfolgt konzeptionell in drei Schritten: Potenzielle geografische Orte werden identi-

fiziert, Mehrdeutigkeit aufgelöst und die Geodaten des zugewiesenen Ortes übernommen.

#### 4.1 Potenzielle geografische Orte identifizieren

In einem ersten Schritt wird festgestellt, welche geografischen Orte mit einer Ortsangabe im AZR gemeint sein könnten. Dieser auch als Gazetteer Matching bezeichnete Vorgang (Goldberg und andere, 2007) ist

vergleichbar mit dem Suchen von Ortsnamen in einem Ortsverzeichnis (englisch: gazetteer). Bei der Auswahl potenzieller geografischer Orte treten unterschiedliche Probleme auf. Einerseits können mit unterschiedlichen textlichen Repräsentationen die gleichen geografischen Orte gemeint sein, zum Beispiel ‚Frankfurt a.M.‘ und ‚Frankfurt am Main‘ (Synonyme). Aufgrund der Erfassung als Freitext enthalten die Daten unterschiedliche Schreibweisen, unterschiedliche Sprachen und auch Schreibfehler. Synonyme Ortsangaben können mit Werkzeugen aus dem Bereich des Natural Language Processing identifiziert werden. Hierbei werden über textliche oder phonetische Distanzmaße unterschiedliche Schreibweisen und Schreibfehler erkannt.

Andererseits können mit einem Ortsnamen unterschiedliche geografische Orte gemeint sein, zum Beispiel könnte ‚Frankfurt‘ sowohl ‚Frankfurt am Main‘ als auch ‚Frankfurt an der Oder‘ betreffen (Homonyme). Um homonyme beziehungsweise mehrdeutige Ortsangaben aufzulösen, ist zunächst eine möglichst komplette Auswahl aller möglichen Bedeutungen zu erstellen. Grundlage ist ein weitgehend vollständiges Verzeichnis aller weltweit bewohnten Siedlungen. Als ein solches weltweites Ortsverzeichnis dient OSM.

Technisch erfolgt der Zugriff auf die OSM-Daten über eine Online-Programmierschnittstelle (englisch: Application Programming Interface, kurz API). APIs ermöglichen eine vollständige Automatisierung und damit die Verarbeitung großer Datenmengen. Die Ortsangaben aus dem AZR werden dabei in einer lokalen PostgreSQL-Datenbank gehalten. Über ein Python-Skript werden diese blockweise aus der Datenbank extrahiert und in einzelnen HTTP-Anfragen an die API geschickt. Die Antwort der API erfolgt in Form eines JSON-Dokuments, aus dem die potenziellen geografischen Orte extrahiert werden.<sup>17</sup>

APIs bieten einen komfortablen Datenzugang, allerdings bestehen unterschiedliche Zugangsbeschränkungen. Kostenlosen OSM-Datenzugang für die Geokodierung bietet beispielsweise [Nominatim](#). Beschränkungen bestehen hier nur im Hinblick auf die Anzahl der Anfragen, die innerhalb einer gewissen Zeit gestellt werden können. Andere Anbieter bieten neben der Geokodierung

weitere, zumeist kostenpflichtige Dienstleistungen. Für die Aufbereitung der Angaben zum Geburtsort von Schutzsuchenden nutzt das Statistische Bundesamt die API von [LocationIQ](#). Dieser Anbieter stellt eine Autocomplete-Funktion zur Verfügung, die automatisch kleinere Schreibfehler und unterschiedliche Schreibweisen eines Ortes erkennt. So erkennt die API in den Originaldaten beispielsweise rund 50 unterschiedliche Repräsentationen der äthiopischen Hauptstadt Addis Abeba und ordnet diese korrekt dem gleichen geografischen Ort zu. Damit ersetzt die Autocomplete-Funktion aufwendiges nutzerseitiges Preprocessing.<sup>18</sup>  **Grafik 3**

Das Preprocessing beschränkt sich daher im Wesentlichen darauf, Angaben ohne verwertbaren Inhalt (ISO-Codes, Staatennamen und unterschiedliche Repräsentationen für unbekannte oder fehlende Angaben) zu identifizieren, um diese von der weiteren Verarbeitung auszuschließen.

Findet die API für eine Geburtsortsangabe keinen passenden Eintrag in den OSM-Daten, erfolgen Korrekturversuche. Zunächst wird überprüft, ob der nicht auffindbare Ort umbenannt wurde oder der Ortsname in einer anderen Sprache zu einem Treffer führt. Hierfür werden Informationen aus Wikipedia genutzt. Beispielsweise lautete der offizielle Ortsname der ukrainischen Stadt Dnipro im Ukrainischen bis Mai 2016 ‚Dnipropetrowsk‘. Ein großer Teil der Bevölkerung von Dnipro spricht russisch und würde als Geburtsort vermutlich ‚Dnepropetrowsk‘ angeben. In Wikipedia führen sowohl Suchanfragen nach ‚Dnipropetrowsk‘ als auch nach ‚Dnepropetrowsk‘ auf den Artikel des heutigen Dnipro. Ebenso verweisen Suchanfragen nach ‚Leningrad‘ auf das heutige Sankt Petersburg.

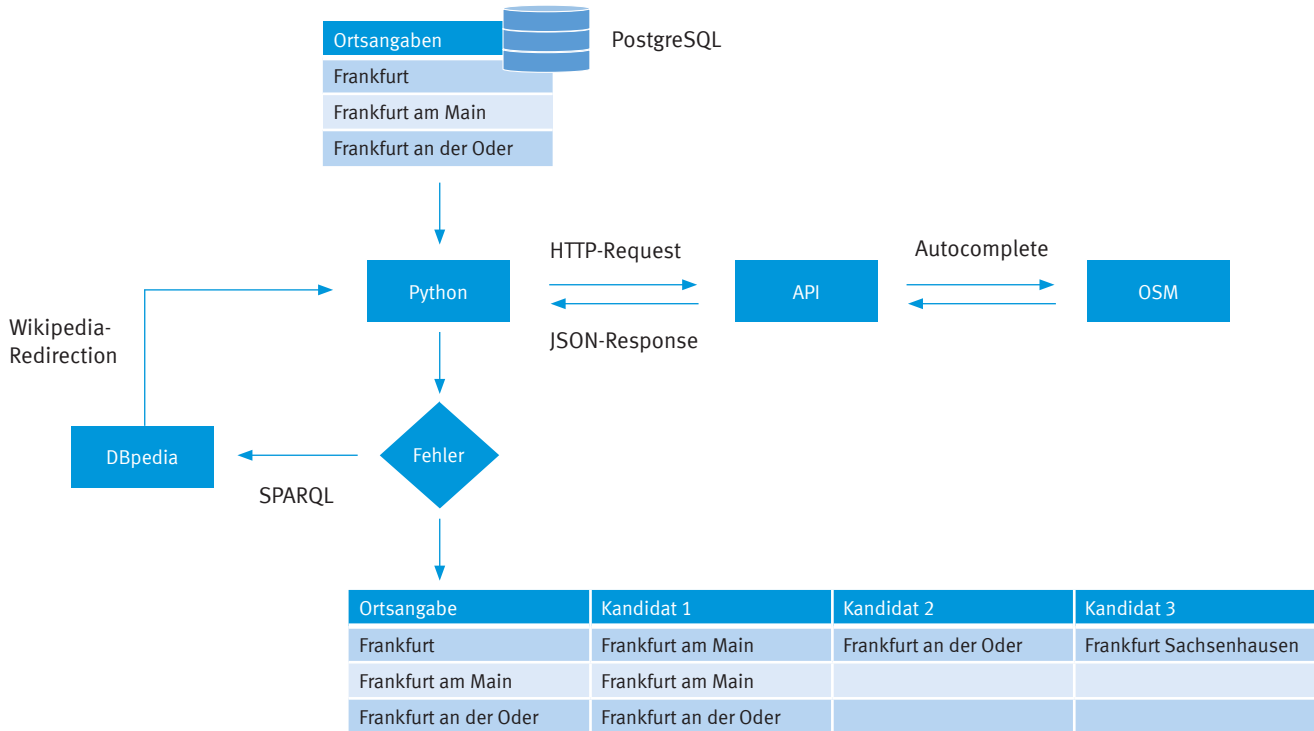
Die Prüfung, ob für eine Ortsbezeichnung in Wikipedia eine Weiterleitung besteht, erfolgt automatisiert über den [SPARQL-Endpunkt von DBpedia](#). Das DBpedia-Projekt extrahiert regelmäßig die Informationen aus Wikipedia und stellt diese in strukturierter Form (RDF-Graph) als Linked Open Data zur Verfügung (Lehmann und andere, 2015). Über die Abfragesprache SPARQL können die Informationen zu Weiterleitungen von Suchanfragen automatisiert abgefragt werden. Führt auch dieser Korrekturschritt nicht zu einem Treffer in

<sup>17</sup> Für jede Ortsangabe werden zehn Kandidaten extrahiert. Die Reihenfolge der Kandidaten legt ein in den OSM-Daten enthaltener Importance Score fest. Dieser Wert wird auf Basis des Wikipedia-Artikels eines Orts berechnet. Dabei wird Orten, auf deren Artikel innerhalb von Wikipedia häufiger verlinkt wird, ein höherer Wert zugeordnet.

<sup>18</sup> Die API ermöglicht täglich 25 000 Anfragen unter Verwendung der Autocomplete-Funktion.

**Grafik 3**

Ablauf der Identifizierung potenzieller geografischer Orte



2023 - 0012

den OSM-Daten, erfolgen bei häufig vorkommenden Angaben manuelle Korrekturen.

## 4.2 Mehrdeutigkeit von Ortsangaben auflösen

Im zweiten Schritt werden mehrdeutige Ortsangaben identifiziert und anschließend einem geografischen Ort zugewiesen. In der Literatur wird dieser Schritt auch als Toponym Resolution bezeichnet (DeLozier und andere, 2015). Ausgangspunkt ist die Liste potenzieller geografischer Orte, die mit einem Ortsnamen (Toponym) gemeint sein könnten, aus dem vorherigen Aufbereitungsschritt. Im Zuge der Toponym Resolution wird für diese Kandidaten eine Wahrscheinlichkeitsverteilung gesucht, anhand derer im Anschluss eine stochastische Zuordnung erfolgen kann.

### Mehrdeutige Ortsangaben identifizieren

Zunächst besteht die Herausforderung, mehrdeutige Ortsangaben automatisiert zu identifizieren. Hierfür erfolgt ein Vergleich der ursprünglichen Ortsangabe mit den Rückmeldungen (Kandidaten) aus den OSM-Daten. Eine Ortsangabe gilt dabei als mehrdeutig, wenn alle einzelnen Wörter, beginnend vom ersten Wort in der richtigen Reihenfolge, in mehreren Kandidaten vorkommt. Dahinter steht die Annahme, dass Namenszusätze, die die Mehrdeutigkeit auflösen, immer am Ende eines Ortsnamens angehängt werden. Weiterhin werden Kandidaten, die in OSM nicht als bewohnte Siedlungen (englisch: populated places) annotiert sind, ausgeschlossen. Dieses Vorgehen verhindert die Verwechslung von Städten mit gleichnamigen Distrikten oder Provinzen. Ebenso sollen Stadtteile, beispielsweise ‚Frankfurt Sachsenhausen‘, nicht mit mehrdeutigen Ortsangaben wie ‚Frankfurt am Main‘ und ‚Frankfurt an der Oder‘ verwechselt werden. Daher werden Kandidaten ausgeschlossen, die eine Distanz von weniger als zehn Kilometer zu einem bereits identifizierten Kandidaten aufweisen. Letztlich wird nicht von mehrdeutigen Ortsangaben ausgegan-

gen, wenn Ortsangabe und Kandidat auf dem ersten Rang, also der Kandidat mit dem höchsten OSM-Importance-Score, identisch sind.<sup>19</sup>

### Mehrdeutige Ortsangaben auflösen

Durch die bereits aufbereiteten Angaben zum Geburtsstaat (Canan/Eberle, 2022) ist die potenzielle internationale Mehrdeutigkeit der Ortsangaben (Paris in Texas oder Paris in Frankreich) bereits aufgelöst. Bestehen bleibt die Mehrdeutigkeit von Ortsnamen innerhalb eines Staates. Für die in Deutschland mehrdeutige Ortsangabe Frankfurt ergibt sich beispielsweise folgende unbekannte Wahrscheinlichkeitsverteilung:

$$\begin{aligned} (1) \quad & P(\text{Geburtsort} = \text{Frankfurt am Main} \mid \text{Ortsangabe} = \text{„Frankfurt“}) + \\ & P(\text{Geburtsort} = \text{Frankfurt an der Oder} \mid \text{Ortsangabe} = \text{„Frankfurt“}) = 1 \end{aligned}$$

Zur Auflösung der Mehrdeutigkeit von Ortsnamen innerhalb von Staaten gibt es unterschiedliche Strategien.

<sup>19</sup> Bei der Ortsangabe „Berlin“ wird beispielsweise unmittelbar von der gleichnamigen deutschen Hauptstadt ausgegangen, obwohl auch eine Ortschaft „Berlin Seedorf“ in Schleswig-Holstein existiert.

Einen Überblick bietet Leidner (2008). Beispielsweise können die Wahrscheinlichkeiten (P) als Funktion (F) der Bevölkerungsgröße dargestellt werden.

$$\begin{aligned} (2) \quad & P(\text{Geburtsort} = \text{Frankfurt am Main} \mid \text{Ortsangabe} = \text{„Frankfurt“}) + \\ & P(\text{Geburtsort} = \text{Frankfurt an der Oder} \mid \text{Ortsangabe} = \text{„Frankfurt“}) \\ & = F(\text{Bevölkerung Frankfurt am Main}) + \\ & F(\text{Bevölkerung Frankfurt an der Oder}) \end{aligned}$$

Im Zusammenhang mit Geburtsorten von Schutzsuchenden ist diese Vorgehensweise aus verschiedenen Gründen nicht geeignet: Zum einen stehen für Geburtsorte von Schutzsuchenden oftmals keine aktuellen und verlässlichen Bevölkerungszahlen zur Verfügung. Zum anderen ist fraglich, ob die Größe der Gesamtbevölkerung eines Ortes auch die gesuchte Wahrscheinlichkeitsverteilung für Schutzsuchende in Deutschland approximiert.

Alternativ ist es möglich, die beobachteten Häufigkeitsverteilungen in den vorhandenen Daten zu nutzen. Dabei wird sich zunutze gemacht, dass die Mehrdeutigkeit von Ortsangaben durch identifizierende Namenszusätze, wie die zusätzliche Nennung des Bundes-

### Grafik 4

#### Schritte zur Auflösung der Mehrdeutigkeit von Ortsangaben

| Ortsangabe            | Kandidat 1            | Kandidat 2            | Kandidat 3              |
|-----------------------|-----------------------|-----------------------|-------------------------|
| Frankfurt             | Frankfurt am Main     | Frankfurt an der Oder | Frankfurt Sachsenhausen |
| Frankfurt am Main     | Frankfurt am Main     |                       |                         |
| Frankfurt an der Oder | Frankfurt an der Oder |                       |                         |

↓ mehrdeutige Ortsangaben und valide Kandidaten identifizieren

| Ortsangabe | Kandidat 1        | Kandidat 2            |
|------------|-------------------|-----------------------|
| Frankfurt  | Frankfurt am Main | Frankfurt an der Oder |

↓ Wahrscheinlichkeitsverteilung aus beobachteten Daten ableiten

| Ortsangabe | Kandidat 1        | Anzahl Kandidat 1 | Kandidat 2            | Anzahl Kandidat 2 |
|------------|-------------------|-------------------|-----------------------|-------------------|
| Frankfurt  | Frankfurt am Main | 90                | Frankfurt an der Oder | 10                |

↓ stochastische Zuordnung

| Ortsangabe | Zufallszahl | Prob(Kandidat 1) | Prob(Kandidat 2) | Ort               |
|------------|-------------|------------------|------------------|-------------------|
| Frankfurt  | 0,71        | 0,90             | 0,10             | Frankfurt am Main |

2023 - 0013

staates oder eines Flusses, aufgelöst werden kann. In großen Datensätzen ist davon auszugehen, dass für die meisten mehrdeutigen Ortsangaben auch ausreichend eindeutige Ortsangaben mit identifizierenden Zusätzen enthalten sind, um eine Wahrscheinlichkeitsverteilung über relative Häufigkeiten zu approximieren. ➤ Grafik 4

- (3)  $P(\text{Geburtsort} = \text{Frankfurt am Main} \mid \text{Ortsangabe} = \text{„Frankfurt“}) +$   
 $P(\text{Geburtsort} = \text{Frankfurt an der Oder} \mid \text{Ortsangabe} = \text{„Frankfurt“})$   
 $= h(\text{Frankfurt am Main}) + h(\text{Frankfurt an der Oder})$

## 4.3 Geokodierung

Sobald auch den mehrdeutigen Ortsangaben ein geografischer Ort zugewiesen wurde, können die Geokoordinaten aus OSM angespielt werden. Geokoordinaten können in unterschiedliche Koordinatensysteme eingebettet sein. OpenStreetMap nutzt das weitverbreitete World Geodetic System 1984 (WGS 84) als Grundlage für die Positionsangaben. Das WGS 84 teilt die Erde in 360 Längengrade von –180 bis 180 Grad und 180 Breitengrade von –90 bis 90 Grad. Die Geokoordinaten werden als Dezimalgrad angegeben. Der Standort Wiesbaden des Statistischen Bundesamtes wird in diesem Koordinatensystem beispielsweise mit Breitengrad 50,0708 und Längengrad 8,2520 verortet.

## 5

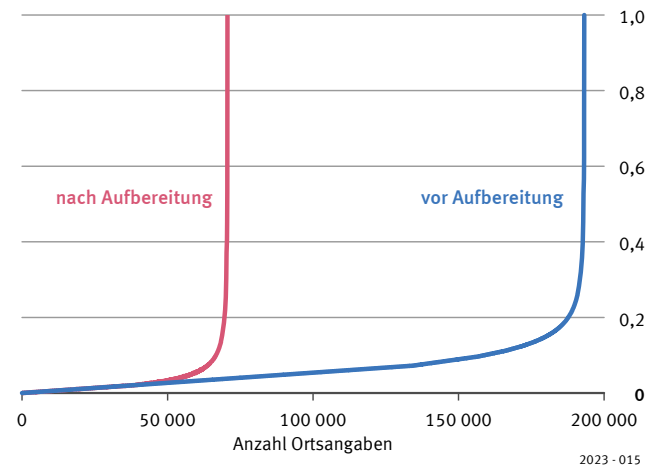
### Qualitätsbewertung der aufbereiteten Daten

Vor der Aufbereitung befanden sich rund 193 000 unterschiedliche Angaben zu Geburtsorten in den AZR-Daten. Nach der Aufbereitung sind es noch knapp 71 000. Der Anteil an einzigartigen Ortsangaben sinkt von 70 % in den Originaldaten auf 55 % in den aufbereiteten Daten. Gleichzeitig ist der Anteil der zehn häufigsten Angaben von 22 auf 30 % gestiegen. ➤ Grafik 5

Die gestiegene Konzentration spiegelt einerseits die Auflösung vieler unterschiedlicher Schreibweisen in den Originalangaben wider. Andererseits ist in einigen Regionen eine systematische Verzerrung von ländlichen

**Grafik 5**

Kumulative Häufigkeitsverteilung der Ortsangaben im Ausländerzentralregister vor und nach der Aufbereitung  
Anteil an den Schutzsuchenden in %



Gebieten in Richtung von Provinz- beziehungsweise Distrikthauptstädten zu beobachten. Ursache hierfür ist, dass in einigen Staaten administrative Verwaltungseinheiten die gleichen Namen wie deren Hauptstädte tragen. Syrien ist beispielsweise in 14 Gouvernements unterteilt, die alle den gleichen Namen tragen wie ihre Hauptstädte. Aleppo ist zum Beispiel gleichzeitig der Name des Gouvernements im Norden Syriens und der Name der Hauptstadt dieses Gouvernements. Auch in Afghanistan, dem Irak und der Türkei tragen die höchsten administrativen Gebietsgliederungen teilweise die gleichen Namen wie ihre Hauptstädte. Wurde neben dem Ortsnamen auch der Name der administrativen Verwaltungseinheit registriert, zum Beispiel ‚Afrin in Aleppo‘, wird teilweise fälschlich die Hauptstadt zugewiesen. Tragen administrative Verwaltungseinheiten die gleichen Namen wie ihre Hauptstädte, sollte die regionale Verteilung deshalb nicht unterhalb der Ebene dieser Verwaltungseinheiten ausgewertet werden.<sup>10</sup>

Insgesamt konnten 42 000 Ortsangaben keinem geografischen Ort zugeordnet werden. Die nicht zugeordneten Angaben repräsentieren 124 000 Schutzsuchende.

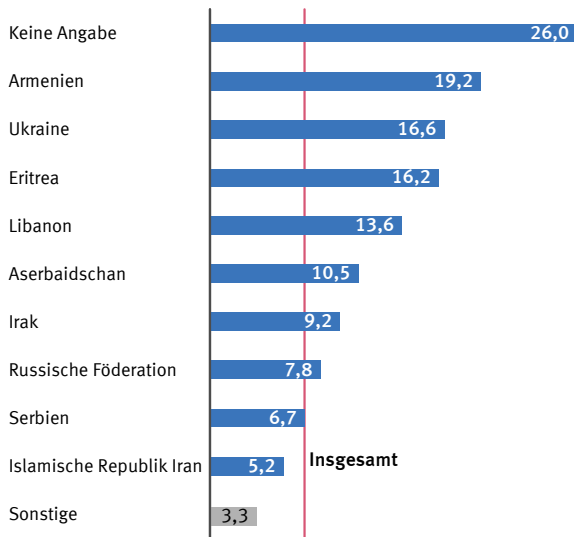
<sup>10</sup> In diesen Fällen können den Ortsangaben die falschen Geokoordinaten zugeordnet sein – der Ortsangabe „Afrin in Aleppo“ beispielsweise die Geokoordinaten der Stadt Aleppo. Neben den Geokoordinaten kann den OSM-Adressdaten aber auch eine Zuordnung zu höheren administrativen Ebenen entnommen werden. Bei einer Analyse auf Ebene der Gouvernements stimmt die Zuordnung wieder, da sowohl Afrin als auch Aleppo im Gouvernement Aleppo liegen.

Daraus ergibt sich ein Anteil fehlender Werte von 6,7 %. Höhere Anteile zeigen sich für die Geburtsstaaten, in denen nicht lateinische Alphabete genutzt werden.

➤ Grafik 6

**Grafik 6**

**Fehlende Werte nach Geburtsstaaten nach der Aufbereitung in %**



2023 - 016

Eine weitere Qualitätseinschätzung liefert ein Vergleich der Ergebnisse mit einer Geokodierung deutscher Geburtsorte über den Geokodierungsdienst des Bundesamtes für Kartographie und Geodäsie (BKG; BKG Geocoder Version 1.1). Um eine möglichst große Vergleichsgruppe zu erhalten, werden für diesen Vergleich die Geburtsorte aller am 31. Dezember 2020 im AZR registrierten und in Deutschland geborenen Ausländerinnen und Ausländer herangezogen.

Für 95 % der rund 188 000 in Deutschland geborenen Schutzsuchenden ergeben beide Vorgehensweisen eine übereinstimmende geografische Ortsangabe.<sup>11</sup> Die Analyse der abweichend kodierten Zuweisungen zeigt, dass bei mehrdeutigen Ortsangaben unterschiedliche Zuweisungen erfolgen. Die meisten Abweichungen entstehen durch die mehrdeutige Ortsangabe ‚Oberhausen‘. Bei dieser Angabe ergibt sich aus den OSM-Daten

<sup>11</sup> Als Übereinstimmung gelten Fälle, deren Angaben zum Breitengrad nicht mehr als 0,045 Grad (5 km) und deren Angaben zum Längengrad nicht mehr als 0,065 Grad (zwischen 4,1 km und 4,9 km) abweichen.

als wahrscheinlichster Kandidat die kreisfreie Großstadt Oberhausen in Nordrhein-Westfalen. Der Geokodierungsdienst des BKG hingegen entscheidet sich für die Gemeinde Oberhausen bei Kirn in Rheinland-Pfalz.

## 6

### Ergebnisse

Die Ergebnisse werden nachträglich als interaktives Kartenangebot zur Verfügung gestellt. Die meisten Menschen, die sich Ende 2020 als Schutzsuchende in Deutschland aufhielten, kamen aus Syrien, genauer gesagt aus den Gouvernements Aleppo, Damaskus und Al-Hasaka.<sup>12</sup> Mit Nordrhein-Westfalen ist auch ein deutsches Bundesland unter den häufigsten Herkunftsregionen vertreten. Aufgrund der jungen Altersstruktur der 2015 und 2016 zugewanderten Schutzsuchenden werden seitdem vermehrt Kinder als Schutzsuchende in Deutschland geboren (Statistisches Bundesamt, 2021a). ➤ Tabelle 2

**Tabelle 2**

**Häufigste Geburtsorte von Schutzsuchenden**

| Geburtsort          | Geburtsstaat | Anzahl           | Anteil in % |
|---------------------|--------------|------------------|-------------|
| Aleppo              | Syrien       | 131 260          | 7,1         |
| Damaskus            | Syrien       | 119 235          | 6,4         |
| Al-Hasaka           | Syrien       | 91 310           | 4,9         |
| Nordrhein-Westfalen | Deutschland  | 53 615           | 2,9         |
| Kabul               | Afghanistan  | 36 600           | 2,0         |
| Ninawa              | Irak         | 36 260           | 2,0         |
| Teheran             | Iran         | 33 280           | 1,8         |
| Homs                | Syrien       | 28 325           | 1,5         |
| Bagdad              | Irak         | 26 770           | 1,4         |
| Idlib               | Syrien       | 26 745           | 1,4         |
| Sonstige            |              | 1 278 491        | 68,6        |
| <b>Insgesamt</b>    |              | <b>1 856 785</b> | <b>100</b>  |

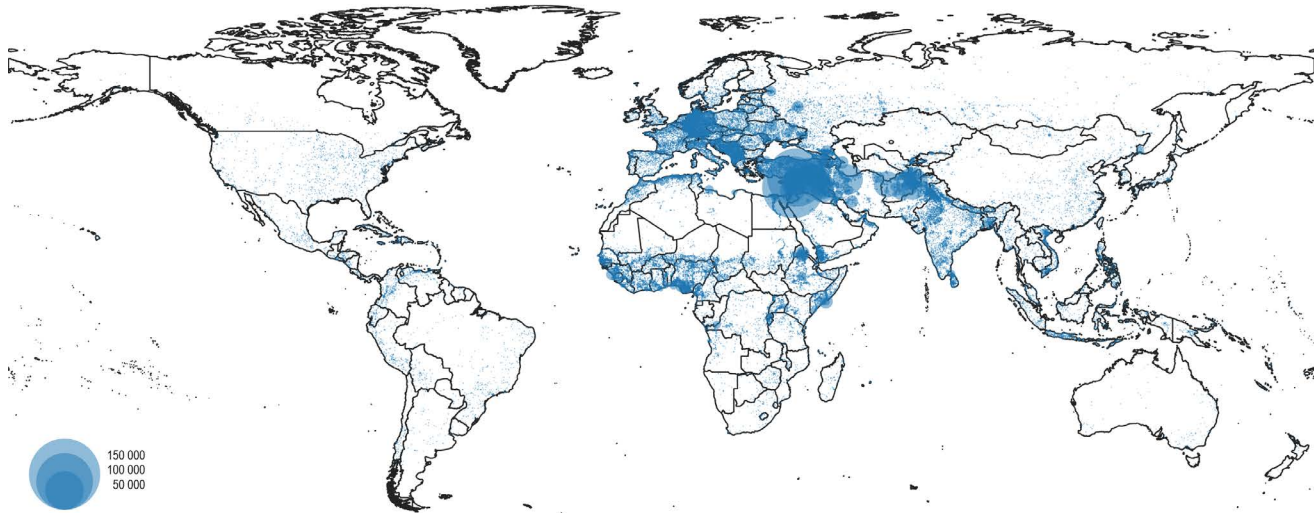
Auswertung auf der höchsten innerstaatlichen administrativen Ebene.

Die weltweite Verteilung der Geburtsorte von Schutzsuchenden ist Ende 2020 geprägt von der hohen Anzahl an Schutzsuchenden aus Syrien, Afghanistan und dem Irak. Dennoch sind auch weitere Herkunftsregionen wie die Balkanstaaten, Nordafrika und die Sahelzone

<sup>12</sup> Der Geburtsort ist nur ein Indikator für die Herkunft und muss nicht mit dem Ort übereinstimmen, an dem die Person aufgewachsen ist oder ihren Lebensmittelpunkt hat oder hatte.

### Grafik 7

Weltweite Verteilung der Geburtsorte von Schutzsuchenden



sichtbar. Die Geburtsorte liegen aber auch in Deutschland und anderen EU-Mitgliedstaaten. Dies ist ein Beleg dafür, dass der Geburtsstaat für viele Schutzsuchende nicht mehr der Staat ist, aus dem die Eltern ursprünglich geflohen sind. [↘ Grafik 7](#)

Die aufbereiteten Daten zu den Geburtsorten geben erstmals Einblicke in die regionale Herkunft dieser Menschen innerhalb einzelner Staaten. Beispielsweise kommt knapp die Hälfte aller in Syrien geborenen Schutzsuchenden aus dem nordsyrischen Gouvernement Aleppo (25 %) oder dem Gouvernement Damaskus (23 %). [↘ Grafik 8](#) auf Seite 38

Die regionale Verteilung innerhalb der Staaten kann Hinweise auf unterschiedliche Fluchtursachen enthalten. Dies wird am Beispiel der Türkei deutlich. Nach dem gescheiterten Putschversuch im Juli 2016 folgte in der Türkei ein zweijähriger Ausnahmezustand, in dem Staatspräsident Erdoğan staatliche und gesellschaftliche Institutionen ohne parlamentarische Beteiligung gezielt umstrukturierte (Gieler, 2021). Im Jahr 2017 wurden die Befugnisse des Staatspräsidenten per Verfassungsreferendum mit knapper Mehrheit auch langfristig erheblich erweitert. Seit 2016 steigt die Zahl türkischer Schutzsuchender in Deutschland stetig an, nachdem sie zuvor rückläufig war. Der umstrittene und polarisierende Umbau von Politik, Verwaltung und Gesellschaft spiegelt sich auch in der regionalen Herkunft von türkischen Schutzsuchenden in Deutschland wider. Schutz-

suchende, die vor 2016 nach Deutschland kamen, sind großteils in den kurdisch geprägten Gebieten im Südosten der Türkei geboren. Seit 2016 kommen türkische Schutzsuchende hingegen häufiger aus Istanbul, Ankara und Izmir, und damit aus nordwestlichen Provinzen, die 2017 mehrheitlich gegen die Verfassungsänderungen gestimmt hatten. [↘ Grafik 9](#) auf Seite 39

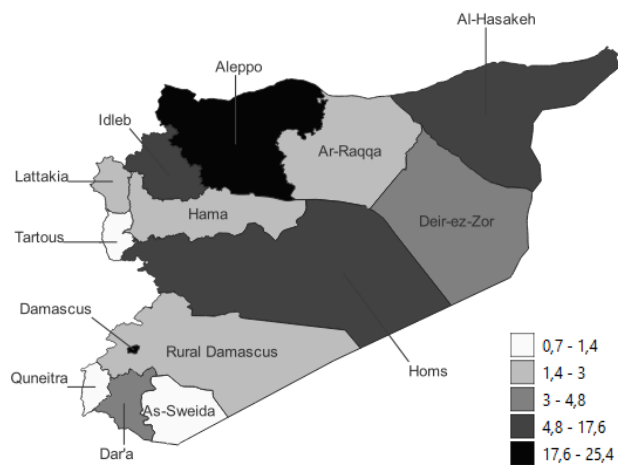
Angaben zu Geburtsorten ermöglichen darüber hinaus, Kettenmigration, Migrationsnetzwerke und Cluster in der regionalen Verteilung der ausländischen Bevölkerung detaillierter zu analysieren. Beispielsweise liefern die Daten Hinweise dafür, dass Schutzsuchende aus unterschiedlichen türkischen Provinzen unterschiedliche Präferenzen im Hinblick auf ihren Wohnort in Deutschland haben. So lebte mit 4,7 % der größte Teil der Schutzsuchenden aus Istanbul Ende 2020 in Berlin. Türkische Schutzsuchende aus der kurdisch geprägten Provinz Mardin hingegen waren zumeist in Bremen registriert (4,9 %). [↘ Grafik 10](#) auf Seite 40

## Grafik 8

### Verteilung der Geburtsorte in Syrien, dem Irak und Afghanistan

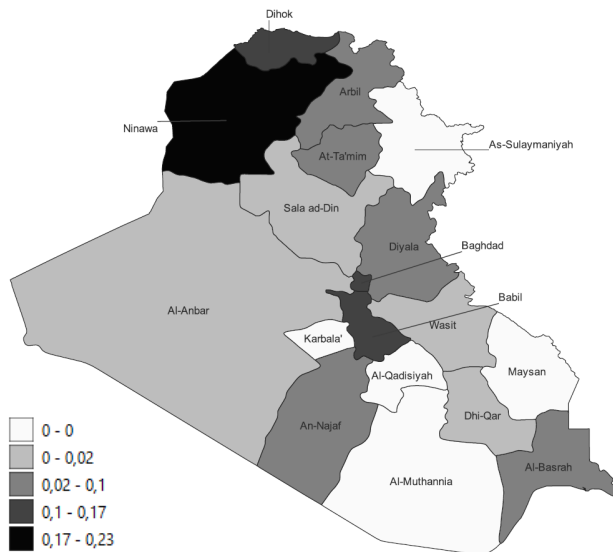
#### Syrien

Regionale Ebene: 14 Gouvernements



#### Irak

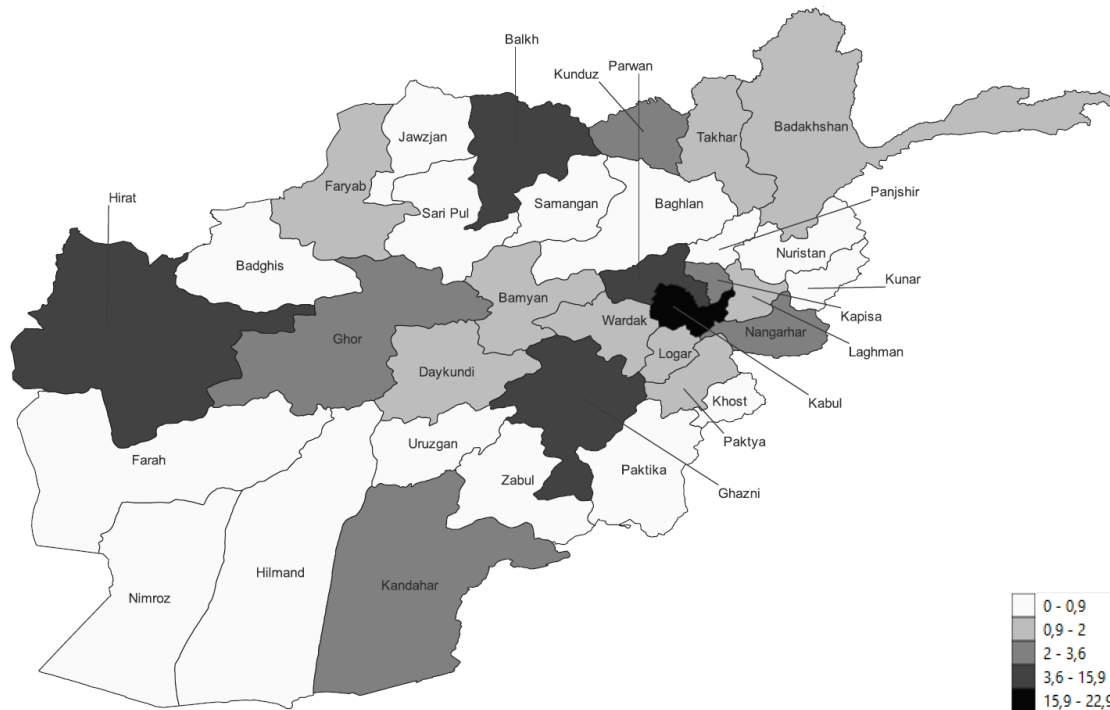
Regionale Ebene: 18 Gouvernements<sup>1</sup>



<sup>1</sup> 5 900 Angaben zum Geburtsort autonome Region Kurdistan wurden anteilmäßig auf die Gouvernements Erbil, Dahuk und Halabdscha aufgeteilt.

#### Afghanistan

Regionale Ebene: 34 Distrikte

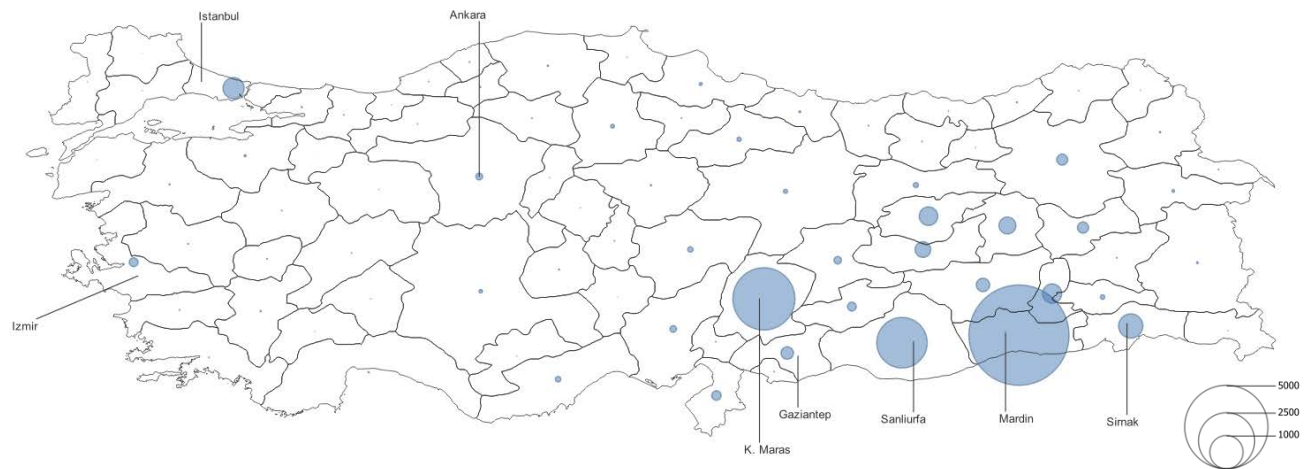


## Grafik 9

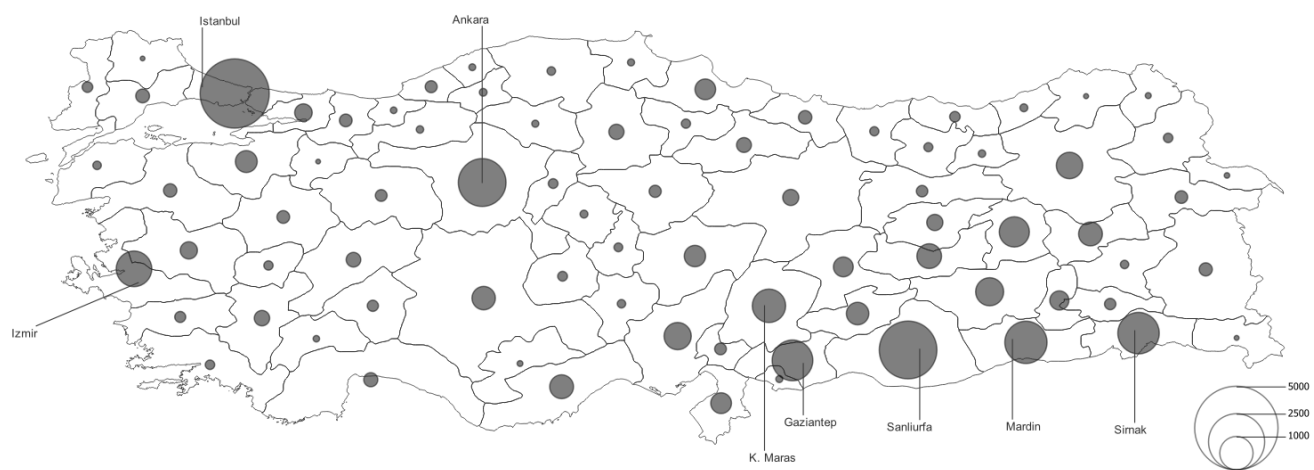
### Regionale Verteilung von in der Türkei geborenen Schutzsuchenden

Regionale Ebene: 81 Provinzen

mit Ersteinreise vor 2016



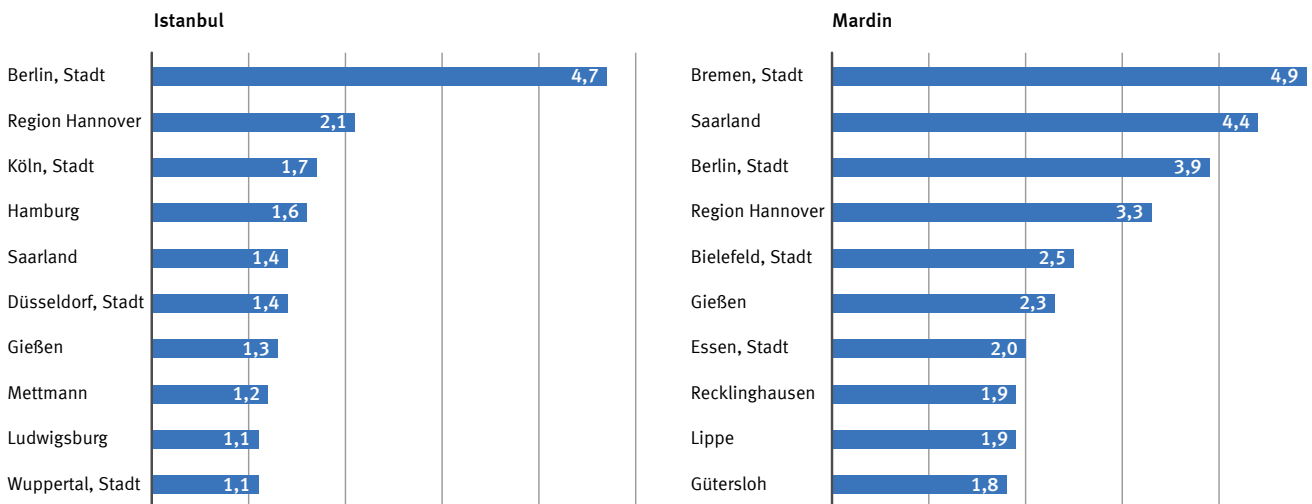
mit Ersteinreise seit 2016



2023 - 0018

**Grafik 10**

**Regionale Verteilung von türkischen Schutzsuchenden aus Istanbul und Mardin nach Wohnorten in Deutschland Ende 2020 in %**



2023 - 020

## 7

### Fazit

Informationen zum Geburtsort im AZR bieten vielfältige neue Auswertungsmöglichkeiten in der Ausländerstatistik und der Statistik über Schutzsuchende. Analysen zur Herkunft können damit in einem kleinräumigeren regionalen Kontext innerhalb der Staatsgrenzen durchgeführt werden. Über die Geokoordinaten können die AZR-Daten beispielsweise mit Klimadaten kombiniert werden und Beiträge zur Klimafolgenforschung leisten.

Voraussetzung für die statistische Nutzung ist eine sorgfältige Aufbereitung und Qualitätssicherung der unstrukturiert als Freitext erhobenen Angaben. Im Zuge der Aufbereitung werden unterschiedliche Repräsentationen der gleichen geografischen Orte (synonyme Ortsbezeichnungen) und gleiche Repräsentationen unterschiedlicher geografischer Orte (homonyme beziehungsweise mehrdeutige Ortsbezeichnungen) aufgelöst. Von zentraler Bedeutung für den Aufbereitungsprozess ist eine möglichst vollständige Vergleichsdatenbasis. Mit rund 4,7 Millionen Einträgen zu bewohnten Siedlungen bietet OpenStreetMap eine der größten weltweiten Geodatenansammlungen und stellt diese unter der Open Database License zur freien Verfügung.

Getestet wurde der Aufbereitungsprozess anhand der Angaben zu Geburtsorten von knapp 1,9 Millionen Schutzsuchenden, die am 31. Dezember 2020 im AZR registriert waren. Dieser Anwendungsfall entspricht einem Härtefall, da die Hauptherkunftsländer im arabischen Sprachraum liegen und bei der Registrierung von Schutzsuchenden offizielle Dokumente häufig fehlen. Nach der Aufbereitung verblieb mit 6,7 % dennoch nur ein geringer Anteil an Angaben, die keinem geografischen Ort zugewiesen werden konnten. Erwartungsgemäß sind die Anteile an nicht zugeordneten Ortsangaben in Staaten höher, in denen ein nicht lateinisches Alphabet genutzt wird. Dabei enthalten die Angaben, die keinem geografischen Ort zugeordnet werden konnten, zumeist auch keinen verwertbaren Hinweis. Oftmals handelt es sich hierbei um ISO-Codes oder Angaben ohne jeden semantischen Inhalt. Probleme entstehen dort, wo administrative Verwaltungseinheiten (Regionen, Provinzen, Distrikte) den gleichen Namen tragen wie ihre Hauptstädte. Hier sollte die regionale Verteilung nicht unterhalb der Ebene dieser administrativen Einheiten ausgewertet werden.

Unter den genannten Einschränkungen zeigt die Machbarkeitsstudie: Die administrativen Daten zu Geburtsorten können mithilfe der OSM-Geodaten aufbereitet und für Analysen genutzt werden. 

### LITERATURVERZEICHNIS

---

Canan, Coşkun/Eberle, Jan. *Geburtsstaat und Geburtsort im Ausländerzentralregister – Nutzungsmöglichkeiten für die amtliche Statistik*. In: WISTA Wirtschaft und Statistik. Ausgabe 2/2022, Seite 53 ff.

Conti, Cinzia/Cimbelli, Alessandro. *Exploring international migration at subnational scale: the Italian case*. In: UNECE Session on Migration Statistics Working Paper. 2018. [Zugriff am 15. Dezember 2022]. Verfügbar unter: [unece.org](https://unece.org)

DeLozier, Grant/Baldrige, Jason/London, Loretta. *Gazetteer-independent toponym resolution using geographic word profiles*. In: Twenty-Ninth AAAI Conference on Artificial Intelligence. 2015. Seite 2382 ff.

Eberle, Jan. *Schutzsuchende*. In: WISTA Wirtschaft und Statistik. Ausgabe 1/2019, Seite 19 ff.

Fleischer, Henning. *Entwicklung der Ausländerzahl seit 1987*. In: Wirtschaft und Statistik. Ausgabe 9/1989, Seite 594 ff.

Gieler, Wolfgang. *Politische Entwicklungen in der Türkei seit dem Putschversuch*. In: Henrich, Christian Johannes/Gieler, Wolfgang (Herausgeber). Die Außenpolitik der Türkei im Mittleren Osten. Wiesbaden 2021, Seite 27 ff.

Goldberg, Daniel W./Wilson, John P./Knoblock, Craig A. *From text to geographic coordinates: the current state of geocoding*. In: URISA Journal. Jahrgang 19. Ausgabe 1/2007, Seite 33 ff.

Herfort, Benjamin/Lautenbach, Sven/Porto de Albuquerque, João/Anderson, Jennings/Zipf, Alexander. *The evolution of humanitarian mapping within the OpenStreetMap community*. In: Scientific Reports. Jahrgang 11. Artikel 3037/2021.

Lehmann, Jens/Isele, Robert/Jakob, Max/Jentzsch, Anja und weitere. *DBpedia – a large-scale, multilingual knowledge base extracted from Wikipedia*. In: Semantic web 6.2.2014. Seite :167 ff. DOI: [10.3233/SW-140134](https://doi.org/10.3233/SW-140134)

Leidner, Jochen L. *Toponym Resolution in Text: Annotation, Evaluation and Applications of Spatial Grounding of Place Names*. 2007.

OpenStreetMap Foundation. *Licence Community Guideline Geocoding*. 2017. [Zugriff am 15. Dezember 2022]. Verfügbar unter: [wiki.osmfoundation.org](https://wiki.osmfoundation.org)

Seto, Toshikazu/Kanasugi, Hiroshi/Nishimura, Yuichiro. *Quality verification of volunteered geographic information using osm notes data in a global context*. In: International Journal of Geo-Information. Jahrgang 9. Ausgabe 6/2020, Seite 372 ff.

Šimbera, Jan/Drbohlav, Dušan/Štych, Přemysl. *Geocoding Freeform Placenames: An Example of Deciphering the Czech National Immigration Database*. In: ISPRS International Journal of Geo-Information. Ausgabe 10.5/2021, Seite 335 ff.

Statistisches Bundesamt. *Digitale Agenda des Statistischen Bundesamtes. Version 2.1. 03/2019*. [Zugriff am 15. Dezember 2022]. Verfügbar unter: [www.destatis.de](https://www.destatis.de)

## LITERATURVERZEICHNIS

---

Statistisches Bundesamt. *Seit 2015 werden mehr Schutzsuchende in Deutschland geboren*. Pressemitteilung Nr. N 39 vom 17. Juni 2021. 2021a. [Zugriff am 15. Dezember 2022]. Verfügbar unter: [www.destatis.de](http://www.destatis.de)

Statistisches Bundesamt. *Qualitätsbericht Ausländerstatistik*. 2021b. [Zugriff am 15. Dezember 2022]. Verfügbar unter: [www.destatis.de](http://www.destatis.de)

Statistisches Bundesamt. *Methodische Hinweise zur Ausländerstatistik*. 2021c. [Zugriff am 15. Dezember 2022]. Verfügbar unter: [www.destatis.de](http://www.destatis.de)

Statistisches Bundesamt. *Staats- und Gebietssystematik*. 2021d. [Zugriff am 15. Dezember 2022]. Verfügbar unter: [www.destatis.de](http://www.destatis.de)

## RECHTSGRUNDLAGEN

---

Gesetz über das Ausländerzentralregister (AZR-Gesetz) vom 2. September 1994 (BGBl. I Seite 2265), das zuletzt durch Artikel 5c des Gesetzes vom 23. Mai 2022 (BGBl. I Seite 760) geändert worden ist.

Gesetz zur Verbesserung der Registrierung und des Datenaustausches zu aufenthalts- und asylrechtlichen Zwecken (Datenaustauschverbesserungsgesetz) vom 2. Februar 2016 (BGBl. I Seite 130).

Gesetz zur Weiterentwicklung des Ausländerzentralregisters vom 9. Juli 2021 (BGBl. I Seite 2467).

Zweites Gesetz zur Verbesserung der Registrierung und des Datenaustausches zu aufenthalts- und asylrechtlichen Zwecken (Zweites Datenaustauschverbesserungsgesetz – 2. DAVG) vom 4. August 2019 (BGBl. I Seite 1131).

# BEDEUTUNG RAUMSTRUKTURELLER UND ARBEITSMARKTRELEVANTER FAKTOREN BEI INNERDEUTSCHEN WANDERUNGEN VON GEFLÜCHTETEN

Johannes Weber



Johannes Weber

ist Volkswirt und seit 2018 für das Forschungszentrum Migration, Integration und Asyl des Bundesamtes für Migration und Flüchtlinge tätig. Seine derzeitigen Forschungsschwerpunkte umfassen quantitative Analysen zu Integrationsprozessen in ländlichen Räumen sowie zur Binnenmobilität von Geflüchteten.

➤ **Schlüsselwörter:** Geflüchtete – Binnenmobilität – Ausländerzentralregister – Urbanisierung – Arbeitslosigkeit – Integration

## ZUSAMMENFASSUNG

Der Beitrag präsentiert vertiefte Analysen über den Zusammenhang zwischen raumstrukturellen sowie arbeitsmarktrelevanten Merkmalen und der Binnenmobilität von Geflüchteten mit Schutzstatus in Deutschland in den Jahren 2015 bis 2020. Genutzt wird dabei eine vom Institut für Arbeitsmarkt- und Berufsforschung erstellte Klassifizierung von Arbeitsagenturbezirken. Demzufolge sind die beobachtbaren Muster vor allem auf Wanderungen im Jahr 2016 zurückzuführen, als die großstädtisch geprägten Kreise mit hoher Arbeitslosigkeit im Westen Deutschlands große Wanderungsgewinne verzeichneten. Diese resultieren vor allem aus Fortzügen von Geflüchteten mit Schutzstatus aus ländlichen Kreisen im Osten Deutschlands, während die dortigen städtischen Kreise nur marginale Wanderungsverluste aufweisen.

➤ **Keywords:** refugees – internal mobility – Central Register of Foreigners – urbanisation – unemployment – integration

## ABSTRACT

*The article presents in-depth analyses of the relationship between spatial-structural and labour market relevant variables and the internal mobility of refugees with protection status in Germany between 2015 and 2020. A classification of employment agency districts set up by the Institute of Employment Research is used. The patterns observed are mainly due to migration movements in 2016 when large migration gains were recorded for the metropolitan districts with high unemployment in western Germany. They resulted mainly from departures of refugees with migration status from rural districts in eastern Germany while only marginal migration losses were observed in the urban districts there.*

## 1

### Einleitung

Im Vergleich zu anderen (migrantischen und nicht migrantischen) Bevölkerungsgruppen stellt die Wohnortwahl von Geflüchteten eine Besonderheit dar, da sowohl der Wohnort als auch die Art der Unterkunft in Deutschland bis zu drei Jahre nach Erteilung eines Schutzstatus gesetzlich reguliert werden. Einschränkungen der Wohnortwahl betreffen sowohl Personen im Asylverfahren (Residenzpflicht) als auch diejenigen mit anerkanntem oder abgelehntem Schutzstatus.<sup>1</sup> Bei Menschen mit einer positiven Entscheidung im Asylverfahren spielt insbesondere die 2016 mit dem Integrationsgesetz eingeführte und 2019 entfristete Wohnsitzregelung nach § 12a Aufenthaltsgesetz eine Rolle: Diese bindet sie drei Jahre lang an das Bundesland, in dem das Asylverfahren durchgeführt wurde. Zudem erfolgt in einem Teil der Bundesländer eine Zuweisung der Schutzberechtigten in eine bestimmte Kommune, auch gibt es Zuzugs-sperren für bestimmte Kommunen. Trotz der Bedeutung des Wohnumfelds für gelingende Integrationsprozesse von Geflüchteten und der gesetzlichen Bestimmungen, denen sie zu Beginn ihres Aufenthalts bei der Wohnortwahl unterliegen, liegen empirische Erkenntnisse über das Wanderungsverhalten und die Wohnsituation von Geflüchteten bislang nur vereinzelt vor (unter anderem Tanis, 2022; Weber, 2022).

Erste Erkenntnisse über die Binnenmobilität von ausländischen Staatsangehörigen in Deutschland lieferte Brückner (2019), der erstmals das Ausländerzentralregister (AZR) als Längsschnittdatensatz nutzte, um Wanderungsbiografien bundesweit zu rekonstruieren. Dessen Auswertungen zur Wanderungsrichtung und -entfernung basieren auf den jährlichen „Bestandsdaten für Statistikzwecke“ des AZR für den Zeitraum 2007 bis 2017. Sie deuten darauf hin, dass sich die Binnenmobilität von ausländischen Staatsangehörigen in Deutschland „größtenteils in der näheren Umgebung“ abspielt (Brückner, 2019, hier: Seite 52). Die Mehrheit der Wanderungen zwischen den Kreisen fand im genannten Zeitraum innerhalb des jeweiligen Bundeslands statt. Die zehn (Land-)Kreise mit der stärksten Abwanderung

lagen sämtlich im Osten, die zehn Kreise beziehungsweise kreisfreien Städte mit der geringsten Abwanderung hingegen alle im Westen Deutschlands (Brückner, 2019, hier: Seite 50).

Der Autor dieses Beitrags stellte im Jahr 2022 anhand einer explorativen Analyse von Daten aus dem Gesamtbestand des AZR, die sich auf der Vorgehensweise von Brückner stützt, erste Erkenntnisse über das Binnenwanderungsverhalten von Geflüchteten mit Schutzstatus<sup>2</sup> vor (Weber, 2022). Im Mittelpunkt der Studie standen innerdeutsche, kreisübergreifende Wanderungen der Jahre 2015 bis einschließlich 2020. Die Analysen zur räumlichen Verteilung, zu den Netto-Wanderungssalden und zur Wanderungsrichtung zeigten, dass (Groß-)Städte mit einer überdurchschnittlich hohen Arbeitslosigkeit bei Geflüchteten mit Schutzstatus eine hohe Attraktivität als Wohnort haben. Dagegen hatten ländliche Kreise eine geringere Anziehungskraft. Dies galt insbesondere für sehr ländliche Kreise, aus denen eine starke Abwanderung von Geflüchteten mit Schutzstatus stattfand. Der Befund, dass Geflüchtete mit Schutzstatus mehrheitlich in urbane Regionen ziehen, stand dabei nicht im Widerspruch zu den Erkenntnissen über den festgestellten positiven Zusammenhang zwischen Arbeitslosenquote und Binnenmobilität. Vielmehr bestätigten sich diese beiden Beobachtungen, da kreisfreie Städte meist eine höhere Arbeitslosenquote aufweisen als ländliche Kreise. So zählten zu den Kreisen mit der höchsten Arbeitslosenquote mehrheitlich Großstädte, darunter auch die drei Stadtstaaten Berlin, Hamburg und Bremen.

Eine Erklärung für dieses Wanderungsmuster könnte sein, dass zu Beginn des Aufenthalts in Deutschland für die Gruppe der Geflüchteten mit Schutzstatus die lokalen Arbeitsmarktbedingungen (noch) eher nachrangig sind. Dagegen könnten andere Faktoren, wie ein stärkerer Ausbau der lokalen migrationsspezifischen Infrastrukturen oder ethnische Netzwerke, eine bedeutendere Rolle spielen als ökonomische Faktoren. Möglicherweise bestehen auch Informationsdefizite, da die Geflüchteten über keine vertieften Kenntnisse über den regionalen Arbeitsmarkt verfügen. Inwieweit Effekte der rechtlichen Regularien (staatlich gesteuerte Anfangs-

1 Zu den rechtlichen Regelungen der Wohnortwahl für alle genannten Gruppen siehe im Detail Tanis (2022) sowie Rösch und andere (2020).

2 Im Folgenden bezieht sich die Bezeichnung „Geflüchtete mit Schutzstatus“ auf Personen, die in Deutschland einen Asylantrag gestellt und einen Aufenthaltstitel aus dem humanitären Bereich innehaben. Zur detaillierten Beschreibung der Untersuchungsgruppe sowie der Schutzformen siehe Abschnitt 2.2.

verteilung von Asylsuchenden, Wohnsitzregelung für diejenigen mit Schutzstatus) sowie Überlagerungen der Arbeitslosenquote mit der Dimension Stadt-Land ursächlich für diese Beobachtungen sind, bleibt in dieser Studie unbeantwortet.

Daran anknüpfend verfolgt der vorliegende Beitrag vielmehr das Ziel, anhand einer SGB-III-Typisierung von Arbeitsagenturbezirken (Blien/Hirschenauer, 2017) vertiefte Erkenntnisse über den Zusammenhang zwischen raumstrukturellen sowie arbeitsmarktrelevanten Merkmalen und dem innerdeutschen Wanderungsverhalten von geflüchteten Menschen mit Schutzstatus zu generieren. Dabei stehen folgende Fragen im Mittelpunkt:

1. In welchem Zusammenhang stehen räumlich-strukturelle Merkmale und die regionale Arbeitsmarktlage mit der räumlichen Verteilung und dem Wanderungsverhalten von Geflüchteten mit Schutzstatus?
2. Können zeitliche Unterschiede bei der Wanderungshäufigkeit, das heißt Differenzen je nach betrachtetem Jahr, sowie regionale Wandermuster beobachtet werden?

Zunächst thematisiert Kapitel 2 die Datengrundlage und Methodik zur Identifikation von Wanderungsbewegungen sowie die Untersuchungsgruppe. Zudem wird die SGB-III-Typisierung der Agenturbezirke vorgestellt. Kapitel 3 rückt die räumliche Verteilung der Geflüchteten mit Schutzstatus sowie ihr Wanderungsverhalten in den Mittelpunkt. Dabei erfolgt hinsichtlich der Analysen der Binnenwanderungen eine zeitliche sowie regionale Differenzierung. Das vierte Kapitel fasst die Erkenntnisse über den Zusammenhang zwischen raumstrukturellen und arbeitsmarktrelevanten Merkmalen der Kreise und der räumlichen Verteilung sowie des Wanderungsverhaltens der Untersuchungsgruppe zusammen und diskutiert deren Aussagekraft.

## 2

---

### Methodik, Untersuchungsgruppe und SGB-III-Typisierung der Agenturbezirke

---

#### 2.1 Längsschnittdaten aus dem Ausländerzentralregister

---

Im AZR werden bundesweit Angaben zu allen Personen gespeichert, die sich nicht nur vorübergehend, das heißt in der Regel länger als drei Monate, in Deutschland aufhalten und nicht die deutsche Staatsangehörigkeit besitzen. Geflüchtete werden mit der Registrierung eines Asylgesuchs im AZR erfasst, also noch bevor ein Asylantrag formell gestellt wurde. Neben dem Asylstatus und den nach dem Verfahren erteilten Aufenthaltstiteln werden weitere Sachverhalte im AZR gespeichert, beispielsweise das Jahr der Einreise, das Alter, das Geschlecht, der Familienstand und die Staatsangehörigkeit der registrierten Personen.

Auf Basis der im Jahr 2006 eingeführten anonymisierten AZR-Nummer ist eine Analyse der individuellen Mobilität von Geflüchteten mithilfe der Speichersachverhalte „aktenführende Ausländerbehörde“ (§ 3 Absatz 1 Nummer 1 der Verordnung zur Durchführung des Gesetzes über das Ausländerzentralregister [AZRG-DV]) sowie „Zuzug/Fortzug“ (§ 3 Absatz 1 Nummer 6 AZRG-DV) möglich. Bei einer Wanderung über eine Kreisgrenze hinweg wechselt die Zuständigkeit von der Ausländerbehörde des vorherigen Kreises auf die des neuen Wohnsitzes. Solch eine Wanderung wird im Regelfall von der Ausländerbehörde im AZR hinterlegt, in deren Zuständigkeit die über eine Kreisgrenze gewanderte Person am neuen Wohnort liegt. Die Ausländerbehörde fungiert damit in den Auswertungen als Proxy für den Wohnort auf Kreisebene, wodurch es möglich ist, Wanderungen über Kreis- sowie Bundesländergrenzen hinweg zu rekonstruieren.<sup>3</sup> Für die sechs Kreise im Saarland sowie die kreisfreien Städte Kassel und Cottbus und die umliegenden Landkreise Kassel und Spree-Neiße ist keine klare Kreiszuordnung der Wohnorte von Geflüchteten

---

<sup>3</sup> Umzüge innerhalb eines Kreises können anhand der verfügbaren Informationen im AZR nicht identifiziert werden. Für eine detaillierte Beschreibung der Rekonstruktion der Wanderungen siehe Brückner (2019) und Weber (2022).

möglich. Demzufolge beziehen sich die Aussagen zur Binnenwanderung der Untersuchungsgruppe auf Kreisebene auf Auswertungen zu insgesamt 391 Kreisen in Deutschland (286 Landkreise, 105 kreisfreie Städte).

## 2.2 Untersuchungsgruppe

Im Mittelpunkt der Untersuchung steht die innerdeutsche Mobilität der großen Zahl von Geflüchteten, die in den vergangenen Jahren nach Deutschland eingereist sind. Wie schon bei Weber (2022) bezieht die Analyse ausschließlich Personen mit ein, die in den Jahren 2015 bis 2019 erstmals nach Deutschland eingereist sind, zu diesem Zeitpunkt volljährig waren und einen Asylantrag gestellt haben.<sup>14</sup> Die Analyse betrachtet die Binnenwanderungen dieser Personengruppe bis zum 31. Dezember 2020 mit zwei zentralen Eingrenzungen: Sie schließt Wanderungen innerhalb der ersten 90 Tage nach Ersteinreise aus der Betrachtung aus, da es sich dabei vorwiegend um eine sogenannte „Zuweisungsmobilität“<sup>15</sup> handelt, die vor allem rechtlichen Regularien und weniger individuellen Präferenzen folgt. Zudem werden ausschließlich Wanderungen von Personen betrachtet, deren Asylantrag positiv beschieden wurde<sup>16</sup>. Grund für diese zweite Einschränkung ist, dass diese Personengruppe abgesehen von der Wohnsitzregelung nach § 12a Aufenthaltsgesetz einen größeren individuellen Entscheidungsspielraum bei kreisübergreifenden Wanderungen hat als Personen im Asylverfahren beziehungsweise mit abgelehntem Asylantrag. Deren Mobilität ist aufgrund rechtlicher Bestimmungen stärker

eingeschränkt (beispielsweise durch die Residenzpflicht nach § 56 Asylgesetz). ➔ **Tabelle 1**

Die Untersuchungsgruppe der Geflüchteten mit Schutzstatus umfasst insgesamt 470 823 Personen<sup>17</sup>. Hinsichtlich der individuellen Merkmale Alter, Aufenthaltsdauer sowie Geschlecht entsprechen diese Personen den allgemeinen Mustern der Fluchtmigration der vergangenen Jahre, die auch den Publikationen des Statistischen Bundesamtes zu Schutzsuchenden zu entnehmen sind (Statistisches Bundesamt, 2020). Demzufolge sind etwa zwei von drei Geflüchteten Männer: In den Jahren 2015 und 2016 sind vor allem männliche Schutzsuchende nach Deutschland eingereist, dieses Ungleichgewicht hat sich zwischen 2017 und 2019 etwas abgemildert, als auch vermehrt Frauen nach Deutschland flüchteten. Als Resultat dieser Entwicklung ist festzuhalten, dass sich geflüchtete Männer mit einem Schutzstatus durchschnittlich länger in Deutschland aufhalten als Frauen und zudem auch etwas jünger sind.

## 2.3 Vergleichstypen der SGB-III-Typisierung 2018

Um vertiefte Erkenntnisse über den Zusammenhang zwischen ökonomischen sowie raumstrukturellen Merkmalen und dem Wanderungsverhalten zu sammeln, greift der vorliegende Beitrag auf eine vom Institut für Arbeitsmarkt- und Berufsforschung (IAB) erstellte SGB-III-Typisierung der Agenturbezirke der Bundesagentur für Arbeit (BA) zurück (Blien/Hirschenauer, 2017). Die BA nutzt diese Agenturtypisierung bei Leistungsvergleichen

4 Für weitere Informationen zur Definition der Untersuchungsgruppe siehe Weber (2022).

5 Weiterführende Erläuterungen sowie Auswertungen über die sogenannte Zuweisungsmobilität bei Geflüchteten siehe Weber (2022).

6 Dies umfasst Personen, die im Besitz einer Asylberechtigung nach Artikel 16a Absatz 1 Grundgesetz oder einer Anerkennung als Flüchtling nach der Genfer Konvention (§ 3 Absatz 1 AsylG) sind beziehungsweise subsidiären Schutz (§ 4 Absatz 1 AsylG) oder ein nationales Abschiebeverbot (§ 60 Absatz 5 oder 7 AufenthG) zuerkannt bekommen haben.

7 Die Untersuchungsgruppe in der Studie von Weber (2022), auf der dieser Beitrag aufbaut, umfasst 471 250 Personen und weicht damit um 427 Personen von der hier untersuchten Gruppe von Geflüchteten mit Schutzstatus ab. Dies ist auf eine bei der aktuellen Analyse identifizierte Eintragung einer Meldebehörde im AZR zurückzuführen, die keine räumliche Zuordnung ermöglicht. Dies führt darüber hinaus zu einer Abweichung bei der Gesamtzahl an kreisübergreifenden Wanderungen.

**Tabelle 1**

Differenzierung der Geflüchteten mit Schutzstatus nach demografischen Merkmalen

|                                              | Insgesamt | Frauen  | Männer  |
|----------------------------------------------|-----------|---------|---------|
| Anzahl                                       | 470 823   | 167 577 | 303 015 |
| Durchschnittliches Alter in Jahren           | 34,8      | 36,1    | 34,1    |
| Durchschnittliche Aufenthaltsdauer in Jahren | 4,8       | 4,6     | 4,9     |

Quellen: Ausländerzentralregister (Stand: 31.12.2020), eigene Berechnungen. Bei 228 Personen lagen keine Informationen zum Geschlecht vor, drei gaben die Option „divers“ an.

zwischen den Arbeitsagenturen und ermöglicht damit, regionale Arbeitsmarktdisparitäten mit zu berücksichtigen. Ein Mehrwert der Typisierung besteht darin, räumliche und ökonomische Regionalindikatoren auf Ebene der Agenturbezirke, die auf Kreisebene abgebildet werden können, zu verknüpfen. Diese Verknüpfung stellt für das vorliegende Forschungsvorhaben eine geeignete Möglichkeit dar, innerdeutsche Wanderungsmuster von Geflüchteten mit Schutzstatus differenziert nach raumstrukturellen und arbeitsmarktrelevanten Kriterien zu untersuchen.

Die Klassifikation enthält insgesamt acht Indikatoren mit Zeitbezug zum Jahr 2016: die Arbeitslosenquote, den Erwerbspersonenanteil ohne Berufsabschluss, den Beschäftigtenanteil im Dienstleistungssektor (Tertiärisierungsgrad), den Beschäftigtenanteil am Arbeitsort in Relation zur Bevölkerung im Alter von 15 bis 64 Jahren (Arbeitsplatzbesatz), den Beschäftigtenanteil in Betrieben mit unter 100 Beschäftigten, zwei Variablen zur saisonalen Dynamik des regionalen Arbeitsmarkts und den Anteil der Arbeitslosen aus den acht wichtigsten nichteuropäischen Asylherkunftsländern<sup>8</sup> in Relation zu allen Arbeitslosen. Die Variable zur Fluchtmigration ist im Vergleich zu vorherigen Typisierungen neu hinzugekommen. Sie hat einen signifikant – wenn auch vergleichbar geringen – negativen Effekt auf die SGB-III-Integrationsquote im Jahr 2016<sup>9</sup> und zeigt damit, dass die Arbeitsmarktintegration von Geflüchteten besondere Herausforderungen mit sich bringt. Dies bedeutet zugleich, dass die SGB-III-Typisierung, in der die Variable zur Fluchtmigration enthalten ist, nicht unabhängig

ist von dem Binnenwanderungsverhalten und der damit einhergehenden räumlichen Verteilung der hier beobachteten Geflüchteten mit Schutzstatus. Bei der Interpretation der Ergebnisse ist diese Sachlage kritisch zu betrachten – auch wenn die Variable zur Fluchtmigration einen relativ geringen Effekt auf die Integrationsquote hat und die Situation im Jahr 2016 abbildet, während die Binnenwanderungen für die Jahre 2015 bis 2020 und die räumliche Verteilung der hier untersuchten Geflüchteten mit Schutzstatus zum Stichtag 31. Dezember 2020 analysiert werden.

Anhand der acht Variablen erfolgte eine zweistufige Clusteranalyse zur Bildung von sogenannten Strategie- und Vergleichstypen. Die Typisierung umfasst fünf Strategietypen, die sich wiederum in insgesamt 12 Vergleichstypen untergliedern. In diesem Beitrag wird lediglich der Strategietyp V noch weiter untergliedert analysiert (siehe Übersicht 1 sowie Grafik 1 und Tabelle 2). Dabei variiert die Zahl der umfassten Agenturbezirke je Strategietyp zwischen 6 und 49 Agenturbezirken. Es ist eine grundsätzliche Differenzierung einerseits zwischen ost- und westdeutschen Agenturbezirken und andererseits zwischen Stadt und Land festzustellen. Zudem bestehen in Westdeutschland Unterschiede entlang der Nord-Süd-Achse. ➤ Übersicht 1, Grafik 1

Die beiden städtischen Strategietypen in Westdeutschland unterscheiden sich hinsichtlich der Arbeitsmarktsituation. Während Strategietyp I „Großstädtische Agenturbezirke mit günstiger Arbeitsmarktlage“ insgesamt sechs Bezirke (10 Kreise, siehe Tabelle 2) umfasst, befinden sich in Strategietyp II „Großstädtisch geprägte Bezirke mit hoher Arbeitslosigkeit“ insgesamt 21 Agenturbezirke (36 Kreise). Diese liegen vor allem in Nord- und Westdeutschland und zu ihnen zählen auch die drei Stadtstaaten Berlin, Hamburg und Bremen.

<sup>8</sup> Afghanistan, Eritrea, Irak, Iran, Nigeria, Pakistan, Somalia und Syrien.

<sup>9</sup> Die Integrationsquote misst den Anteil der in Erwerbstätigkeit integrierten Personen an allen erwerbsfähigen Personen im Rechtskreis SGB III (Drittes Buch Sozialgesetzbuch) des jeweiligen Agenturbezirks (BA, 2020).

### Übersicht 1

#### Definition der Strategie- und Vergleichstypen der SGB-III-Typisierung 2018

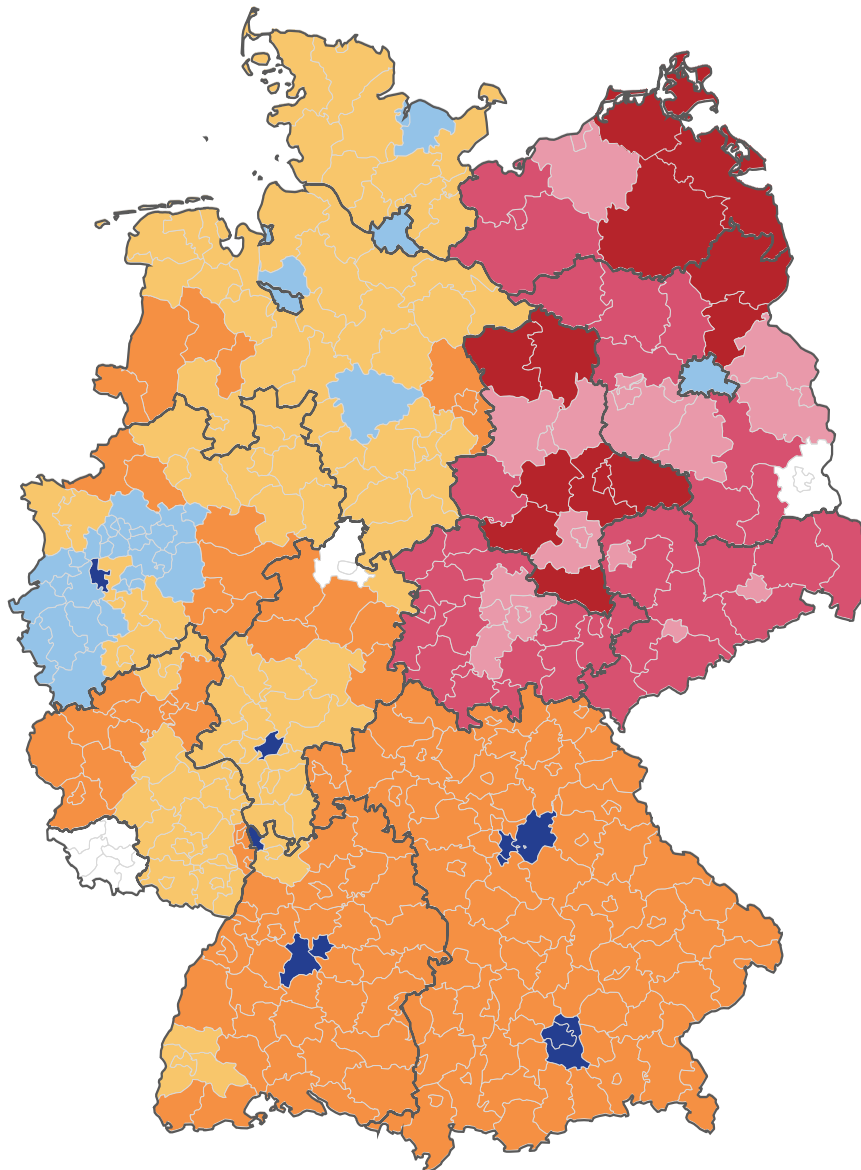
|                  |                                                                    |
|------------------|--------------------------------------------------------------------|
| Strategietyp I   | Großstädtische Bezirke mit günstiger Arbeitsmarktlage              |
| Strategietyp II  | Großstädtisch geprägte Bezirke mit hoher Arbeitslosigkeit          |
| Strategietyp III | Verdichtete und ländliche Bezirke mit mittlerer Arbeitslosigkeit   |
| Strategietyp IV  | Verdichtete und ländliche Bezirke mit niedriger Arbeitslosigkeit   |
| Vergleichstyp Va | Verdichtete und städtische Bezirke mit schlechter Arbeitsmarktlage |
| Vergleichstyp Vb | Gering verdichtete Bezirke mit schlechter Arbeitsmarktlage         |
| Vergleichstyp Vc | Ländliche Bezirke mit schlechter Arbeitsmarktlage                  |

Quelle: Blien/Hirschenauer (2017).

**Grafik 1**

**SGB-III-Typisierung 2018 der Agenturbezirke der Bundesagentur für Arbeit**

- Strategietyp I
- Strategietyp II
- Strategietyp III
- Strategietyp IV
- Vergleichstyp Va
- Vergleichstyp Vb
- Vergleichstyp Vc



Die Definition der Strategie- und Vergleichstypen der SGB-III-Typisierung 2018 enthält Übersicht 1. – Bei den sechs Kreisen im Saarland sowie den kreisfreien Städten Kassel und Cottbus und den umliegenden Landkreisen Kassel und Spree-Neiße ist keine klare Kreisuordnung der Wohnorte von Geflüchteten möglich.

Quelle: Blien/Hirschenauer (2017), eigene Darstellung mit Kreis- und Bundesländergrenzen.

© GeoBasis-DE / BKG 2018

2023 - 001

**Tabelle 2**

SGB-III-Typisierung der Agenturbezirke 2018

|                                                                        | Agenturbezirke | Kreise | Anteil der Arbeitslosen aus acht Asylherkunftsländern <sup>1</sup> |
|------------------------------------------------------------------------|----------------|--------|--------------------------------------------------------------------|
|                                                                        | Anzahl         |        | %                                                                  |
| Großstädtische Bezirke mit günstiger Arbeitsmarktlage (I)              | 6              | 10     | 6,4                                                                |
| Großstädtisch geprägte Bezirke mit hoher Arbeitslosigkeit (II)         | 21             | 36     | 5,8                                                                |
| Verdichtete und ländliche Bezirke mit mittlerer Arbeitslosigkeit (III) | 43             | 119    | 6,0                                                                |
| Verdichtete und ländliche Bezirke mit niedriger Arbeitslosigkeit (IV)  | 49             | 160    | 6,4                                                                |
| Bezirke mit schlechter Arbeitsmarktlage (V)                            | 35             | 76     | 3,2                                                                |
| Verdichtete und städtische Bezirke (Va)                                | 9              | 22     | 4,8                                                                |
| Gering verdichtete Bezirke (Vb)                                        | 17             | 41     | 3,9                                                                |
| Ländliche Bezirke (Vc)                                                 | 9              | 13     | 2,8                                                                |
| Insgesamt                                                              | 154            | 401    | 5,5                                                                |

<sup>1</sup> Afghanistan, Eritrea, Irak, Iran, Nigeria, Pakistan, Somalia und Syrien.  
Quelle: Blien/Hirschenauer (2017), Angaben beziehen sich auf das gesamte Bundesgebiet.

Bei den Strategietypen, die verdichtete und ländliche Agenturbezirke im Westen umfassen, ist ein Nord-Süd-Unterschied zu beobachten. Der Strategietyp III „Verdichtete und ländliche Bezirke mit mittlerer Arbeitslosigkeit“ umfasst 43 Agenturbezirke (119 Kreise), die zumeist außerhalb der süddeutschen Bundesländer Bayern und Baden-Württemberg liegen. Sie weisen einen in der Regel überdurchschnittlichen Erwerbspersonenanteil ohne Berufsabschluss auf. Die 49 Agenturbezirke (160 Kreise, siehe Tabelle 2) des Strategietyps IV „Verdichtete und ländliche Bezirke mit niedriger Arbeitslosigkeit“ sind gekennzeichnet von allgemein niedrigen Arbeitslosenquoten, der Anteil an Arbeitslosen aus den acht häufigsten Asylherkunftsländern liegt allerdings über dem bundesweiten Durchschnitt. [↗ Tabelle 2](#)

Die 35 Agenturbezirke (76 Kreise) im Osten sind sämtlich anhand des Strategietyps V „Bezirke mit schlechter Arbeitsmarktlage“ zusammengefasst. Sie weisen unterdurchschnittliche Arbeitsmarktindikatoren auf. Der Anteil an Arbeitslosen aus den Asylherkunftsländern liegt mit Ausnahme der Großstädte Halle, Magdeburg, Dresden und Leipzig in allen ostdeutschen Agenturbezirken unter den durchschnittlichen Werten.

## 3

### Räumliche Verteilung und Binnenmobilität differenziert nach der SGB-III-Typisierung 2018

#### 3.1 Räumliche Verteilung und Wanderungssalden

Die nach den fünf Strategietypen differenzierte räumliche Verteilung der in der Studie betrachteten Geflüchteten mit Schutzstatus weist im Vergleich zur Gesamtbevölkerung einige Unterschiede auf: Knapp ein Drittel (31 %) der zwischen 2015 und 2019 eingereisten Personen wohnte zum 31. Dezember 2020 in den Kreisen der großstädtisch geprägten Agenturbezirke mit hoher Arbeitslosigkeit (Strategietyp II). Der entsprechende Anteil der in Deutschland lebenden Gesamtbevölkerung (22 %) war um 9 Prozentpunkte geringer. Demgegenüber lebten im Vergleich zur Gesamtbevölkerung in den Kreisen der verdichteten und ländlichen Agenturbezirke mit niedriger Arbeitslosigkeit (Strategietyp IV) sowie den ostdeutschen Kreisen mit schlechter Arbeitsmarktlage (Strategietyp V) anteilig weniger Geflüchtete mit Schutzstatus (Strategietyp IV: 24 % gegenüber 30 %, Strategietyp V: 11 % gegenüber 15 %). Damit zeigt sich einerseits, dass diese Personengruppe in großstädtisch geprägten Agenturbezirken überrepräsentiert ist – vor allem in den Agenturbezirken mit hoher Arbeitslosigkeit. Andererseits lebten zum 31. Dezember 2020 im

Vergleich zur Gesamtbevölkerung weniger Geflüchtete mit Schutzstatus in den ostdeutschen Bundesländern. Inwieweit dies aus der anfänglich bundesweiten Verteilung von Geflüchteten resultiert oder ob vielmehr kreisübergreifende Wanderungen die Unterschiede bei der räumlichen Verteilung erklären können, macht ein Vergleich der Zu- und Fortzüge, das heißt der jeweilige Wanderungssaldo, sichtbar. [↗ Tabelle 3](#)

Während des Betrachtungszeitraums verzeichnen die Kreise der Agenturbezirke des Strategietyps II, die großstädtisch geprägt sind und eine hohe Arbeitslosigkeit aufweisen, mit 33 218 Personen den höchsten Netto-Wanderungsgewinn. In Relation zur Anzahl an Personen, die in diesen Kreisen ihren ersten Wohnsitz hatten, hat damit die Zahl der Geflüchteten mit Schutzstatus um 30 % zugenommen. Ebenfalls hohe Wanderungsgewinne von Geflüchteten mit Schutzstatus weisen die Kreise der Agenturbezirke des Strategietyps I auf, die auch großstädtisch geprägt sind, jedoch eine günstige Arbeitsmarktlage haben (absolut: + 4 087, relativ: + 19 %).

Demgegenüber registrieren die Kreise der ostdeutschen Agenturbezirke (Strategietyp V, ohne Berlin) mit schlechter Arbeitsmarktlage mit 23 994 Geflüchteten den höchsten Wanderungsverlust. Dies entspricht einem Rückgang von fast einem Drittel (– 31 %) an Geflüchteten mit Schutzstatus während des Betrachtungszeitraums. Kreise der verdichteten und ländlichen Agenturbezirke in Westdeutschland mit niedriger Arbeitslosigkeit (Strate-

giety IV, absolut: – 11 253, relativ: – 9 %) zeigen ebenfalls Wanderungsverluste, wenn auch in geringerem Umfang. In den Kreisen der verdichteten und ländlichen Agenturbezirke mit mittlerer Arbeitslosigkeit (Strategietyp III) hat sich die Anzahl an Geflüchteten mit Schutzstatus über den Betrachtungszeitraum hinweg nur leicht verändert (absolut: – 2 058; relativ: – 1 %).

### 3.2 Jährliche Wanderungsgewinne und -verluste

Die zum Teil sehr unterschiedlichen Wanderungssalden deuten darauf hin, dass neben der Anfangsverteilung vor allem das Wanderungsverhalten der Geflüchteten mit Schutzstatus die Diskrepanzen hinsichtlich der räumlichen Verteilung gegenüber der Gesamtbevölkerung erklärt. Aufgrund dessen rücken im Folgenden die Wanderungen selbst in den Fokus der Analyse. Bevor diese differenziert nach der SGB-III-Typisierung der Agenturbezirke analysiert werden, erfolgt eine zeitlich differenzierte Betrachtung nach den einzelnen Jahren. Von besonderem Interesse ist dabei die Frage, ob zeitliche Unterschiede bei der Wanderungshäufigkeit und -richtung beobachtet werden können.

Mit mehr als 100 000 Wanderungen fand mehr als ein Drittel aller hier betrachteten kreisübergreifenden Umzüge im Jahr 2016 statt. In diesem Jahr wurde – rückwirkend zum 1. Januar 2016 – die Wohnsitzregelung

**Tabelle 3**

**Räumliche Verteilung und Wanderungssaldo der Geflüchteten mit Schutzstatus sowie räumliche Verteilung der Gesamtbevölkerung nach SGB-III-Typisierung 2018**

|                                                                        | Geflüchtete mit Schutzstatus |     |                               |                | Bevölkerung <sup>1</sup> |     |
|------------------------------------------------------------------------|------------------------------|-----|-------------------------------|----------------|--------------------------|-----|
|                                                                        | Stichtag: 31.12.2020         |     | Wanderungssaldo 2015 bis 2020 |                | Stichtag: 31.12.2020     |     |
|                                                                        | Anzahl                       | %   | Anzahl                        | % <sup>2</sup> | Anzahl                   | %   |
| Großstädtische Bezirke mit günstiger Arbeitsmarktlage (I)              | 25 303                       | 5   | + 4 087                       | + 19           | 5 283 180                | 6   |
| Großstädtisch geprägte Bezirke mit hoher Arbeitslosigkeit (II)         | 144 548                      | 31  | + 33 218                      | + 30           | 17 582 109               | 22  |
| Verdichtete und ländliche Bezirke mit mittlerer Arbeitslosigkeit (III) | 135 911                      | 29  | – 2 058                       | – 1            | 21 923 241               | 27  |
| Verdichtete und ländliche Bezirke mit niedriger Arbeitslosigkeit (IV)  | 112 189                      | 24  | – 11 253                      | – 9            | 24 444 748               | 30  |
| Bezirke mit schlechter Arbeitsmarktlage (V)                            | 52 872                       | 11  | – 23 994                      | – 31           | 12 288 003               | 15  |
| Insgesamt <sup>1</sup>                                                 | 470 823                      | 100 | ± 0                           | ± 0            | 81 521 281               | 100 |

Quellen: Ausländerzentralregister (Stand: 31.12.2020), Blien/Hirschenauer (2017), Statistisches Bundesamt (2021), eigene Berechnungen

1 Die Angaben beziehen sich auf die 391 untersuchten Kreise zum 31.12.2020 und entsprechen damit nicht der Gesamtfläche und -bevölkerung in Deutschland (siehe auch Abschnitt 2.1).

2 Der Wanderungssaldo in % zeigt an, in welchem Umfang sich die Zahl der Geflüchteten mit Schutzstatus in den jeweiligen Regionen im Betrachtungszeitraum verändert hat. Die Bezugsbevölkerung entspricht dabei der Gesamtzahl an Personen, die ihren ersten Wohnort in Kreisen dieses Strategietyps hatten. Rechenbeispiel zu Strategietyp V: – 23 994 / [52 872 – (– 23 994)] = – 0,312; ein Wanderungsverlust von 23 994 Personen entspricht in den Kreisen der Agenturbezirke des Strategietyps V somit einem Rückgang um 31 %.

# Bedeutung raumstruktureller und arbeitsmarktrelevanter Faktoren bei innerdeutschen Wanderungen von Geflüchteten

**Tabelle 4**

Wanderungsverhalten der Geflüchteten mit Schutzstatus in den Jahren 2015 bis 2020

|                                                                             | Einheit | 2015   | 2016    | 2017   | 2018   | 2019   | 2020   |
|-----------------------------------------------------------------------------|---------|--------|---------|--------|--------|--------|--------|
| Wanderungen                                                                 | Anzahl  | 11 136 | 101 693 | 51 633 | 43 382 | 41 514 | 36 358 |
| Anteil an den kreis-<br>übergreifenden Umzügen<br>im Zeitraum 2015 bis 2020 | %       | 3,9    | 35,6    | 18,1   | 15,2   | 14,5   | 12,7   |

Quellen: Ausländerzentralregister (Stand: 31.12.2020), eigene Berechnungen

für anerkannte Geflüchtete eingeführt. Umzüge, die vom Jahresanfang bis zur Verabschiedung der gesetzlichen Grundlage im August 2016 erfolgten, mussten nur punktuell rückgängig gemacht werden. In den darauffolgenden Jahren 2017 bis 2020 bewegt sich die Zahl der kreisübergreifenden Umzüge zwischen 36 000 und 52 000. Diese Unterschiede können einerseits ein Resultat der hohen Zuwanderung von Geflüchteten in den Jahren 2015 und 2016 und der damit einhergehenden anfänglich hohen Mobilität sein. Andererseits könnten sich darin die Wirkungen der Wohnsitzregelung, die die Mobilität deutlich einschränkte, widerspiegeln.

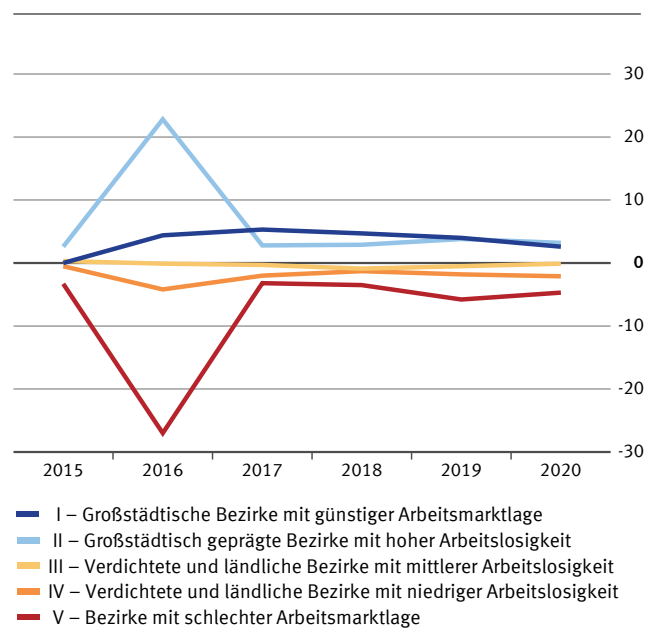
↪ Tabelle 4

Die Kreise der großstädtisch geprägten Agenturbezirke (Strategietypen I und II) weisen über den gesamten Betrachtungszeitraum positive Wanderungssalden auf. Dagegen sind in den ostdeutschen Agenturbezirken des Strategietyps V durchweg Wanderungsverluste festzustellen. Eine deutliche Diskrepanz der Wanderungssalden ist im Jahr 2016 zu beobachten. Demnach ist in den Agenturbezirken des Strategietyps V, die sich allesamt im Osten Deutschlands befinden (ohne Berlin) und eine schlechte Arbeitsmarktlage aufweisen, ein Wanderungsverlust von 27 % zu konstatieren. Im Vergleich dazu verzeichnen die Kreise der großstädtisch geprägten Agenturbezirke mit einer hohen Arbeitslosigkeit (Strategietyp II) hohe Wanderungsgewinne (23 %). In den Agenturbezirken der übrigen Strategietypen zeichnen sich über den gesamten Beobachtungszeitraum hingegen eher moderate jährliche Wanderungsgewinne beziehungsweise -verluste zwischen + 5 % und – 4 % ab.

↪ Grafik 2

**Grafik 2**

Jährliche Wanderungsgewinne und -verluste nach der SGB-III-Typisierung 2018  
in %



Quellen: Ausländerzentralregister (Stand: 31.12.2020), Blien/Hirschenauer (2017), eigene Berechnungen.

2023 - 002

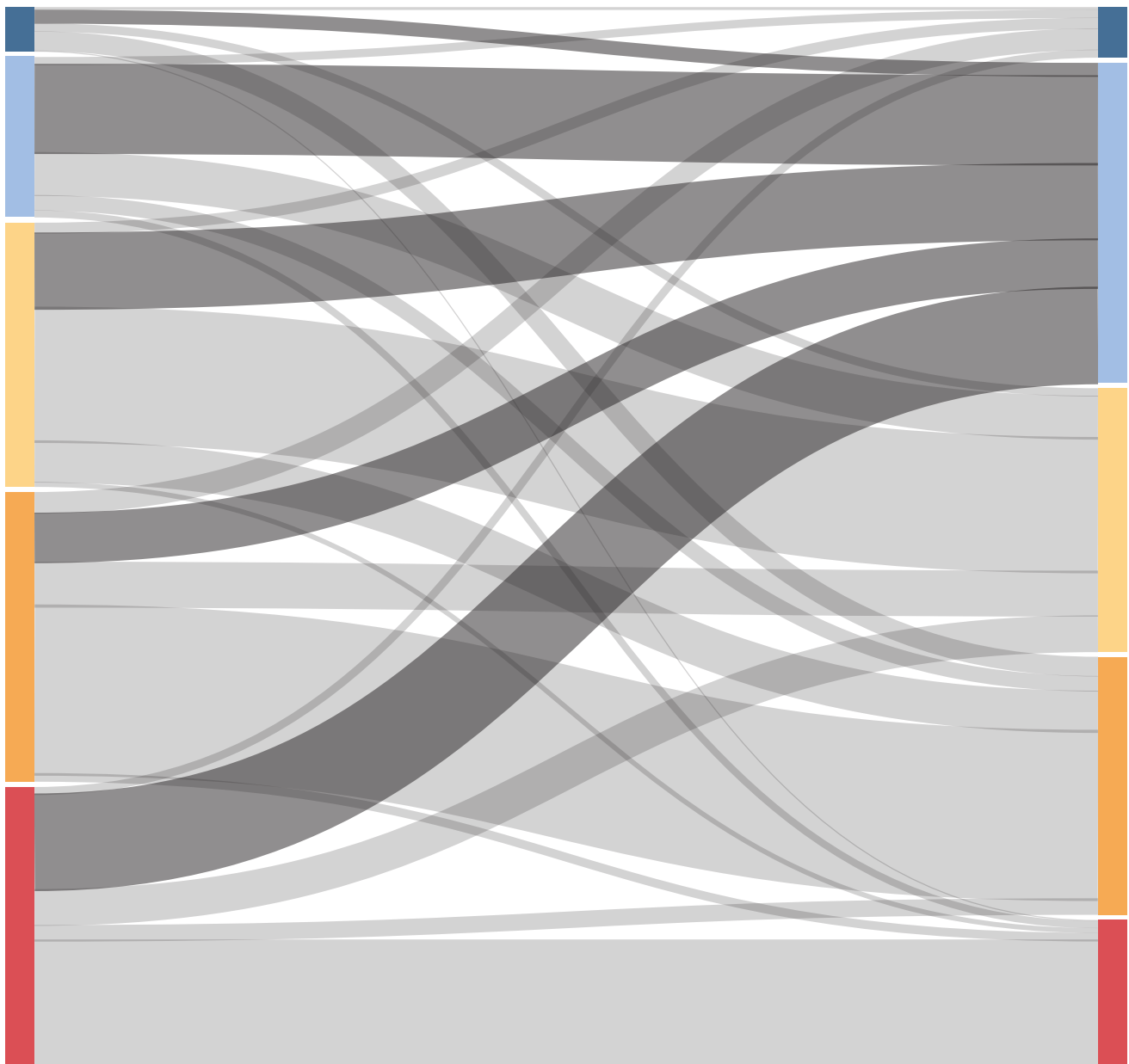
## 3.3 Wanderungsrichtung im Jahr 2016

Aufgrund der hohen Wanderungszahlen und sehr unterschiedlichen Wanderungssalden erfolgt im nächsten Schritt eine Auswertung der Wanderungen nur für das Jahr 2016 anhand der jeweiligen Start- und Zielregion. Die Wanderungsverläufe zwischen den Strategietypen sind dabei von links nach rechts zu lesen und entsprechen in ihrer Breite dem Anteil der jeweiligen Wanderung an allen registrierten Wanderungen, die Diskrepanz der Balken links und rechts hinsichtlich ihrer Höhe symbolisiert den Umfang der Zu- oder Abnahme der im jeweiligen Typ wohnhaften Geflüchteten mit Schutzstatus.

↪ Grafik 3

**Grafik 3**

Nach der SGB-III-Typisierung 2018 differenzierte Wanderungen im Jahr 2016



- I – Großstädtische Bezirke mit günstiger Arbeitsmarktlage
- II – Großstädtisch geprägte Bezirke mit hoher Arbeitslosigkeit
- III – Verdichtete und ländliche Bezirke mit mittlerer Arbeitslosigkeit
- IV – Verdichtete und ländliche Bezirke mit niedriger Arbeitslosigkeit
- V – Bezirke mit schlechter Arbeitsmarktlage

Die zugrunde liegenden Werte enthält Tabelle 5. Hervorgehoben sind die Zuzüge in Kreise der großstädtisch geprägten Agenturbezirke mit hoher Arbeitslosigkeit des Strategietyps 2.  
Quelle: Ausländerzentralregister (Stand: 31.12.2020), Blien/Hirschenauer (2017), eigene Berechnungen.

2023 - 003

## Bedeutung raumstruktureller und arbeitsmarktrelevanter Faktoren bei innerdeutschen Wanderungen von Geflüchteten

Die Wanderungsverläufe verdeutlichen einerseits, dass die meisten Wanderungen innerhalb derselben Strategietypen stattfinden; in den Strategietypen II, III und IV gilt dies für mindestens jeden zweiten Fortzug. Dies kann ein Indiz dafür sein, dass viele Umzüge als Ziel benachbarte Kreise haben und damit innerhalb desselben, mehrere Kreise umfassenden Agenturbezirks bleiben. Ausnahme bilden hierbei Fortzüge aus großstädtisch geprägten Agenturbezirken mit guter Arbeitsmarktlage, die als Ziel häufig umliegende Kreise der Großstädte haben, die den Agenturbezirken des Strategietyps IV zugeordnet werden können.

Andererseits stellen Fortzüge in den Strategietyp II, das heißt in großstädtisch geprägte Agenturbezirke mit einer hohen Arbeitslosigkeit, in allen anderen Strategietypen den jeweils zweithäufigsten Anteil an kreisübergreifenden Umzügen dar. Dies lässt auf eine hohe Attraktivität dieser Regionen für Geflüchtete unabhängig ihres vorherigen Wohnortes schließen. Besonders attraktiv scheinen die städtisch geprägten Kreise des Strategietyps II dabei für Geflüchtete mit Schutzstatus zu sein, die in den ostdeutschen Agenturbezirken mit schlechter Arbeitsmarktlage leben. Jede dritte Wanderung aus einem der ostdeutschen Kreise (ohne Berlin) im Jahr 2016 hatte als Ziel einen Kreis der großstädtisch geprägten Agenturbezirke mit hoher Arbeitslosigkeit (Strategietyp II). Dabei ist festzuhalten, dass diese Wanderungen angesichts der vergleichsweise schlechten Arbeitsmarktlage in den Agenturbezirken der Strategietypen II und V wohl zu

keiner nennenswerten Verbesserung der Arbeitsmarktchancen für die Geflüchteten geführt haben. Dennoch trägt dieses Wanderungsmuster in Kombination mit vergleichsweise geringen Zuzugszahlen maßgeblich zu dem hohen Wanderungsverlust der ostdeutschen Agenturbezirke im Jahr 2016 bei. ➤ **Tabelle 5**

In einem letzten Schritt wird näher untersucht, ob es hinsichtlich des Zu- und Fortzugsverhaltens im Osten raumstrukturelle Stadt-Land-Differenzen gibt. Dafür werden alle Wanderungen im Jahr 2016, die einen der ostdeutschen Kreise entweder als Start- oder Zielregion aufweisen, anhand der weiter ausdifferenzierten Vergleichstypen Va, Vb und Vc hinsichtlich ihrer raumstrukturellen Merkmale gesondert betrachtet und analysiert. ➤ **Grafik 4**

Dabei zeigen sich Unterschiede zwischen den verdichteten und städtischen Bezirken (Vergleichstyp Va) und den gering verdichteten und ländlichen Bezirken (Vergleichstypen Vb und Vc). Während in den Letzteren erhebliche Wanderungsverluste zu beobachten sind, ist der Wanderungssaldo in den verdichteten und städtischen Bezirken in den ostdeutschen Bundesländern ausgeglichen. Dies bedeutet, dass die Fortzüge aus den gering verdichteten und ländlichen Kreisen zu einem großen Teil den Wanderungsverlust im Osten dominieren. Ein besonders großer Anteil der Fortzüge (70 %) aus den ländlichen Bezirken (Vergleichstyp Vc, siehe dazu die markierten Wanderungsverläufe in Grafik 4) hatte dabei

**Tabelle 5**

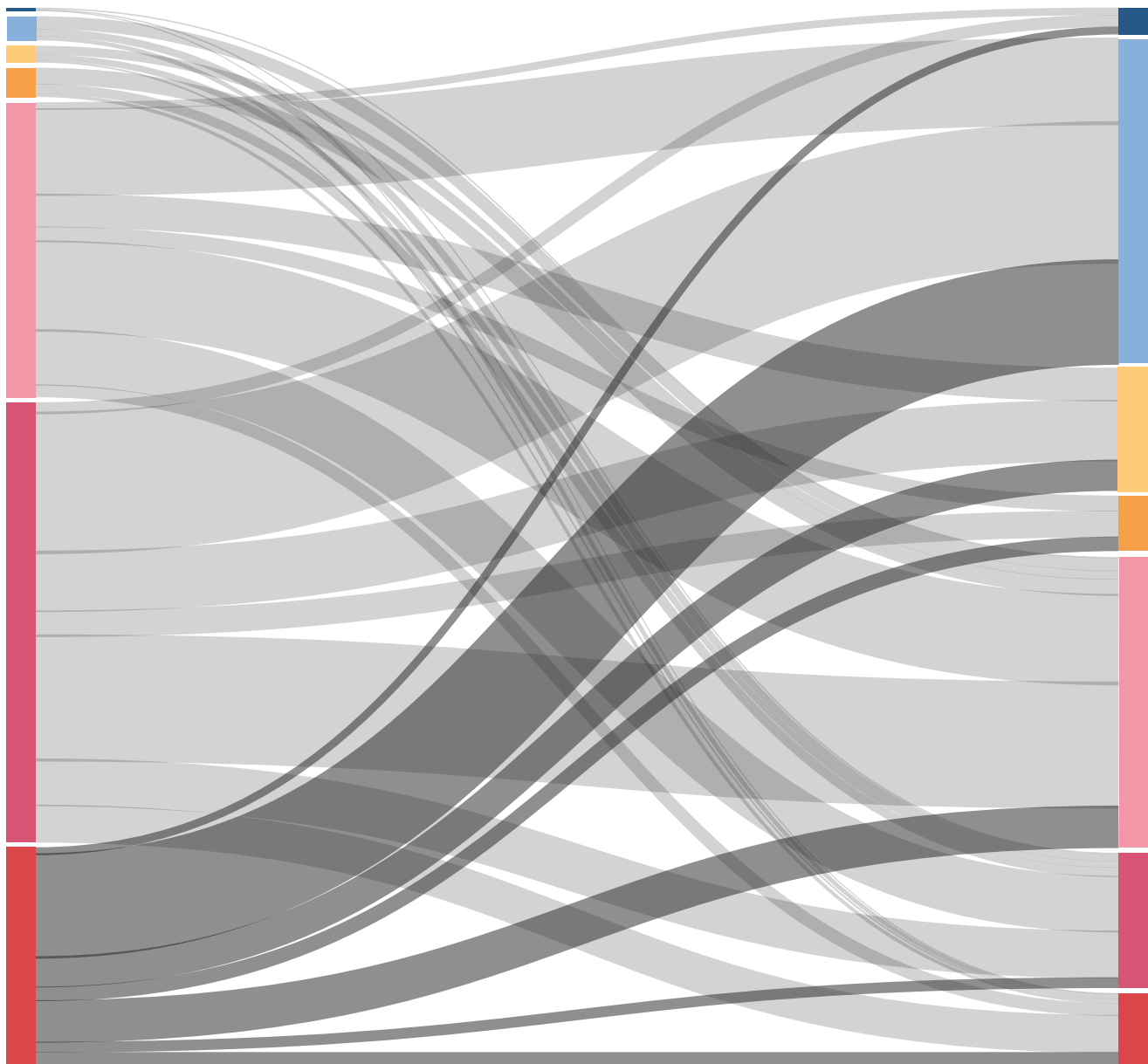
Nach der SGB-III-Typisierung 2018 differenzierte Wanderungsrichtungen 2016

| Ziel der Wanderung \ Start der Wanderung                     | Strategietyp I | Strategietyp II | Strategietyp III | Strategietyp IV | Strategietyp V | Fortzüge insgesamt |
|--------------------------------------------------------------|----------------|-----------------|------------------|-----------------|----------------|--------------------|
| Strategietyp I                                               | 275            | 1 330           | 791              | 1 876           | 95             | 4 367              |
| Strategietyp II                                              | 787            | 8 584           | 4 177            | 1 483           | 686            | 15 717             |
| Strategietyp III                                             | 1 056          | 7 358           | 12 983           | 4 023           | 478            | 25 898             |
| Strategietyp VI                                              | 2 066          | 4 807           | 4 374            | 16 304          | 828            | 28 379             |
| Strategietyp V                                               | 765            | 9 291           | 3 499            | 1 578           | 12 199         | 27 332             |
| Zuzüge insgesamt                                             | 4 949          | 31 370          | 25 824           | 25 264          | 14 286         | 101 693            |
| Anteil an allen Fortzügen aus der jeweiligen Zielregion in % |                |                 |                  |                 |                |                    |
| Strategietyp I                                               | 6              | 30              | 18               | 43              | 2              | 100                |
| Strategietyp II                                              | 5              | 55              | 27               | 9               | 4              | 100                |
| Strategietyp III                                             | 4              | 28              | 50               | 16              | 2              | 100                |
| Strategietyp VI                                              | 7              | 17              | 15               | 57              | 3              | 100                |
| Strategietyp V                                               | 3              | 34              | 13               | 6               | 45             | 100                |
| Zuzüge insgesamt                                             | 5              | 31              | 25               | 25              | 14             | 100                |

Quellen: Ausländerzentralregister (Stand: 31.12.2020), Blien/Hirschenauer (2017), eigene Berechnungen.

Grafik 4

Fort- und Zuzüge aus bzw. in ostdeutsche Kreise nach der SGB-III-Typisierung 2018 im Jahr 2016



- I – Großstädtische Bezirke mit günstiger Arbeitsmarktlage
- II – Großstädtisch geprägte Bezirke mit hoher Arbeitslosigkeit
- III – Verdichtete und ländliche Bezirke mit mittlerer Arbeitslosigkeit
- IV – Verdichtete und ländliche Bezirke mit niedriger Arbeitslosigkeit

- Va – Verdichtete und städtische Bezirke mit hoher Arbeitslosigkeit
- Vb – Gering verdichtete Bezirke mit hoher Arbeitslosigkeit
- Vc – Ländliche Bezirke mit schlechter Arbeitsmarktlage

Die zugrunde liegenden Werte enthält Tabelle 6. Hervorgehoben sind die Fortzüge aus Kreisen der ländlichen Agenturbezirke mit schlechter Arbeitsmarktlage in Ostdeutschland.  
Quelle: Ausländerzentralregister (Stand: 31.12.2020), Blien/Hirschenauer (2017), eigene Berechnungen.

2023 - 004

# Bedeutung raumstruktureller und arbeitsmarktrelevanter Faktoren bei innerdeutschen Wanderungen von Geflüchteten

**Tabelle 6**

Nach der SGB-III-Typisierung 2018 differenzierte Fort- und Zuzüge aus beziehungsweise in ostdeutsche(n) Kreise(n) 2016

| Start der Wanderung \ Ziel der Wanderung                     | Strategietyp |       |       |       | Vergleichstyp |       |       | Fortzüge insgesamt |
|--------------------------------------------------------------|--------------|-------|-------|-------|---------------|-------|-------|--------------------|
|                                                              | I            | II    | III   | IV    | Va            | Vb    | Vc    |                    |
| Strategietyp I                                               |              |       |       |       | 41            | 37    | 17    | 95                 |
| Strategietyp II                                              |              |       |       |       | 355           | 209   | 122   | 686                |
| Strategietyp III                                             |              |       |       |       | 250           | 178   | 50    | 478                |
| Strategietyp IV                                              |              |       |       |       | 462           | 268   | 98    | 828                |
| Vergleichstyp Va                                             | 212          | 2 424 | 932   | 437   | 2 530         | 1 564 | 356   | 8 455              |
| Vergleichstyp Vb                                             | 334          | 3 948 | 1 701 | 739   | 3 519         | 1 322 | 1 050 | 12 613             |
| Vergleichstyp Vc                                             | 219          | 2 919 | 866   | 402   | 1 176         | 298   | 384   | 6 264              |
| Zuzüge insgesamt                                             | 765          | 9 291 | 3 499 | 1 578 | 8 333         | 3 876 | 2 077 | 29 419             |
| Anteil an allen Fortzügen aus der jeweiligen Zielregion in % |              |       |       |       |               |       |       |                    |
| Strategietyp I                                               |              |       |       |       | 43            | 39    | 18    | 100                |
| Strategietyp II                                              |              |       |       |       | 52            | 30    | 18    | 100                |
| Strategietyp III                                             |              |       |       |       | 52            | 37    | 10    | 100                |
| Strategietyp IV                                              |              |       |       |       | 56            | 32    | 12    | 100                |
| Vergleichstyp Va                                             | 3            | 29    | 11    | 5     | 30            | 18    | 4     | 100                |
| Vergleichstyp Vb                                             | 3            | 31    | 13    | 6     | 28            | 10    | 8     | 100                |
| Vergleichstyp Vc                                             | 3            | 47    | 14    | 6     | 19            | 5     | 6     | 100                |
| Zuzüge insgesamt                                             | 3            | 32    | 12    | 5     | 28            | 13    | 7     | 100                |

Quellen: Ausländerzentralregister (Stand: 31.12.2020), Blien/Hirschenauer (2017), eigene Berechnungen.

im Jahr 2016 westdeutsche Bundesländer, darunter vor allem Großstädte mit hoher Arbeitslosigkeit (47 %), zum Ziel. Aus den gering verdichteten Bezirken (Vergleichstyp Vb) zogen Geflüchtete mit Schutzstatus vor allem in Großstädte mit hoher Arbeitslosigkeit (31 %) im Westen sowie in verdichtete und städtische Bezirke (28 %) im Osten Deutschlands. ➔ Tabelle 6

## 4

### Fazit

Ziel des vorliegenden Beitrags war es, den Zusammenhang zwischen raumstrukturellen und arbeitsmarktrelevanten Merkmalen und der räumlichen Verteilung sowie dem Wanderungsverhalten dieser Gruppe detailliert zu beschreiben. Dafür wurde eine Typisierung von Arbeitsagenturbezirken genutzt, die die beiden Dimensionen, das heißt Stadt-Land-Unterschiede und die regionale Arbeitsmarktlage, miteinander verknüpft (Blien/Hirschenauer, 2017). Die Auswertungen beruhen außerdem auf den Vorarbeiten von Brückner (2019), der erst-

mals AZR-Daten<sup>10</sup> als Längsschnittdatensatz nutzte und damit Wanderungsbiografien von ausländischen Staatsangehörigen systematisch untersuchte.

Die Analysen zeigen, dass Geflüchtete mit Schutzstatus im Vergleich zur deutschen Bevölkerung zum 31. Dezember 2020 überproportional oft in großstädtisch geprägten Bezirken mit einer hohen Arbeitslosigkeit lebten. Zudem lag der Anteil der Geflüchteten in den ostdeutschen Kreisen unter dem entsprechenden Anteil in der Gesamtbevölkerung (11 % gegenüber 15 %). Diese Unterschiede sind weniger das Resultat einer ungleichen Anfangsverteilung, sondern vielmehr auf individuelle Wanderungsentscheidungen zurückzuführen. So verzeichnen die großstädtischen Kreise mit hoher Arbeitslosigkeit zwischen 2015 und 2020 einen Wanderungsgewinn von insgesamt rund 33 000 Geflüchteten mit Schutzstatus, was einer Steigerung um etwa 30 % entspricht. Im Gegensatz

<sup>10</sup> Seit August 2021 können Forschungseinrichtungen Daten aus dem AZR für Forschungszwecke über das Forschungsdatenzentrum des Bundesamtes für Migration und Flüchtlinge (BAMF-FDZ) erhalten. Das Datenangebot umfasst den „AZR-Forschungsdatensatz 2021“, eine 20-prozentige Zufallsstichprobe aus dem Gesamtdatenbestand des AZR, personenbezogene Adressdaten einschließlich erforderlicher Strukturmerkmale für die Durchführung von wissenschaftlichen Befragungen sowie anonymisierte Auswertungen aus dem Gesamtdatenbestand. Weitere Informationen unter: [www.bamf.de/fdz](https://www.bamf.de/fdz)

dazu beläuft sich der Netto-Wanderungsverlust in den ostdeutschen Bundesländern (ohne Berlin) im gleichen Zeitraum auf knapp 24 000 Geflüchtete, das bedeutet einen Rückgang um 31 %.

Zur Erklärung dieser sehr unterschiedlichen Wanderungssalden kommt dem Wanderungsgeschehen im Jahr 2016 große Bedeutung zu, in dem zum einen die mit Abstand meisten kreisübergreifenden Umzüge erfolgten. Zum anderen unterscheiden sich die Wanderungssalden der einzelnen Strategietypen im Jahr 2016 stark, während in den übrigen Jahren die jährlichen Wanderungsgewinne und -verluste sich in einem moderaten Bereich bewegen. Offensichtlich war 2016 – das Jahr nach der sehr starken Zuwanderung von Geflüchteten nach Deutschland im Jahr 2015, in dem aber erst im zweiten Halbjahr mit der Wohnsitzregelung nach § 12a Aufenthaltsgesetz eine Mobilitätsbeschränkung für Geflüchtete mit Schutzstatus eingeführt wurde – das Jahr des „großen Kommens und Gehens“ innerhalb Deutschlands.

Bei einer Betrachtung der etwa 100 000 Wanderungen im Jahr 2016 wird deutlich, dass die Wanderungsgewinne in den großstädtisch geprägten Kreisen mit hoher Arbeitslosigkeit vor allem aus Zuzügen aus einem der ostdeutschen Kreise resultieren. Diese vergleichsweise sehr hohe Anzahl an Geflüchteten mit Schutzstatus, die aus einem der ostdeutschen Kreise in einen großstädtischen, jedoch von hoher Arbeitslosigkeit gekennzeichneten (westdeutschen) Kreis sowie Berlin ziehen, ist auch maßgeblich für die im Jahr 2016 festgestellten hohen Wanderungsverluste der ostdeutschen Kreise. Eine regionale Betrachtung von Wanderungen in und aus ostdeutschen Kreisen im Jahr 2016 verdeutlicht darüber hinaus, dass der dortige Wanderungsverlust vor allem aus einer starken Abwanderung aus gering verdichteten sowie ländlichen Kreisen im Osten resultiert. Die städtisch geprägten ostdeutschen Kreise weisen dagegen nur marginale Wanderungsverluste auf.

Eine weitere Erklärung für die Wanderungsgewinne der großstädtisch geprägten Kreise mit hoher Arbeitslosigkeit könnte sein, dass zu Beginn des Aufenthalts in Deutschland für die Gruppe der Geflüchteten mit Schutzstatus die regionale Arbeitsmarktlage bei der Wohnortwahl eher nachrangig ist. Andere Faktoren, wie ein stärkerer Ausbau der lokalen migrationsspezifischen Infrastrukturen oder ethnische Netzwerke, die vor allem

in (Groß-)Städten zu finden sind, spielen eine bedeutendere Rolle als ökonomische Faktoren. Über die Bedeutung von ethnischen Netzwerken bei innerdeutschen Wanderungsentscheidungen von Geflüchteten liegen bislang nur wenige empirische Erkenntnisse vor, daher ist weitere Forschung zu diesem Bereich erforderlich. Beispielsweise könnten mit den im AZR verfügbaren Informationen über die Staatsangehörigkeit beziehungsweise das Herkunftsland räumliche Cluster bestimmter herkunftsspezifischer Gruppen identifiziert werden, die eine Analyse über mögliche Zusammenhänge mit dem Wanderungsverhalten von Geflüchteten zulassen. Zwar wäre eine Einschränkung bei diesem Vorgehen, dass deutsche Staatsangehörige mit Migrationshintergrund in den Analysen nicht mitberücksichtigt werden, da diese nicht im AZR erfasst werden. Dennoch wäre es ein zu prüfender Forschungsansatz, um neben der Bedeutung raumstruktureller und arbeitsmarktrelevanter Faktoren mehr über die Rolle von ethnischen Netzwerken bei innerdeutschen Wanderungsentscheidungen von Geflüchteten zu erfahren. 

## LITERATURVERZEICHNIS

---

Bundesagentur für Arbeit. *Glossar zum Bericht zur Rahmenzielvereinbarung SGB III*. 2022. [Zugriff am 10. Januar 2023]. Verfügbar unter: [www.bmas.de](http://www.bmas.de)

Blien, Uwe/Hirschenauer, Franziska. *Vergleichstypen 2018: Aktualisierung der SGB-III-Typisierung*. Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg. IAB-Forschungsbericht 11/2017.

Brückner, Gunter. *Das Ausländerzentralregister als Längsschnittdatenquelle*. In: WISTA Wirtschaft und Statistik. Ausgabe 1/2019, Seite 35 ff.

Rösch, Tabea/Schneider, Hanne/Weber, Johannes/Worbs, Susanne. *Integration von Geflüchteten in ländlichen Räumen*. Forschungsbericht 36 des Forschungszentrums Migration, Integration und Asyl im Bundesamt für Migration und Flüchtlinge. Nürnberg 2020.

Statistisches Bundesamt. *Bevölkerung und Erwerbstätigkeit. Schutzsuchende. Ergebnisse des Ausländerzentralregisters 2019*. Fachserie 1, Reihe 2.4.

Statistisches Bundesamt. *Daten aus dem Gemeindeverzeichnis. Kreisfreie Städte und Landkreise nach Fläche, Bevölkerung und Bevölkerungsdichte, Gebietsstand: 31.12.2020*.

Tanis, Kerstin. *Die Wohnhistorie Geflüchteter in Deutschland*. Ausgabe 01/2022 der Kurzanalysen des Forschungszentrums Migration, Integration und Asyl des Bundesamtes für Migration und Flüchtlinge. Nürnberg 2022.

Weber, Johannes. *Binnenmobilität von Geflüchteten in Deutschland. Eine explorative Analyse auf Basis des Ausländerzentralregisters*. Forschungsbericht 39 des Forschungszentrums Migration, Integration und Asyl des Bundesamtes für Migration und Flüchtlinge. Nürnberg 2022.

## RECHTSGRUNDLAGEN

---

Asylgesetz in der Fassung der Bekanntmachung vom 2. September 2008 (BGBl. I Seite 1798), das zuletzt durch Artikel 1 des Gesetzes vom 21. Dezember 2022 (BGBl. I Seite 2817) geändert worden ist.

Gesetz über den Aufenthalt, die Erwerbstätigkeit und die Integration von Ausländern im Bundesgebiet (Aufenthaltsgesetz – AufenthG) in der Fassung der Bekanntmachung vom 25. Februar 2008 (BGBl. I Seite 162), das zuletzt durch Artikel 5 des Gesetzes vom 21. Dezember 2022 (BGBl. I Seite 2847) geändert worden ist.

Grundgesetz für die Bundesrepublik Deutschland in der im Bundesgesetzblatt Teil III, Gliederungsnummer 100-1, veröffentlichten bereinigten Fassung, das zuletzt durch Artikel 1 des Gesetzes vom 19. Dezember 2022 (BGBl. I Seite 2478) geändert worden ist.

Verordnung zur Durchführung des Gesetzes über das Ausländerzentralregister (AZRG-Durchführungsverordnung – AZRG-DV) vom 17. Mai 1995 (BGBl. I Seite 695), die zuletzt durch Artikel 5 des Gesetzes vom 19. Dezember 2022 (BGBl. I Seite 2632) geändert worden ist.

# TRUSTED SMART SURVEYS: FRAGEBOGEN TRIFFT NEUE DIGITALE DATEN

Shari Stehrenberg, Johannes Volk

➤ **Schlüsselwörter:** Trusted Smart Surveys – Trusted Smart Statistics – neue digitale Daten – Erhebungsmethoden – Sensordaten – ESSnet

## ZUSAMMENFASSUNG

„Smart Surveys“ sind Befragungen, die mithilfe von mobilen Geräten, typischerweise Smartphones oder Tablets, durchgeführt werden. Sie kombinieren interaktive Daten, die explizit von den Auskunftgebenden zur Verfügung gestellt werden, mit passiven, im Hintergrund von Sensoren erfassten Daten. Dadurch bieten Smart Surveys das Potenzial, die Qualität von Befragungen zu verbessern und den Befragungsaufwand bei zusätzlichem Informationsgewinn zu verringern. Das europäische Projekt ESSnet Smart Surveys leistete vorbereitende Arbeiten für die Entwicklung einer europaweiten Plattform zum Austausch von Smart-Survey-Lösungen und -komponenten. Pilotprojekte in den Bereichen Konsum, Zeitverwendung, Gesundheit und Lebensbedingungen ergänzten das ESSnet-Smart-Surveys-Projekt.

➤ **Keywords:** Trusted Smart Surveys – Trusted Smart Statistics – New Digital Data – survey methodology – sensor data – ESSnet

## ABSTRACT

*Smart Surveys are surveys conducted with the help of mobile devices, typically smartphones or tablets. They combine interactive data explicitly provided by respondents with passive data collected in the background by sensors. In this way, Smart Surveys offer the potential to improve the quality of surveys and reduce the response burden while collecting additional information. The European project ESSnet Smart Surveys carried out preparatory work for the development of a Europe-wide platform for the exchange of Smart Survey solutions and components. In addition, pilot projects were conducted in the areas of consumption, time use, health and living conditions.*



**Shari Stehrenberg**

hat einen Master of Science Wirtschaftsprsychologie und ist Referentin im Referat „Europäische und internationale Koordination“ des Statistischen Bundesamtes. Zuvor war sie als Referentin im Institut für Forschung und Entwicklung in der Bundesstatistik tätig und in dieser Position für die Gesamtprojektleitung des ESSnet-Smart-Surveys-Projekts verantwortlich.



**Johannes Volk**

studierte Politologie und Statistik und ist Referent im Referat „Design und Weiterentwicklung von Erhebungsinstrumenten“ des Statistischen Bundesamtes. Einer seiner Arbeitsschwerpunkte in den letzten Jahren war das ESSnet-Smart-Surveys-Projekt.

## 1

### Einleitung

Die rasante und transformative Entwicklung des Internets der Dinge und sensorgesteuerter smarterer Geräte sowie das Vordringen digitaler Technologien in den Alltag haben zu einer regelrechten Datenflut geführt. Das Volumen, die Menge und die Dimensionalität dieser Daten sind um ein Vielfaches höher als bei Datenerhebungen im vordigitalen Zeitalter (Ricciato und andere, 2019).

Dadurch bieten neue digitale Daten die Chance, sich als dritte Datenquelle neben Umfragedaten und Verwaltungsdaten zu etablieren.<sup>1</sup> Zudem eröffnen sie die Möglichkeit, die umfragebasierte Datenerhebung durch den Einsatz von Trusted Smart Surveys maßgeblich zu ergänzen und zu verbessern.

Dieser Beitrag geht einleitend allgemein auf Trusted Smart Surveys ein und stellt danach das ESSnet<sup>2</sup>-Smart-Surveys-Projekt vor, bei dem das Statistische Bundesamt neben der inhaltlichen Beteiligung auch mit der Gesamtkoordination betraut war.

## 2

### Chancen und Herausforderungen von Trusted Smart Surveys

Der Begriff „Smart Surveys“ bezeichnet Befragungen, die mithilfe von mobilen Geräten, typischerweise mit Smartphones oder Tablets, durchgeführt werden. Sie kombinieren explizit von den Auskunftgebenden zur Verfügung gestellte interaktive Daten (beispielsweise Antworten auf Fragen) mit passiven Daten, die im Hintergrund von Sensoren (zum Beispiel von Beschleunigungsmessern, Temperatursensoren, Mikrofonen) erfasst werden (Ruppert und andere, 2018).

1 Mehrere Forschungsaktivitäten und Pilotprojekte im Europäischen Statistischen System (ESS) haben bereits das Potenzial der Nutzung neuer digitaler Datenquellen (dazu zählen Mobilitätsdaten von Mobilfunknetzbetreibern, Satellitenbilder oder durch Webscraping automatisiert gewonnene, frei verfügbare Daten aus dem Internet) für die amtliche Statistik aufgezeigt. Als einschlägige Beispiele sind die Projekte ESSnet Big Data I und II zu nennen. Siehe [ec.europa.eu](https://ec.europa.eu) und [ec.europa.eu](https://ec.europa.eu)

2 ESSnet = European Statistical System Network.

Darüber hinaus werden bei „Trusted Smart Surveys“ technische Lösungen zum Schutz der Privatsphäre (Privacy Enhancing Technologies) eingesetzt. Sie sollen personenbezogene Daten besser und automatisiert schützen und eine strenge Kontrolle über die Datenverarbeitung und die Extraktion von Informationen aus den Daten etablieren, sodass vollständige Transparenz und Nachprüfbarkeit der Datenverarbeitungsmethoden und -algorithmen gewährleistet sind (Ricciato und andere, 2020). Der Einsatz solcher Technologien ermöglicht es zudem, Daten zu analysieren, ohne dass diese im Auswertungssystem vollständig vorliegen müssen. Diese Maßnahme soll nicht nur die rechtlichen Anforderungen an den Datenschutz erfüllen, sondern auch die Akzeptanz und Teilnahmebereitschaft bei den Auskunftgebenden fördern. Dies ist besonders wichtig, da bei smarten Erhebungen viel mehr Daten erhoben werden als bei herkömmlichen Erhebungen, diese Daten weniger spezifisch für die endgültige Statistik sind und viele zusätzliche Informationen über die Befragten enthalten. Daraus resultiert ein größeres Risiko für die Privatsphäre der Auskunftgebenden als bei herkömmlichen Umfragen, was die Notwendigkeit sicherer Systeme unterstreicht.

Smart Surveys bieten das Potenzial, sehr viele, sehr präzise Daten verarbeiten zu können und somit eine bessere Datenqualität zu erreichen als mit herkömmlichen Umfragen. Diese können durch negative Effekte wie schlechtes Erinnerungsvermögen, schlechtes Schätzvermögen oder soziale Erwünschtheit verzerrt sein. Allerdings ist zu bedenken, dass verschiedene Fehlerquellen (wie mangelnde Repräsentativität, technische Probleme, ausgeschaltete Geräte oder geringe Qualität der Sensormessungen) die Qualität von Sensordaten ebenfalls negativ beeinflussen können (Struminskaya und andere, 2020).

Neben den erwähnten Risiken liegen die Chancen von Smart Surveys vor allem in der Möglichkeit, passiv und aktiv erhobene Daten miteinander zu kombinieren. Passiv gewonnene Sensordaten sind zwar umfangreich und können beispielsweise das Reiseverhalten (Standorte, Dauer, Geschwindigkeit, Verkehrsmittel) sehr genau erfassen, sie ermöglichen aber beispielsweise keinen Einblick in die Motivation, die zu einer Reise führte. Die Kombination der Sensordaten mit Umfragedaten, zum Beispiel zum Grund der Reise, bietet den großen Mehrwert, sehr genaue Messdaten mit zusätzlichen, qualitativen Informationen analysieren zu können (Struminskaya und andere, 2020).

Zudem können Smart Surveys die Belastung der Befragten reduzieren: Da entsprechende Informationen passiv über Sensoren gewonnen werden, müssen weniger Fragen gestellt werden (Keusch und andere, 2019). Dies gilt insbesondere, wenn die Datenabfrage sehr komplex und/oder mit großem Zeitaufwand verbunden ist.<sup>13</sup> Einen weiteren positiven Effekt für die Befragten könnte die Möglichkeit einer kontinuierlichen Interaktion mit ihnen darstellen, zum Beispiel indem sie individualisierte Berichte und Auswertungen als Feedback erhalten (Ricciato und andere, 2020).

### 3

## Das ESSnet-Smart-Surveys-Projekt

### 3.1 Ziel und Aufbau

Mit dem von der Europäischen Kommission geförderten Projekt „ESSnet Smart Surveys“ hat ein Konsortium von zwölf nationalen statistischen Ämtern die Einsatzmöglichkeiten von Trusted Smart Surveys in der amtlichen Statistik untersucht (Stehrenberg/Giannakouris, 2022). Das Konzept von Trusted Smart Surveys ist dabei ein Teil der europäischen Trusted-Smart-Statistics-Initiative. Hierbei wird Trusted Smart Statistics als ein von intelligenten Systemen bereitgestellter Dienst verstanden, der prüfbare und transparente Datenlebenszyklen umfasst, die Gültigkeit und Genauigkeit der Ergebnisse gewährleistet und die Privatsphäre der betroffenen Personen respektiert sowie einen hohen Datenschutz gewährleistet (Ricciato und andere, 2020).<sup>14</sup>

Das Projekt „ESSnet Smart Surveys“ leistete in diesem Kontext vorbereitende Arbeiten, um eine europaweite Plattform zu entwickeln. Eine solche Plattform dient zum Austausch und zur Wiederverwendung von Trusted-Smart-Survey-Lösungen und -Komponenten, die sowohl unabhängig von bestimmten Statistikbereichen als auch

flexibel und modular sein sollen. Die Vorentwicklungsphase umfasste folgende Schritte: Anforderungen sammeln, Designprinzipien identifizieren, einen konzeptionellen Rahmen und eine Referenzarchitektur entwickeln, Spezifikationen und Proof of Concepts formulieren (Stehrenberg/Giannakouris, 2022).

Das Projekt gliederte sich in drei Arbeitspakete:

- › Kommunikation und Koordinierung (Arbeitspaket 1; Leitung: Statistisches Bundesamt)
- › Pilotprojekte (Arbeitspaket 2; Leitung: CBS [Centraal Bureau voor de Statistiek – Statistikamt der Niederlande])
- › Konzeptioneller Rahmen für eine Plattform für Trusted-Smart-Survey-Komponenten (Arbeitspaket 3; Leitung: Istat [Istituto Nazionale di Statistica – Statistikamt Italiens])

Neben diesen Arbeitspaketen gründeten sich im Verlauf des Projekts Arbeitsgruppen, um die Arbeiten zwischen den Arbeitspaketen 2 und 3 besser zu verknüpfen. Der folgende Abschnitt 3.2 stellt die Arbeitsgruppe zu rechtlichen Fragestellungen vor.

Dieser Beitrag kann nicht auf alle Teilaspekte des Projekts eingehen. Die detaillierten Projektergebnisse der beiden Arbeitspakete 2 und 3 stehen auf der [CROS-Seite des Projekts](#) zur Verfügung; darauf basieren die Abschnitte 3.3 bis 3.5.

### 3.2 Arbeitsgruppe zu rechtlichen Fragen

Zu Beginn des Projektes zeichnete sich ab, dass rechtliche Aspekte eine ernst zu nehmende Herausforderung für das Projekt, aber auch generell für die europäische Zusammenarbeit im Bereich von Trusted Smart Surveys darstellen. Bei der Vertragsgestaltung zur Nutzung einer App zur Zeitverwendungserhebung zeigte sich, dass die einzelnen Mitgliedstaaten die für alle Beteiligten geltende europäische Datenschutz-Grundverordnung (DSGVO) unterschiedlich interpretieren. Zudem stellt nationales Recht, zum Beispiel das deutsche IT-Grundschutz-Kompendium, weitere Anforderungen.

Um dieses Thema umfassender zu berücksichtigen, wurde während der Projektlaufzeit beschlossen, eine zusätzliche Arbeitsgruppe zu rechtlichen und ethi-

<sup>13</sup> So kann beispielsweise bei der Einkommens- und Verbrauchsstichprobe das Abfotografieren eines Kassenzettels zur Erfassung der Konsumausgaben einfacher und zeitsparender sein, als jede Position des Belegs einzeln einzutippen. Außerdem können Zeitpunkt und Inhalt der gestellten Fragen dynamisch auf der Grundlage von aus den Sensordaten abgeleiteten Kontextinformationen festgelegt werden.

<sup>14</sup> Siehe zudem auch die Informationen auf der Seite von Eurostat, dem Statistischen Amt der Europäischen Union: [ec.europa.eu](https://ec.europa.eu/eurostat)

schen Fragestellungen einzusetzen. Sie setzte sich aus Rechts- und Erhebungsdesignexperten und -expertinnen zusammen.

Die Arbeitsgruppe hat einen ausführlichen Bericht erstellt über neue Aspekte, die bei Trusted Smart Surveys im Vergleich zu herkömmlichen Umfragen zu beachten sind, und welche technischen und organisatorischen Maßnahmen dazu zu ergreifen sind. Der Bericht enthält zudem einen Entwurf für eine Datenschutz-Folgenabschätzung für Smart Surveys.

In einem Diskussionspapier setzte sich die Gruppe mit zwei wichtigen Designentscheidungen bei Trusted Smart Surveys auseinander:

- (1) mit der Einbeziehung von Dritten, beispielsweise Vertragspartnern, zur Bereitstellung von Apps, und
- (2) der Datenverarbeitung auf den zur Datenerhebung eingesetzten smarten Geräten gegenüber der Verarbeitung auf Servern, zum Beispiel in den Statistikämtern.

Zwar waren diese Arbeiten im Voraus nicht geplant, sie stellen jedoch wichtige grundlegende Überlegungen für die weitere Zusammenarbeit im Bereich Trusted Smart Surveys dar, auf denen andere Projekte aufbauen können. Dies ist nicht nur von erheblicher Bedeutung, um die rechtlichen Anforderungen an den Datenschutz zu erfüllen, sondern auch, um die Akzeptanz und Teilnahmerebereitschaft in der Gesellschaft zu fördern.

### 3.3 Arbeitspaket 1: Kommunikation und Koordinierung

---

Das Projekt startete Anfang 2020, also etwa zu der Zeit, als das Coronavirus Europa erreichte. Die Pandemie sorgte dafür, dass persönliche Treffen während fast der gesamten Laufzeit nicht möglich waren. Die Koordinations- und Halbezeittreffen mussten daher alle virtuell stattfinden. Zudem kam es aufgrund der Pandemie zu Verzögerungen, da Personal in den nationalen statistischen Ämtern anderweitig gebunden war. Persönliche Befragungen mithilfe von Interviewerinnen und Interviewern waren in Arbeitspaket 2 geplant, konnten aber nicht im ursprünglich vorgesehenen Zeitrahmen durchgeführt werden, sodass die Projektlaufzeit um ein halbes Jahr verlängert wurde.

Schließlich war es im Mai 2022 möglich, das Projekt mit einem [Abschlussworkshop](#) in Wien zu beenden. Das Projekt wurde darüber hinaus auf einigen (virtuellen) internationalen Konferenzen vorgestellt, etwa dem Weltkongress 2021 des Internationalen Statistischen Instituts (ISI)<sup>5</sup>, der [NTTS 2021](#)<sup>6</sup> sowie der [IAOS-Konferenz 2022](#)<sup>7</sup> und der [Q2022-Konferenz](#)<sup>8</sup>, dort mit einer eigenen Special Session.

### 3.4 Arbeitspaket 2: Pilotprojekte

---

Arbeitspaket 2 umfasste die empirische und angewandte Komponente des Projekts und war in vier Pilotprojekte zu den Themen Konsum, Zeitverwendung, Gesundheit und Lebensbedingungen gegliedert. Befragungen zum Konsum (Beispiel: Einkommens- und Verbrauchsstichprobe – EVS) und zur Zeitverwendung (Beispiel: Zeitverwendungserhebung – ZVE) sind sehr aufwendig für Befragte und gleichzeitig anfällig für fehlerhafte Angaben, zum Beispiel aufgrund von Erinnerungslücken. Dies zeigt sich beispielsweise in den Erhebungen, die einen tagebuchgestützten Erhebungsansatz verfolgen. Gerade in diesen Fällen erscheint der Einsatz von Apps angebracht. Fragen rund um die Verwendung solcher Apps waren daher Thema in den ersten beiden Pilotprojekten: Wie unterschiedlich sind die Anforderungen der verschiedenen nationalen Statistikämter? Wie können smarte Instrumente in den Apps die Belastung der Befragten reduzieren? Wie müssen Apps gestaltet und aufgebaut sein, damit die Befragten sie leicht und gut nutzen können? Kann eine App auch für andere, ähnliche thematische Fragestellungen genutzt werden? Wie ist die Datenqualität der per App erhobenen Daten? Wie können Befragte am besten für die Nutzung einer App gewonnen werden?

Die Pilotprojekte drei und vier waren kleiner und experimenteller. In beiden wurden Sensordaten mittels anderer Geräte als dem Smartphone erfasst, die Qualität der

---

5 International Statistical Institute World Statistical Congress.

6 Die New Techniques and Technologies for Statistics (NTTS) ist eine internationale zweijährliche wissenschaftliche Konferenzreihe, die vom Statistischen Amt der Europäischen Union (Eurostat) organisiert wird.

7 IAOS: International Association for Official Statistics.

8 Q2022: 10th European Conference on Quality in Official Statistics.

Daten bewertet, aber auch die Bereitschaft von Befragten für solche Erhebungen thematisiert.<sup>9</sup>

### Konsum

Im größten Pilotprojekt aus dem zweiten Arbeitspaket stand das Erfassen von Konsumausgaben, vor allem von Ausgaben des täglichen Gebrauchs, wie Einkäufen im Supermarkt, im Mittelpunkt. Es war daher thematisch eng mit dem europäischen Household Budget Survey (HBS) (in Deutschland: EVS), verbunden. Das Pilotprojekt ist das einzige im ESSnet-Smart-Surveys-Projekt, in dem auch groß angelegte, länderübergreifende Feldtests durchgeführt wurden: Die nationalen Statistikämter der Niederlande, Spaniens und Luxemburgs testeten die vom CBS (Niederlande) entwickelte @HBS App<sup>10</sup> mit insgesamt über 4 000 Personen in der Brutto-Stichprobe.

Zentraler Bestandteil der App sind die beiden Möglichkeiten, Ausgaben zu erfassen: manuell durch eigenes Eingeben der Ausgaben über eine Suchfunktion und automatisiert durch das Abfotografieren von Kassenzetteln. Hinter beiden Möglichkeiten verbergen sich smarte Funktionen: zum einen der Suchalgorithmus bei der Freitext-Eingabe, zum anderen das OCR-Scanverfahren zum Auslesen des Textes auf den Kassenzetteln und die anschließende Kodierung der Produkte anhand der [COICOP-Klassifikation](#). Etwa ein Viertel der in den Feldtests gemeldeten Einkäufe wurde durch die automatisierte Eingabe erfasst. Dabei zeigte sich, dass Befragte zusätzlich die Möglichkeit haben sollten, die eingescannten Daten zu sichten und bei Bedarf zu bearbeiten.<sup>11</sup>

Das Pilotprojekt hat zudem Folgendes verdeutlicht: Die Entwicklung von auf maschinellem Lernen basierten Modellen für die korrekte Kodierung und COICOP-Klassifizierung der Produkte sowie das Erstellen der umfangreichen Produktlisten in den jeweiligen Landessprachen

stellen entscheidende Herausforderungen dar für eine erfolgreiche Verwendung einer Haushaltstagebuch-App.

Die App bietet personalisierte Feedback-Möglichkeiten in Form von Auswertungsübersichten, die die eigenen Ausgaben der Befragten grafisch aufbereiten. Bei den Feldtests ließen sich jedoch keine eindeutigen Anzeichen für höhere Antwortquoten, eine geringere Abbruchquote und/oder eine bessere Datenqualität erkennen, wenn ein personalisiertes Feedback angeboten wird. Nichtsdestotrotz wurde es in den Usability-Tests sehr geschätzt und sollte daher beibehalten werden.

### Zeitverwendung

Das zweite Pilotprojekt beschäftigte sich mit der Erfassung aller Aktivitäten im Tagesablauf und war daher inhaltlich mit dem Harmonised European Time Use Survey (HETUS) (in Deutschland: ZVE) verknüpft. Ähnlich wie beim ersten Pilotprojekt kam auch hier eine bereits zuvor entwickelte App zum Einsatz, die getestet und weiterentwickelt wurde: die [Motus-App](#). Kernelement dieser App ist ein Tagebuch, in dem Befragte ihre Aktivitäten einschließlich einiger Zusatzinformationen eintragen. Zentral ist auch hier eine Suchfunktion, um die entsprechenden Tätigkeiten zu finden und der korrekten Kategorie zuzuordnen.

In mehreren Usability-Testrunden wurde die App getestet. Der Fokus variierte dabei: Neben der klassischen Zeitverwendungserhebung wurden auch Versionen getestet, die eher die Themen Arbeitszeit und Mobilität aufgriffen. In Richtung einer breiteren Einsatzmöglichkeit zielte auch der Test einer neu entwickelten Funktion, die mittels „Geo Fencing“ Geodaten erfasst und eine Aktion auslöst, sobald sich das Smartphone in der Nähe eines bestimmten Ortes befindet.<sup>12</sup> Zwar befindet sich diese Funktion noch in einem frühen Entwicklungsstadium und die Tests zeigten deutlichen Verbesserungsbedarf<sup>13</sup>, jedoch ist das grundlegende Motus-System insgesamt hinsichtlich Funktionalität und Nutzerfreundlichkeit bereits weit ausgereift. So konnte bei der deutschen Zeitverwendungserhebung 2022 dieses System mit einer angepassten App erstmals eingesetzt werden.

9 Die im Folgenden dargestellten Ergebnisse basieren auf den Projektergebnissen des Arbeitspaketes 2. Diese sind auf der CROS-Seite des Projektes zu finden: [ec.europa.eu](https://ec.europa.eu)

10 Siehe Endbericht zu Arbeitspaket 2, Seite 6: [ec.europa.eu](https://ec.europa.eu)

11 In den Feldtests war dies eine der Experimentierbedingungen: Manche Testpersonen hatten die Möglichkeit, ihre eingescannten Daten zu bearbeiten, andere nicht. Die Auswertung zeigt, dass die Datenqualität bei manueller Nachbearbeitung besser ist. Ausführlicher thematisiert wird dies in [Deliverable 2.3](#).

12 Ein Beispiel für eine solche Aktion ist das Senden einer Push-Nachricht mit dem Hinweis, einen kurzen Fragebogen zu beantworten.

13 [ec.europa.eu](https://ec.europa.eu)

## Gesundheit

Beim Thema Gesundheit lag der Fokus auf der Messung der körperlichen Aktivität. Kleinere Machbarkeitsstudien wurden in Belgien, den Niederlanden und Polen mittels eines einfachen Fitnesstrackers, der nur über einen Beschleunigungsmesser verfügt, durchgeführt. Daneben gab es drei größere Online-Befragungen in den Niederlanden<sup>14</sup>. Sie zielten darauf ab, die Akzeptanz solcher Erhebungen in der Öffentlichkeit zu ermitteln, Strategien zu untersuchen, mit denen Menschen zu einer Teilnahme animiert werden können, sowie herauszufinden, ob möglicherweise private Fitnesstracker genutzt werden können. Hierbei zeigte sich, dass die Befragten grundsätzlich bereit sind, an solchen Erhebungen teilzunehmen, dass aber zusätzliche Maßnahmen nötig sind, um die Akzeptanz und damit die Teilnahmebereitschaft zu erhöhen.

Für die Auswertung der Daten wurde versucht, eigene Modelle zu trainieren und Algorithmen abzuleiten mit dem Ziel, körperliche Aktivität erfassen und zuordnen zu können. Die Ergebnisse waren noch nicht zufriedenstellend. Hier ist weitere Forschung unumgänglich, da bereits aus den oben angeführten Gründen zu Vertrauen, Verantwortung und Transparenz eine alleinige Nutzung der von kommerziellen Anbietern zur Verfügung gestellten Analyse-Algorithmen nicht sinnvoll ist.

## Lebensbedingungen

Im vierten und letzten Pilotprojekt von Arbeitspaket 2 beschränkten sich die untersuchten Lebensbedingungen auf die Luftqualität in geschlossenen Räumen. Da Befragte in der Regel nicht über Messgeräte zum Erfassen der Luftqualität verfügen, wurden für diesen Pilotversuch entsprechende Geräte ([uHoo](#)) verteilt, die die Teilnehmenden für eine gewisse Zeit in ihrer privaten Wohnung aufstellen sollten.

In mehreren kleineren Machbarkeitsstudien – durchgeführt in Deutschland, den Niederlanden und Belgien – standen die Akzeptanz und die Benutzerfreundlichkeit der Messgeräte im Fokus. Dabei gaben manche Teilnehmende an, dass die Werte, die ihnen auf einer App zu ihrer Luftqualität angezeigt wurden, bei ihnen zu (dauer-

haften) Verhaltensänderungen führten, vor allem bezüglich des Lüftens und Heizens. Das ist für eine objektive Datenerfassung, welche die Situation ohne den durch die Messgeräte verursachten Einfluss erheben will, problematisch.

In der Wahrnehmung der Befragten steigt die Sensibilität der Daten abhängig von der Art und Menge der Verknüpfung mit Kontextdaten.<sup>15</sup> Zudem zeigte sich ein Unterschied zwischen den Ländern hinsichtlich der wahrgenommenen Sensibilität der Daten, wobei Befragte aus Deutschland und den Niederlanden die Daten als sensibler einschätzten als jene aus Belgien.

## 3.5 Arbeitspaket 3: Konzeptioneller Rahmen für eine Trusted-Smart-Surveys-Plattform

---

Das Hauptziel von Arbeitspaket 3 war, einen konzeptionellen Rahmen und technische Spezifikationen für eine europäische Plattform zum Austausch und zur Wiederverwendung von Smart-Survey-Lösungen und -Komponenten zu entwickeln. Dieser allgemeine Rahmen betrachtet die beteiligten Erhebungsprozesse aus verschiedenen Blickwinkeln (methodisch, technologisch, architektonisch) und bietet ein neues und nützliches Referenzschema für die nationalen statistischen Ämter.

Der Entwurf einer Referenzarchitektur ging von bestehenden Rahmenwerken aus. Im nächsten Schritt wurden Proofs of Concepts in Form von modularen Prototyp-Elementen für wesentliche Aspekte der Architektur durchgeführt, deren Ergebnisse zur Weiterentwicklung des konzeptionellen Rahmens dienten. Im letzten Schritt erfolgten auf Basis dieser Erkenntnisse Anpassung und Erweiterung des zuvor entworfenen konzeptionellen Rahmens für die Plattform.

Nach einem Blick auf die Referenzarchitektur thematisiert dieser Abschnitt noch den Einsatz von Modellen mit maschinellem Lernen und Technologien zum Schutz der Privatsphäre.<sup>16</sup>

---

14 Die drei Projekte in den Niederlanden waren Online-Befragungen, an deren Ende auch die Bereitschaft zur Teilnahme an Studien mit Fitnesstrackern abgefragt wurde. Genauere Informationen finden sich im Endbericht des Arbeitspaketes 2 ([ec.europa.eu](#), hier: Seite 10).

15 Beispielsweise mit sozioökonomischen Daten der Befragten oder Adressdaten.

16 Die im Folgenden dargestellten Ergebnisse basieren auf den Projektergebnissen des Arbeitspaketes 3. Diese sind auf der CROS-Seite des Projektes zu finden: [ec.europa.eu](#)

### Architektur

Die Gestaltung der Referenzarchitektur für Trusted Smart Surveys zielte darauf ab, die Beziehungen zwischen der Architektur, methodischen Komponenten und technologischen Überlegungen zu verdeutlichen. Dazu wurden ein „Business Layer“ und ein „Application Layer“ modelliert und die „Building Blocks“ beschrieben.

Das Hauptziel der Modellierung des Business Layers ist, die wichtigsten Aktivitäten und Verhaltensweisen zu identifizieren und zu standardisieren, die von den nationalen statistischen Ämtern im Zuge der Datenerfassung und -verarbeitung bei Trusted Smart Surveys durchzuführen sind. Dabei standen die folgenden Fragen im Vordergrund: Welche Auswirkungen haben Trusted Smart Surveys auf die GSBPM-Phasen<sup>17</sup> und Prozessschritte? Welche Schritte des Prozesses können standardisiert und in einer gemeinsamen Infrastruktur ausgeführt werden? Welche Tätigkeiten können in der Plattform oder auf den Geräten der Auskunftgebenden ausgeführt und welche Aufgaben müssen von den nationalen statistischen Ämtern intern durchgeführt werden? Was ist der beste Ansatz, um die spezifischen nationalen Bedürfnisse und die Besonderheiten der an einem Prozess beteiligten Akteure zu harmonisieren?

Der Entwurf des Application Layers ist der erste Schritt zur Definition einer vorläufigen Reihe von Bausteinen (Building Blocks), die für die Ausführung der Prozessschritte eines allgemeinen Trusted Smart Surveys erforderlich sind. Diese Bausteine werden in Form von Funktionalitäten beschrieben, die von der Plattform angeboten werden könnten. Sie dienen als Basis, um Annahmen über die Bereitstellung und gemeinsame Nutzung von Diensten abzuleiten.

Durch die Arbeiten im Arbeitspaket 3 zeigte sich, dass eine allumfassende, von den Fachgebieten unabhängige Plattform für Trusted Smart Surveys, wie sie anfangs angedacht war, äußerst komplex und unrealistisch ist. Zudem wäre sie nicht für alle Bereiche notwendig. Komplett unabhängige, auf einzelne Erhebungen zugeschnittene Anwendungen zu entwickeln, erscheint jedoch ebenfalls nicht zielführend. Der Vorschlag ist daher, die Anwendungen für Trusted Smart Surveys flexibel zu ent-

wickeln und klar zwischen allgemeinen Komponenten und dem, was bereichsabhängig ist, zu unterscheiden. Die konkrete Ausgestaltung der Erfassung und Verarbeitung von Smart-Survey-Daten je nach verwendeten Sensoren und Datenquellen sollte angepasst werden, ebenso Aspekte wie das Monitoring der Erhebung.

Die Modularität ist dabei der zentrale Punkt, um ein Gleichgewicht zwischen Verallgemeinerung und Spezifität herzustellen und so die Implementierung wiederverwendbarer Komponenten zu fördern. In der konkreten Umsetzung könnte mit gemeinsam nutzbaren Bausteinen begonnen werden.

Die technische Infrastruktur muss dabei dieser Modularität Rechnung tragen. Das bedeutet, dass sie für die verschiedenen Arten von Anwendungsschnittstellen (APIs) für die Bereitstellung von Daten (verschiedene Sensoren, Umfragedaten) ausgelegt sein muss. Zudem muss sie es ermöglichen, technische Komponenten für Trusted Smart Surveys einheitlich zu spezifizieren. Und zwar unabhängig davon, ob die Daten im Gerät oder auf der Plattform verarbeitet werden oder ob die bereitgestellten Daten unstrukturiert oder strukturiert sind.

### Machine Learning in Smart Surveys

Modelle, die auf maschinellem Lernen (Machine Learning – ML) basieren, können zum einen eingesetzt werden, um von smarten Geräten gelieferte Eingabedaten (zum Beispiel Bilder, Signale, Sprache) während der Datenerfassung zu verarbeiten. Zum anderen können sie hilfreich sein, um Sensordaten zusammenzuführen und Fehler (Anomalien in den Daten, fehlende Daten und Ähnliches) in der Datenverarbeitungsphase zu beheben. Die ML-Anwendungen können dabei entweder auf den Geräten der Befragten oder zentral auf Systemen der nationalen statistischen Ämter (beispielsweise in einer Cloud oder auf deren Servern) ausgeführt werden. Welche der beiden Lösungen zu wählen ist, hängt von den Vor- und Nachteilen des jeweiligen Ansatzes und den Anforderungen der Erhebung ab.

Für ML-Anwendungen direkt in der App spricht zum einen der Schutz persönlicher Daten, da Rohdaten das Gerät der Nutzenden gar nicht erst verlassen, sondern nur die relevanten Daten in aufbereiteter Form. Zum anderen sind ML-Anwendungen direkt in der App nicht auf eine Internetverbindung angewiesen. Die direkte Verarbeitung vor Ort bietet außerdem schnellere Reak-

17 Generic Statistical Business Process Model, siehe hierzu [statswiki.unece.org](https://www.statswiki.unece.org). Das deutsche GMAS-Konzept (GMAS = Geschäftsprozessmodell Amtliche Statistik) ist hiervon abgeleitet.

tionsmöglichkeiten, zum Beispiel zur Qualitätskontrolle oder für Rückmeldungen an Befragte.

Bei ML-Anwendungen an einem zentralen Ort, beispielsweise in einer Cloud oder auf den Servern der nationalen statistischen Ämter, können die Modelle komplexer sein. Sie können mehr Speicherplatz und Rechenleistung in Anspruch nehmen, da beides weniger beschränkt ist als auf mobilen Endgeräten. Dies könnte sich positiv auf die Qualität der Ergebnisse auswirken. Wartung und Aktualisierung der Algorithmen sind intern möglich, ohne dass die Anwendung der Nutzenden geändert oder aktualisiert werden muss.

Im Proof of Concept zum maschinellen Lernen wurde eine verallgemeinerte ML-Komponente entwickelt, die einem unabhängigen Ansatz folgt, um die Daten von verschiedenen Sensortypen zu verarbeiten. Dabei gelang es, aus den verschiedenen Sensordaten mithilfe des ML-Verfahrens mehrere Variablen abzuleiten, zum Beispiel die Art der körperlichen Aktivität oder das Verkehrsmittel der befragten Person. Hierfür wurde eine Wavelet + Convolutional-Neural-Network(CNN)-Pipeline entwickelt. Der Proof of Concept hat gezeigt, dass dieses sehr breit angelegte Instrument genutzt werden kann, um Sensordaten von Endgeräten so zu verarbeiten, dass nützliche, interpretierbare Variablen entstehen.

### Technologien zum Schutz der Privatsphäre

Wie bereits erläutert, enthalten Trusted-Smart-Surveys-Daten ein Übermaß an Informationen über die Befragten, was ein größeres Risiko für die Privatsphäre mit sich bringt als herkömmliche Umfragen. Traditionell werden Methoden wie die Pseudonymisierung verwendet, um zu verhindern, dass die privaten Daten der Befragten den Mitarbeiterinnen und Mitarbeitern in den nationalen statistischen Ämtern zugänglich sind. Außerdem werden digital erhobene Daten während der Übertragung verschlüsselt. Diese Maßnahmen sollen natürlich auch bei Smart Surveys angewendet werden. Möglicherweise verursacht die erhöhte Sensibilität der von den intelligenten Sensoren erfassten Daten jedoch strengere Datenschutzanforderungen. Um diesen gerecht zu werden, könnten zusätzliche Maßnahmen erforderlich sein, zum Beispiel die Datenminimierung: Hierbei soll die Menge der in den gesammelten Daten enthaltenen Informationen auf das für einen bestimmten Zweck erforderliche Maß reduziert werden. Die Datenminimie-

rung schützt die Privatsphäre, indem sie die Möglichkeit ausschließt, aus den Daten zusätzliche sensible Informationen zu gewinnen. Dieses Konzept ist nur von begrenzter Bedeutung, wenn es sich um herkömmliche, fragebogengestützte Erhebungen handelt. Es ist jedoch äußerst relevant, wenn es um sensorgestützte Smart Surveys geht, da die Sensordaten (wahrscheinlich) weit mehr als die benötigten Informationen bieten.

Als zweite Methode sei die Verarbeitung und Analyse von Daten genannt, ohne dass die darin enthaltenen Informationen zugänglich werden. Dies ist die bei weitem stärkste Maßnahme zum Schutz der Privatsphäre, aber schwierig zu realisieren. Es gibt verschiedene Techniken, die eine verdeckte Analyse für unterschiedliche Anwendungen und Szenarien ermöglichen. So kann beispielsweise das Training von ML-Modellen durchgeführt werden, ohne dass es notwendig ist, die Daten zentral zu erfassen und zu speichern.<sup>18</sup>

Da solche Anwendungen zum Schutz der Privatsphäre für verschiedene GSBPM-Phasen (beispielsweise die Phasen 4 [Daten gewinnen], 5 [Daten aufbereiten] und 6 [Ergebnisse analysieren]) relevant sein können, ist die Entwicklung dieser Maßnahmen bereits bei der Konzeption der Erhebung mitzudenken. Es ist jedoch zu beachten, dass die Implementierung solcher Technologien mit einer erhöhten Komplexität der Plattformarchitektur und des Umfragedesigns einhergeht. Daher ist es wichtig, die Anforderungen an den Schutz der Privatsphäre auf Grundlage der Sensibilität der Daten genau zu bestimmen, um die Anwendung übermäßig komplexer Technologien zu vermeiden.

Auch im Bereich der Privatsphäre wurde ein Proof of Concept durchgeführt. Dessen Hauptziel bestand darin, die Anwendbarkeit von Technologien zum Schutz der Privatsphäre zu testen, um die Erstellung von Statistiken ohne Zugang zu Mikrodaten der Befragten zu ermöglichen. Dabei wurden einige Überlegungen zur Nutzung von Aggregatdaten in einer Multiparty-Computation(MPC)-Umgebung aufgezeigt und demonstriert, dass die Nutzung einer solchen Technologie grundsätzlich möglich ist.

Bei der Umsetzung zeigte sich, dass die Verwendung von Open-Source-MPC-Software viele zusätzliche Implemen-

---

18 Nähere Informationen dazu bietet der Abschnitt 3.2.3 in Deliverable 3.2.

tierungen erfordert. Einige dieser Implementierungen gehen über das übliche Software-Engineering hinaus und setzen Expertenwissen in Kryptografie oder Statistik voraus, um sie ordnungsgemäß durchführen zu können. Die entwickelte Anwendung erfüllt die für den Proof of Concepts gestellte Anforderung weitgehend, Auswertungen zu ermöglichen, Einzeldaten dabei aber zu schützen und nicht zugänglich zu machen. Allerdings umfasste die statistische Analyse nur sehr grundlegende Aggregationen. Um zu prüfen, ob ein solches Design alle Anforderungen der modernen statistischen Analyse erfüllt, sind weitere, tiefergehende Tests erforderlich.

## 4

---

### Ausblick

---

Der Einsatz von Apps und das Erheben von Sensordaten öffnet die Tür in eine Welt voller neuer Erhebungsmöglichkeiten. Trotz aller großen Potenziale kann dieser Weg aber nur dann erfolgreich zu beschreiten sein, wenn sichere Systeme und Vorgehensweisen bereitstehen, denen die Befragten vertrauen. Schließlich sind sie es, die in ausreichender Zahl und Repräsentativität an den Erhebungen teilnehmen müssen und ihre Daten zur Verfügung stellen sollen – nicht selten durch die Installation einer App auf ihren Smartphones. Nur mit ihrer Mitwirkung ist eine verlässliche Statistikproduktion möglich.

Das Essnet-Smart-Surveys-Projekt war ein erster Schritt. Durch grundlegende Überlegungen und Arbeiten wurden in Arbeitspaket 3 Möglichkeiten und Grenzen einer kombinierten Plattform und einzelner Datenverarbeitungsinstrumente betrachtet. Die vier Pilotprojekte aus Arbeitspaket 2 testeten mögliche Einsatzbereiche und Anforderungen aus der Sicht von Befragten. Insgesamt wurde deutlich, dass es noch einige Herausforderungen auf dem Weg zu einem sicheren, verlässlichen und transparenten Smart-Survey-Datengewinnungs- und -verarbeitungssystem gibt. Gleiches gilt für die Interaktion mit den Befragten an sich, um sie für diese neue Art der Datenerhebung zu gewinnen. Einen weiteren, nächsten Schritt in diese Richtung geht Eurostat mit dem Projekt „Smart Survey Implementation“, das an das ESSnet-Smart-Surveys-Projekt anknüpft. Zusammen mit zehn weiteren Projektpartnern wird sich das Statistische Bundesamt in den nächsten zwei Jahren an diesem Projekt beteiligen. Mit ersten Ergebnissen ist Mitte 2024 zu rechnen. 

## LITERATURVERZEICHNIS

---

Cukier, Kenneth/Mayer-Schoenberger, Viktor. *The Rise of Big Data: How It's Changing the Way We Think About the World*. In: Foreign Affairs. Ausgabe Mai/Juni 2013, Seite 28 ff.

Keusch, Florian/Struminskaya, Bella/Antoun, Christopher/Couper, Mick P./Kreuter, Frauke. *Willingness to Participate in Passive Mobile Data Collection*. In: Public Opinion Quarterly. Ausgabe 83/2019, Seite 210 ff.

Ricciato, Fabio/Wirthmann, Albrecht/Giannakouris, Konstantinos/Reis, Fernando/Skaliotis, Michail. *Trusted smart statistics: Motivations and principles*. In: Statistical Journal of the IAOS. Ausgabe 35/2019, Seite 589 ff.

Ricciato, Fabio/Wirthmann, Albrecht/Hahn, Martina. *Trusted Smart Statistics: How new data will change official statistics*. In: Data & Policy. Ausgabe e7/2020.

Ricciato, Fabio/Giannakouris, Kostas/Wirthmann, Albrecht/Hahn, Martina. *Trusted Smart Surveys: a possible application of Privacy Enhancing Technologies in Official Statistics*. Conference paper: 50th Scientific Meeting of the Italian Statistical Society (SIS 2020). Rom 2020.

Ruppert, Evelyn/Grommé, Francisca/Ustek-Spilda, Funda/Cakici, Baki. *Citizen Data and Trust in Official Statistics*. In: Economics and Statistics. Ausgabe 505–506/2018, Seite 179 ff.

Stehrenberg, Shari/Giannakouris, Konstantinos. *Trusted Smart Surveys: Solutions for the European Statistical system: An overview of the objectives and the main challenges*. Conference paper: European Conference on Quality in Official Statistics (Q2022). Vilnius 2022.

Struminskaya, Bella/Lugtig, Peter/Keusch, Florian/Höhne, Jan Karem. *Augmenting Surveys With Data From Sensors and Apps: Opportunities and Challenges*. In: Social Science Computer Review. Ausgabe Dezember 2020, Seite 1 ff.

## RECHTSGRUNDLAGEN

---

Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung) (Amtsblatt der EU Nr. L 119, Seite 1).

# MODERNISIERUNG DER STATISTIK-PRODUKTION DURCH DIE NEUE GEHEIMHALTUNGS-INFORMATIONSPLATTFORM

Arijana Amina Ramic

➤ **Schlüsselwörter:** Statistikproduktionsprozess – Standardisierung – Geheimhaltung – Geheimhaltungs-Informationenplattform – Metadatenmanagementsystem

## ZUSAMMENFASSUNG

Eine Möglichkeit, die Statistikproduktion zu modernisieren, besteht in der Standardisierung und Optimierung der Prozesse: Die Reproduzierbarkeit von Methoden, Prozessen und Werkzeugen sorgt für verbesserte Qualität, Transparenz und Wirtschaftlichkeit. Die Bereitstellung einer zentralen, metadatengetriebenen Geheimhaltungs-Informationenplattform reduziert die Dokumentationsaufwände der statistischen Fachbereiche erheblich, zudem können sich diese leichter an der Geheimhaltungspraxis vergleichbarer Fachstatistiken orientieren. Der Artikel beschreibt die Entwicklung dieser Geheimhaltungs-Informationenplattform und deren Nutzung bei der Umstellung der zentralen und dezentralen Bundesstatistiken auf eine standardisierte und automatisierte Geheimhaltung.

➤ **Keywords:** *statistical production process – standardisation – statistical disclosure control – information platform for statistical disclosure control – metadata management system*

## ABSTRACT

*One way of modernising statistical production is to standardise and optimise the processes. The reproducibility of methods, processes and tools ensures better quality, transparency and cost-effectiveness. A central, metadata-driven information platform for statistical disclosure control significantly reduces the documentation efforts of the specialised statistics units. Additionally, they can more easily use the disclosure control practices that are applied to comparable statistics. This article describes the development of the information platform for statistical disclosure control and its use in implementing standardised and automated disclosure control in centrally and decentrally produced federal statistics.*



**Arijana Amina Ramic**

studierte Mathematik mit Informationstechnologien und ist seit insgesamt 18 Jahren in der amtlichen Statistik tätig. Im Statistischen Bundesamt war sie im Referat „IT-Gremien, SteP und IT-Verbundkoordination“ mitverantwortlich für die Standardisierung und Optimierung der fachstatistischen Produktionsprozesse einschließlich der Geheimhaltung. Bei der Entwicklung der Geheimhaltungs-Informationenplattform agierte sie als Product Ownerin, bevor sie ins Referat „Technische und fachliche Unterstützung, Automation im Außenhandel“ wechselte.

## 1

### Einleitung

Standardisierung im Kontext dieses Artikels bezeichnet den Prozess der Entwicklung und Implementierung technischer Standards mit dem Ziel, die Prozesse zu optimieren und zu modernisieren. Prozessstandardisierung beschreibt die Festlegung eines Regelwerks, wie eine bestimmte Aufgabe oder Abfolge von Aufgaben ausgeführt werden soll. Durch die Reproduzierbarkeit von Methoden, Prozessen und Werkzeugen sorgt Standardisierung für verbesserte Qualität, Transparenz und Wirtschaftlichkeit des Produktionsprozesses. Das international abgestimmte Generische Statistische Geschäftsprozessmodell (Generic Statistical Business Process Model – GSBPM) beschreibt den typischen Produktionsablauf einer amtlichen Statistik und ist somit unabhängig von den Erhebungsinhalten, der gewählten Methodik oder den Besonderheiten einer Statistik (UNECE, 2022). Das GSBPM teilt den statistischen Leistungsprozess in acht Phasen ein, die jeweils in mehrere Teilprozesse (insgesamt 44) unterteilt sind; zu übergreifenden Prozessen zählen das Qualitäts-, das Metadaten- und das Datenmanagement. Im Statistischen Verbund<sup>1</sup> gibt es unterschiedliche Arbeitsgruppen sowie Projekte und Verfahren, deren Ziele die Standardisierung und Optimierung der Statistik(teil)prozesse sind (Blumöhr und andere, 2017). Einer dieser Prozesse, die standardisiert ablaufen sollen, ist der Teilprozess zur Gewährleistung der statistischen Geheimhaltung.

Das folgende Kapitel 2 schildert die Überlegungen, die zum Aufbau einer Geheimhaltungs-Informationsplattform geführt haben. Deren Entwicklung wird in Kapitel 3 beschrieben, während Kapitel 4 die Vorteile der Plattform und die mit ihrem Betrieb verbundenen Herausforderungen benennt. Der Beitrag schließt mit einem Ausblick auf künftige Nutzungsmöglichkeiten.

<sup>1</sup> Den Statistischen Verbund bilden die Statistischen Ämter des Bundes und der Länder.

## 2

### Überlegungen zum Aufbau einer Geheimhaltungs-Informationsplattform

Der Schutz von Einzelangaben bei der Verarbeitung und Veröffentlichung statistischer Ergebnisse hat für die amtliche Statistik einen hohen Stellenwert (Rothe, 2015). Die Gewährleistung der Geheimhaltung hat daher Eingang in den Verhaltenskodex für europäische Statistiken (Europäische Union, 2017) gefunden und ist national in § 16 Absatz 1 Bundesstatistikgesetz geregelt. Die Verpflichtung zur statistischen Geheimhaltung stellt demnach international ein grundlegendes Prinzip der amtlichen Statistik dar und keine nationale Besonderheit.

Wird eine Statistik erstmals konzipiert oder wird das Auswertungskonzept überarbeitet, sind zu den Erfordernissen der Fachstatistik passende Methoden zur Geheimhaltung und Regelung des Zugangs zu vertraulichen Ergebnissen und Produkten zu konzipieren und einzuführen. Die Sicherstellung der Geheimhaltung sollte dabei ein einheitlicher Prozess im Statistischen Verbund sein und zwischen allen Anwendungsbereichen abgestimmt werden. Einen Prozess, der schrittweise zur Entwicklung eines einheitlichen Geheimhaltungsverfahrens mit möglichst automatisierter Umsetzung führt, hat die Expertengruppe für Statistische Geheimhaltung im Statistischen Verbund innerhalb des Vorgehenskonzepts „Einheitliche Geheimhaltung in allen Statistiken mit Geheimhaltungsbedarf“ vorgeschlagen. Die zuständigen übergeordneten Gremien der deutschen amtlichen Statistik haben diesem Vorschlag im Oktober 2019 zugestimmt, sodass das Vorgehenskonzept seit Anfang 2020 in allen Fachbereichen im Statistischen Verbund Anwendung findet (Gießing und andere, 2020).

Dieser Prozess gliedert sich in vier aufeinander aufbauende Stufen und wird als ein Standardprozess im Statistischen Verbund aufgefasst:

- › Stufe I – Prüfung des Geheimhaltungsbedarfs: umfasst (unter anderem) die Prüfung, ob bestimmte Merkmale oder Veröffentlichungen von der Pflicht der Geheimhaltung ausgenommen sind
- › Stufe II – Bestimmung der Geheimhaltungsmethode: umfasst die Prüfung infrage kommender Geheimhaltungsmethoden

- › Stufe III – Erstellung eines abgestimmten Geheimhaltungskonzepts: umfasst die Entwicklung und Abstimmung eines statistikspezifischen Fachkonzepts für eine im Statistischen Verbund einheitliche, abgestimmte Geheimhaltung
- › Stufe IV – Aufbau und Optimierung eines automatisierten und medienbruchfreien<sup>12</sup> Ablaufs der Geheimhaltung („Digitalisierung“): umfasst die technische Umsetzung des in den Stufen I bis III geplanten Geheimhaltungsprozesses (Gießing und andere, 2020)

Um diesen Standardprozess zur einheitlichen Geheimhaltung in allen Statistiken mit Geheimhaltungsbedarf zu unterstützen, hat die Expertengruppe für Statistische Geheimhaltung im Statistischen Verbund den Aufbau einer Geheimhaltungs-Informationsplattform vorangetrieben. Die Geheimhaltungs-Informationsplattform dient als zentrales Werkzeug für die Dokumentation der statistikspezifischen Geheimhaltung der Fachbereiche des Statistischen Verbunds. Der Zugriff auf die Geheimhaltungs-Informationsplattform wird über das interne Netzwerk des Statistischen Verbunds nur deren Beschäftigten ermöglicht und im Detail und Maße wie mit Rechten und Rollen der Nutzerinnen und Nutzer über die zentrale Benutzer- und Rechteverwaltung im Verbund bestimmt (siehe auch Abschnitte 3.2 und 3.3). Eine nutzerfreundliche Weboberfläche ermöglicht den Fachbereichen eine einfache und zeitnahe Berichterstattung über Stand und Praxis ihrer statistikspezifischen Geheimhaltung. Die Geheimhaltungs-Informationsplattform erleichtert durch die zentrale Bereitstellung relevanter Daten das kontinuierliche Monitoring der Umstellung der Geheimhaltung auf standardisierte und automatisierte Verfahren entsprechend der oben beschriebenen Stufen sowie die Berichterstattung an Leitungsebenen und Gremien des Statistischen Verbunds. Sie wurde dabei als ein weiteres Modul des Metadatenmanagementsystems des Statistischen Verbunds umgesetzt. Metadaten sind Daten, die Informationen über andere Daten enthalten („Daten über Daten“). Sie beschreiben Inhalt und Bedeutung von Daten und sind für die Interpretation statistischer Wertgrößen unabdingbar. Metadaten spielen eine wichtige Rolle und sind innerhalb des statistischen Produktionsprozesses auf operativer Ebene zu verwal-

ten. Gutes Metadatenmanagement ist daher für den effizienten Betrieb statistischer Geschäftsprozesse wesentlich. Im Rahmen des GSBPM liegt der Schwerpunkt des übergreifenden Prozesses der Metadatenverwaltung auf der Erstellung, Überarbeitung, Aktualisierung, Nutzung und Archivierung statistischer Metadaten (Lindenstruth und andere, 2017).

Die Bereitstellung einer zentralen, metadatengetriebenen Geheimhaltungs-Informationsplattform reduziert die Dokumentationsaufwände der Fachbereiche erheblich, zudem können diese sich leichter an der Geheimhaltungspraxis vergleichbarer Fachstatistiken orientieren. Dies unterstützt den Prozess der Umstellung der zentralen und dezentralen Statistiken auf eine standardisierte und automatisierte Geheimhaltung, die in den strategischen Maßnahmen der amtlichen Statistik von Deutschland<sup>13</sup> gefordert ist.

### 3

## Entwicklung einer Geheimhaltungs-Informationsplattform

---

### 3.1 Hintergrund

---

Informationen, Wissen und Dokumentationen zum Thema Geheimhaltung sowie zur Umstellung auf eine automatisierte und verbundeinheitliche Geheimhaltung im Statistischen Verbund liegen sehr unterschiedlich vor. Dies stellt sowohl für Fach- und Leitungsebenen als auch für diejenigen, die solche Informationen suchen, ein großes Problem dar. Davon ausgehend erfolgte die Entwicklung einer Geheimhaltungs-Informationsplattform im Statistischen Verbund, über die auf die für die Geheimhaltung relevanten Informationen zentral über eine Anlaufstelle zugegriffen werden kann. Sie stellt die Informationen systematisiert und maschinell verarbeitet.

---

<sup>12</sup> Ein Medienbruch entsteht in der Informationsverarbeitung, wenn in der Übertragungskette eines Prozesses Daten/Informationen von einem auf ein weiteres/anderes Informationsmedium übertragen werden (müssen).

<sup>13</sup> Standardisierung ist eines der acht Handlungsfelder der strategischen Schwerpunktplanung im Statistischen Verbund. Um die Prozesstransparenz und -standardisierung sowie Qualität zu erhöhen, wurden konkrete Maßnahmen zum Einsatz einheitlicher STEP-Werkzeuge im Erhebungs-, Aufbereitungs- und Auswertungsprozess und zur Umsetzung eines Musterprozessablaufs definiert. Dazu zählen unter anderem: die zentralen und dezentralen Statistiken auf eine standardisierte und automatisierte Geheimhaltung umstellen, Online-Meldeverfahren optimieren, automatischer Adressänderungs-workflow, digitale Assessments durchführen.

bar bereit und ermöglicht darüber hinaus eine weitgehend automatisierte, metadatengetriebene Erstellung beziehungsweise Aktualisierung der Geheimhaltungsleitfäden<sup>4</sup> durch die Fachbereiche. Die statistikspezifischen Geheimhaltungsleitfäden sind ein Instrument des Qualitätsmanagements. Sie dienen dazu, umfassende geheimhaltungsrelevante Informationen aufzubereiten, zu dokumentieren und zu archivieren. Bei dezentralen Statistiken sind diese zu einem Gesamtbericht zusammenzuführen. Die Geheimhaltungsleitfäden ermöglichen damit den Fachbereichen und den zuständigen Leitungsebenen nachträglich die Dokumentation, Beobachtung und Steuerung der Geheimhaltung.

Eine Bestandsaufnahme der geheimhaltungsrelevanten Informationen im Statistischen Verbund erfolgte 2018/2019 durch eine Erhebung über ein standardisiertes IDEV<sup>5</sup>-Formular. Dieses erforderte zur Abbildung der Geheimhaltung in den unterschiedlichen Statistiken und im Statistischen Verbund den Einsatz einer aufwendigen Filter- und Fragebogenführung. Ein statistikspezifischer Geheimhaltungsleitfaden wurde für jede Statistik erzeugt, für die in der IDEV-Abfrage Auskunft erteilt worden war. Der Aufbau der Dokumente orientierte sich dabei eng am Aufbau des IDEV-Formulars. Für die für den Bund-Länder-Abgleich notwendigen Vorbelegungen sowie statistikübergreifende Prüfungen, Aufbereitung und Auswertungen wurden Ad-hoc-Lösungen in unterschiedlichen Programmen und Werkzeugen erstellt, ebenso für die Erzeugung des statistikspezifischen Geheimhaltungsleitfadens auf Einzelstatistikebene als editierbares Word-Dokument. Die allgemeinen Informationen zu Statistiken stammen aus bestehenden Systemen; aufgrund regelmäßiger Aktualisierungen sollten sie auch für die Vorbereitung der statistikspezifischen Geheimhaltungsleitfäden direkt aus diesen Systemen übernommen werden.

Die Weiterentwicklung der Geheimhaltung in den Fachstatistiken der Statistischen Ämter des Bundes und der Länder ist ein laufender Prozess. Als Folgeprozess der Bestandsaufnahme wird die Thematik „Geheimhaltung“ in den nächsten Jahren häufig zu Änderungen in vielen Statistiken führen. Diese kontinuierlichen Wei-

terentwicklungen sind entsprechend der Zielsetzung zur Erstellung der Geheimhaltungsleitfäden zu dokumentieren. Die für die erstmalige Durchführung der Bestandsaufnahme verwendete Ad-hoc-Lösung genügt dabei in keinerlei Hinsicht den Anforderungen des erforderlichen dauerhaften Monitorings. Für ein konstantes Monitoring soll für jede Statistik regelmäßig die Aktualität der Geheimhaltungsleitfäden geprüft und Änderungen dokumentiert werden können. Die Überprüfung soll jederzeit Auswertungen für die betroffenen Fachstatistiken, aber auch für die Leitungsebenen ermöglichen. Die erstmalige Durchführung der Bestandsaufnahme zeigte, dass das bisher eingesetzte Verfahren auf Dauer weder für die Datengewinnung noch für die Zusammenstellung und Analyse der gewonnenen Informationen für den flächendeckenden Einsatz geeignet ist. Dies gilt insbesondere auch für den erforderlichen Abgleich der Angaben aus Bund und Ländern. Die für das Monitoring der Umstellung der Statistiken auf eine einheitliche und automatisierte Geheimhaltung benötigten Informationen wurden in unterschiedlichen Formen und Formaten gesammelt: aus den eingegangenen statistikspezifischen Geheimhaltungsleitfäden sowie aus den schriftlichen und mündlichen Rückmeldungen der Fachbereiche (bei zentralen Statistiken) beziehungsweise Referentenbesprechungen (bei dezentralen Statistiken).

Eine dauerhafte manuelle beziehungsweise aufwendig zu entwickelnde und zu pflegende Aufbereitung und Auswertung der Geheimhaltungsleitfäden durch Ad-hoc-Programme ist unter Berücksichtigung von Wirtschaftlichkeitsaspekten auszuschließen. Daher war eine technische Lösung notwendig, die eine weitgehend automatisierte, metadatengetriebene Erstellung der Geheimhaltungsleitfäden ermöglicht. Für das kontinuierliche Monitoring der Geheimhaltung sollte folglich eine Informationsplattform konzipiert werden, die sich aus dem Metadatenmanagementsystem speist. Da im Metadatenmanagementsystem Metadaten zu jeder Statistik enthalten sind, müssen diese nur einmal gepflegt werden und stehen ebenfalls für die Geheimhaltungsleitfäden zur Verfügung. Die Aktualisierung der Geheimhaltungsleitfäden beschränkt sich so auf fachliche Informationen zur Geheimhaltung, zum Beispiel eine geänderte Geheimhaltungsmethodik.

---

<sup>4</sup> Ein Geheimhaltungsleitfaden fasst die geheimhaltungsspezifischen Angaben der Statistischen Ämter des Bundes und der Länder für eine bestimmte Statistik zusammen.

<sup>5</sup> IDEV (Internet-Datenerhebung im Verbund) ist ein formulargestütztes Standarderhebungswerkzeug der amtlichen Statistik in Deutschland.

### 3.2 Ziel und Inhalte der Geheimhaltungs-Informationsplattform

---

Kernziele der Geheimhaltungs-Informationsplattform sind, die Geheimhaltung in den Fachstatistiken von Bund und Ländern zu dokumentieren sowie die Steuerung und ein einfaches Monitoring der Geheimhaltung zu ermöglichen. Hierfür bündelt die Plattform alle relevanten Metadaten zur statistikspezifischen Geheimhaltung an einer Stelle. Sie ermöglicht so eine zentrale Darstellung und Pflege dieser Informationen, ebenso wie die automatisierte Erstellung von statistikspezifischen Geheimhaltungsleitfäden.

Auf diese Weise

- › wird für alle Statistiken an einer zentralen Stelle die Geheimhaltungsmethodik dokumentiert,
- › kann jede Fachstatistik jederzeit nach Bedarf die Aktualität ihrer eigenen Geheimhaltungsleitfäden über eine grafische Benutzeroberfläche prüfen, Änderungen dokumentieren und archivieren sowie deren Änderungshistorie einsehen,
- › können Geheimhaltungsleitfäden für neue Statistiken angelegt werden,
- › können Fachstatistiken und Leitungsebenen jederzeit (entsprechende Berechtigungen vorausgesetzt) statistikübergreifende Auswertungen der dokumentierten Geheimhaltung durchführen (die Auswertung erfolgt nach ihrer Anforderung automatisiert),
- › können Nutzerinnen und Nutzer die Geheimhaltungsleitfäden aus anderen Fachbereichen betrachten (ohne statistikspezifische geheimzuhaltende Parameter – abhängig von Rollen und Rechten der Nutzenden),
- › werden die Aufwände für Erstellung und Aktualisierung der Geheimhaltungsleitfäden verringert.

Der Aufwand für regelmäßige Berichterstattungen reduziert sich, da die Fachbereiche mit der Geheimhaltungs-Informationsplattform selbst angeben, welcher Stand erreicht ist: Regelmäßige Abfragen zum Monitoring des Umsetzungsstands bei der statistikspezifischen Geheimhaltung entfallen hiermit. Ebenso kann sich die Leitungsebene schnell einen Überblick verschaffen, da alle relevanten Informationen an einem Ort vorgehalten werden und auf Grundlage dieser Informationen automatisierte

oder On-demand-Standardauswertungen ausgegeben werden können.

Die wichtigsten Stakeholder der Geheimhaltungs-Informationsplattform lassen sich in zwei Zielgruppen einteilen:

- › Fachstatistikerinnen und Fachstatistiker, die für die Sicherstellung der Geheimhaltung einer Statistik verantwortlich sind. Sie nutzen die Geheimhaltungs-Informationsplattform, um die statistikspezifischen Geheimhaltungsangaben zu pflegen und sich über die Geheimhaltungspraxis in der eigenen sowie gegebenenfalls in anderen Statistiken zu informieren.
- › Gremien und Leitungsebenen der Statistischen Ämter des Bundes und der Länder: Diesem Nutzerkreis dient die Geheimhaltungs-Informationsplattform vor allem für administrative Tätigkeiten sowie statistikübergreifende Auswertungen der Angaben zur Geheimhaltung zu Monitoringzwecken; er kann dafür auf alle Statistiken und Eintragungen zugreifen.

Da es sich beim Monitoring zur Umsetzung und Fortentwicklung der Geheimhaltungsleitfäden um einen internen Prozess der Statistischen Ämter des Bundes und der Länder handelt, sind externe Nutzerinnen und Nutzer beziehungsweise Stakeholder nicht vorgesehen.

Ausgehend von dem Ziel und den Zielgruppen bietet die Geheimhaltungs-Informationsplattform unterschiedliche Arten von Dokumenten und Informationen an: Eine zugangsbeschränkte Langfassung enthält sämtliche Angaben, in einer allen interessierten Nutzerinnen und Nutzern gleichermaßen zugänglichen Kurzfassung sind insbesondere sensitive Parameter für Geheimhaltungsverfahren ausgeblendet.

### 3.3 Benutzeroberfläche der Geheimhaltungs-Informationsplattform

Ein agiles Team aus dem Statistischen Bundesamt hat die Geheimhaltungs-Informationsplattform innerhalb von 14 Monaten nach dem SCRUM-Verfahren mit der JavaScript-Softwarebibliothek ReactJS (React JavaScript library) entwickelt und getestet.

#### Aufbau

Die grafische Benutzeroberfläche stellt die zentrale Schnittstelle zwischen Nutzenden und der Geheimhaltungs-Informationsplattform dar. Die Benutzeroberfläche der Webanwendung zur Geheimhaltungs-Informationsplattform ist nach einem einheitlichen Aufbau strukturiert: Identitätsbereich, Bereich zur Navigation, Bereich mit dem Inhalt (Content) und Fußbereich.

Im **Identitätsbereich** befinden sich das Logo der Anwendung und der Bearbeitungsbereich für das Profil des jeweils angemeldeten Benutzers oder der jeweils angemeldeten Benutzerin. Über dieses Ausgabeelement sind folgende weitere Optionen verlinkt:

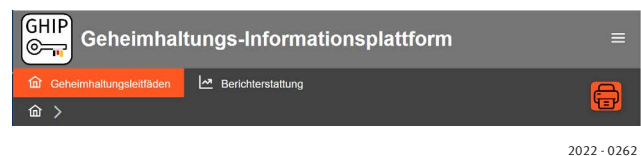
- › Config (unter anderem Direktlink in die zentrale Benutzer- und Rechteverwaltung, um das persönliche Passwort zu ändern)
- › Info (mit den allgemeinen Informationen über Bedeutung und Zweck der Geheimhaltungs-Informationsplattform sowie über die Ansprechpersonen)
- › Abmelden

In der **Hauptnavigation** findet sich der Zugang zu den Hauptdatenobjekten der Anwendung. Die Menüpunkte „Geheimhaltungsleitfäden“ und „Berichterstattung“

werden für alle angemeldeten Nutzerinnen und Nutzer gleichermaßen angezeigt. Standardmäßig wird immer für alle Benutzerinnen und Benutzer der erste Menüpunkt mit der Statistikübersicht nach der Anmeldung geöffnet, da hier zumindest lesende Zugriffe für alle Benutzerinnen und Benutzer bestehen. ➤ **Grafik 1**

**Grafik 1**

Hauptnavigation



Der Aufbau des **Content-Bereichs** unterscheidet zwischen dem Einstiegsdialog (Verzeichnis) und den Bearbeitungsdialogen der zu bearbeitenden Datenobjekte.

Im Verzeichnis werden Datenobjekte entsprechend der jeweiligen Berechtigung aufgelistet, die anhand der Filter näher eingegrenzt werden können. Die Darstellung erfolgt in tabellarischer Form, wobei über die Tabellenköpfe eine Filterung und Umsortierung der aktuellen Verzeichniseinträge möglich ist. ➤ **Grafik 2**

Die Auswahl von Funktionen zu den einzelnen Datenobjekten erfolgt grundsätzlich objektspezifisch über entsprechende Schaltflächen am rechten Rand einer Objektzeile in der Verzeichnisliste.

Im unteren Teil des Statistikbereichs sind die statistikspezifischen Fragebogen aufgelistet. ➤ **Grafik 3**

**Grafik 2**  
Verzeichnis

| <div> <div>&lt;&lt; &lt; 1 2 3 4 5 &gt;&gt; &gt;&gt;</div> <div>50</div> </div> |                                 |             |                 |                  |          |
|---------------------------------------------------------------------------------|---------------------------------|-------------|-----------------|------------------|----------|
| EVAS                                                                            | Name                            | Berichtsweg | Statistikart    | Statistikcluster | Aktionen |
| suche nach EVAS                                                                 | suche nach Name                 | all         |                 |                  |          |
| 11111                                                                           | Feeststellung des Gebietsstands | dezentral   | Bundesstatistik | Klassifikationen |          |
| 11911                                                                           | Gemeindeverzeichnis             | dezentral   | Bundesstatistik | Klassifikationen |          |
| 12111                                                                           | Volkszählung                    | dezentral   | Bundesstatistik | Zensen           |          |
| 12112                                                                           | Testverfahren zur Volkszählung  | dezentral   | Bundesstatistik | Zensen           |          |

2022 - 0263

## Grafik 3

### Statistikbereich

| Nr. | Kürzel | Bundesland              | Status      | Stand                | Aktionen |
|-----|--------|-------------------------|-------------|----------------------|----------|
| 00  | DE     | Statistisches Bundesamt | freigegeben | 12.04.2022, 08:32:27 |          |
| 01  | SH     | Schleswig-Holstein      | freigegeben | 12.04.2022, 08:34:29 |          |
| 02  | HH     | Hamburg                 | freigegeben | 28.04.2022, 12:00:21 |          |
| 03  | NI     | Niedersachsen           | freigegeben | 12.04.2022, 08:34:30 |          |

2022 - 0264

## Fachliche Funktionalitäten

Für jede Statistik gibt es genau einen statistikspezifischen Geheimhaltungsleitfaden (sofern zur Statistik Angaben zur Geheimhaltung in der Geheimhaltungs-Informationsplattform abgelegt wurden). Er enthält alle verfügbaren Angaben der Statistischen Ämter des Bundes und der Länder aus den entsprechend freigegebenen Fragebogen. Zusätzlich gibt es zu jeder Statistik organisationsspezifische Geheimhaltungsleitfäden, die nur die Angaben des entsprechenden Statistikamtes enthalten. Diese sind bei der Leseansicht des zugehörigen Fragebogens einsehbar.

Grundlage für die Generierung eines statistikspezifischen Geheimhaltungsleitfadens sind die Angaben aus den Fragebogen der Statistischen Ämter des Bundes und der Länder, die vom System tabellarisch in Form einer zum Download angebotenen PDF-Datei aufbereitet werden.

Der Zugang zu den statistikspezifischen Geheimhaltungsleitfäden (und den zugrunde liegenden Fragebogen) erfolgt über ein Statistikverzeichnis, das alle zentralen und dezentralen Bundesstatistiken auflistet,

und dessen Inhalt unter verschiedenen Kriterien sortiert und gefiltert werden kann.<sup>16</sup>

Im Statistikverzeichnis werden alle verfügbaren Bundesstatistiken untereinander aufgelistet (Quelle: Metadatenmanagement beziehungsweise Statistikdatenbank).

Im oberen Teil des statistikspezifischen Geheimhaltungsleitfaden-Bereichs befinden sich EVAS-Nummer und Name der Statistik. Ebenso gibt es Möglichkeiten zum Abruf der wichtigsten geheimhaltungsrelevanten Statistikinformationen:

- › Berichtsweg: zentral/dezentral (entsprechend der Angabe in der Statistikdatenbank)
- › Statistikart: Bundesstatistik (entsprechend der Angabe in der Statistikdatenbank)
- › Zuständigkeit beim Statistischen Bundesamt: zuständige Gruppe im Statistischen Bundesamt (entsprechend der Angabe in der Statistikdatenbank)

<sup>16</sup> Es handelt sich hierbei um das Einheitliche Verzeichnis aller Statistiken des Bundes und der Länder (EVAS).

- › Cluster: zugeordneter Statistikcluster (entsprechend der Angabe in der Statistikdatenbank)
- › Geheimhaltungsbedarf: ja/nein
- › Umstellungsprozess: nicht relevant/läuft/spätere (Neu-)Konzeption
- › Umstellungsstand: Angabe der Stufe zur Umstellung (Stufe I, Stufe II, Stufe III, Stufe IV beziehungsweise „Ohne Geheimhaltungs-Bedarf“ und „Prozess abgeschlossen“).
- › Geheimhaltungsverfahren: verwendetes Geheimhaltungsverfahren mit möglichen Werten:
  - MIX: Zellspernung und Datenveränderndes Geheimhaltungsverfahren
  - ZS: Zellspernung mit sekundärer Zellspernung
  - PZS: Zellspernung ohne sekundäre Zellspernung
  - DV (SÜ): Datenveränderndes Geheimhaltungsverfahren mit Stochastischer (Zufalls-)Überlagerung
  - DV (R): Datenveränderndes Geheimhaltungsverfahren mit Rundung von Ergebnissen
  - DV: sonstiges Datenveränderndes Geheimhaltungsverfahren
  - Keine GH: in der Statistik wird kein Geheimhaltungsverfahren benötigt
  - Unbekannt: Geheimhaltungsbedarf besteht, aber ein Verfahren wurde noch nicht festgelegt
- › koordinierte Geheimhaltung: Angabe, ob ein koordiniertes Vorgehen stattfindet
- › automatisierte Geheimhaltung: Angabe, ob eine automatisierte Geheimhaltung umgesetzt wurde
- › Fertigstellung: sofern eine Umstellung angedacht ist, wird hier der vorgesehene Monat mit Jahr angegeben

➤ Grafik 4

Grafik 4

#### Statistikinformationen

×

### Infos

|                                      |                          |
|--------------------------------------|--------------------------|
| <b>Berichtsweg:</b>                  | dezentral                |
| <b>Statistikart:</b>                 | Bundesstatistik          |
| <b>Zuständigkeit beim StBA:</b>      | F 22                     |
| <b>Cluster:</b>                      | Klassifikationen         |
| <b>Geheimhaltungsbedarf:</b>         | Ja                       |
| <b>Umstellungsprozess:</b>           | spätere (Neu-)Konzeption |
| <b>Umstellungsstand:</b>             | Stufe_I                  |
| <b>Geheimhaltungsverfahren:</b>      |                          |
| <b>koordinierte Geheimhaltung:</b>   | Nein                     |
| <b>automatisierte Geheimhaltung:</b> | Nein                     |
| <b>Fertigstellung:</b>               | 1.2022                   |

2022 - 0265

Während die ersten vier Angaben aus dem Metadatenmanagement (beziehungsweise der Statistikdatenbank) übernommen werden, berechnet das System die übrigen Werte aus den Angaben des Statistischen Bundesamtes.

### Anmeldung und Nutzung

Die Anmeldung setzt eine Registrierung für die Anwendung in der zentralen Benutzer- und Rechteverwaltung im Verbund der Statistischen Ämter des Bundes und der Länder voraus. Die Registrierung wird durch die Administration der zentralen Benutzer- und Rechteverwaltung im Statistischen Bundesamt vorgenommen.

Um die Vertraulichkeit und Integrität der in der Geheimhaltungs-Informationsplattform vorgehaltenen Daten sicherzustellen, kann jede Akteurin beziehungsweise jeder Akteur nur diejenigen Funktionalitäten aufrufen, die für ihre beziehungsweise seine Aufgabenerledigung zwingend notwendig sind. Der Aufruf von Funktionalitäten setzt daher eine entsprechende Identifikation gegenüber der Geheimhaltungs-Informationsplattform voraus beziehungsweise eine Zuweisung einer oder mehrerer entsprechender Benutzergruppen in der zentralen Benutzer- und Rechteverwaltung. Das Rechtekonzept des Geheimhaltungs-Informationsplattform-Moduls basiert auf sechs unterschiedlichen Benutzergruppen.

Eine Benutzerin oder ein Benutzer muss mindestens eine, kann aber auch mehrere Benutzergruppen innehaben. Die Geheimhaltungs-Informationsplattform ist also

anonymen Nutzenden, die nicht wenigstens einen Lese-Zugang besitzen, nicht zugänglich. Eine Benutzergruppe (ZBRADMINISTRATOR) dient der Pflege der Rechteinformationen. Alle anderen Benutzergruppen benötigen Aufgaben im direkten Zusammenhang mit der Anwendung der Geheimhaltungs-Informationsplattform.

Die Speicherung der erfassten Informationen erfolgt im zentralen Metadatenmanagementsystem des Statistischen Verbunds. Über spezifische, in einer Nutzerverwaltung vergebene Zugangsrechte, kann das Metadatenmanagementsystem bei Bedarf diese Metadaten für die weitere Verwendung (beispielsweise in Qualitätsberichten und Referentenbesprechungen) im jeweiligen Fachbereich in Bund und Ländern bereitstellen.

## 4

### Vorteile und Herausforderungen bei der Entwicklung einer Geheimhaltungs-Informationsplattform

---

In der Digitalen Agenda des Statistischen Verbunds bietet ein Informationssystem eine zentrale Anlaufstelle zum Thema Geheimhaltung und ermöglicht so eine schnelle, flexible und sichere Geheimhaltung.

Die Fachbereiche können ihre statistikspezifischen Geheimhaltungsleitfäden selbstständig, also ohne Abfragen oder zusätzliche Programmierung, aktualisieren. Die Leitfäden werden im Anschluss automatisiert neu generiert.

Die Eingabe von Geheimhaltungsinformationen (zum Beispiel Verfahren, Parameter und so weiter) erfolgt medienbruchfrei über eine grafische Benutzeroberfläche.

Auswertungen zur statistischen Geheimhaltung lassen sich mithilfe der Geheimhaltungs-Informationsplattform automatisiert erstellen.

Zu den mit der Entwicklung einer Geheimhaltungs-Informationsplattform verbundenen Herausforderungen zählen:

- › Die IT-Sicherheit der Geheimhaltungs-Informationsplattform sowie des Zugriffs gemäß dem Rechte- und Rollenkonzept für Mitarbeiterinnen und Mitarbeiter

der Statistischen Ämter des Bundes und der Länder ist sicherzustellen.

- › Die Umsetzung der Geheimhaltungs-Informationsplattform hängt vom weiteren Ausbau und der Verfügbarkeit der Aufbereitungs-, Auswertungs- und Metadaten-systeme ab.

## 5

### Fazit und Ausblick

---

Die Geheimhaltungs-Informationsplattform wurde als neues Modul des Metadatenmanagementsystems agil entwickelt. Die grafische Benutzeroberfläche stellt die zentrale Schnittstelle zwischen Nutzerinnen und Nutzern sowie der Geheimhaltungs-Informationsplattform dar. Sie dient der Navigation zu unterschiedlichen Komponenten der Geheimhaltungs-Informationsplattform (Übersicht der Statistiken, Übersicht und Änderung der geheimhaltungsrelevanten Metadaten zu Statistiken, Generierung der Geheimhaltungsleitfäden, Auswertungen). Die allgemeinen Informationen zu zentralen und dezentralen Bundesstatistiken werden aus bestehenden Systemen abgerufen und in der Übersicht der Statistiken angezeigt. Die Daten der Geheimhaltungs-Informationsplattform zu Statistiken entstammen mehreren Quellen: der Befragung aus den Jahren 2018/2019 sowie der Übersichtstabelle, die die aktuellen geheimhaltungsrelevanten Metadaten aus empfangenen Geheimhaltungsleitfäden, Berichtsvorlagen oder anderen schriftlichen Rückmeldungen der Fachbereiche zusammenführt. Die Metadaten können in die passenden Geheimhaltungsleitfäden exportiert werden. Unterschiedliche Arten von Auswertungen dieser Metadaten über Geheimhaltung der Statistiken sind möglich, ebenso das Speichern in unterschiedlichen Formaten. Controllingberichte zum Umsetzungsstand können auf einfache Weise generiert werden.

Darüber hinaus soll die Geheimhaltungs-Informationsplattform künftig auch statistikspezifische ebenso wie allgemein für die statistische Geheimhaltung relevante (Geheimhaltungs-)Konzepte, Richtlinien, Vorgaben und Gesetzestexte vorhalten. Um die praktische Umsetzung der Geheimhaltung zu unterstützen, soll die Plattform Geheimhaltungs-Handbücher und -Leitfäden bereitstellen, ebenso die Geheimhaltungswerkzeuge selbst sowie

ihre Dokumentation/Benutzerhandbücher und Installationsvorgaben.

Die Modernisierung der Statistikproduktion sichert der Statistische Verbund ständig durch die Weiterentwicklungen und Standardisierung sowie Optimierung der Prozesse, Werkzeuge und Methoden. Standardisierung und Gewährleistung der Geheimhaltung gehören zu den aufwendigen Prozessen im Statistischen Verbund. Die entwickelte Geheimhaltungs-Informationsplattform als Teil des Metadatenmanagementsystems ist ein Baustein im Statistikproduktionsprozess. 

### LITERATURVERZEICHNIS

---

Blumöhr, Torsten/Teichmann, Corina/Noack, Anke. *Standardisierung der Prozesse: 14 Jahre AG SteP*. In: WISTA Wirtschaft und Statistik. Ausgabe 5/2017, Seite 58 ff.

Europäische Union. *Verhaltenskodex für europäische Statistiken für die nationalen statistischen Ämter und Eurostat*. 2017. [Zugriff am 29. August 2022]. Verfügbar unter: [www.destatis.de](http://www.destatis.de)

Gießing, Sarah/Höninger, Julia/Ramic, Arijana Amina/Rohde, Johannes. *Ein Implementierungsprozess für Geheimhaltungsverfahren in vier Stufen*. In: Zeitschrift für amtliche Statistik Berlin Brandenburg. Ausgabe 2/2020, Seite 4 ff., hier: Seite 4. [Zugriff am 19. September 2022]. Verfügbar unter: [www.statistik-berlin-brandenburg.de](http://www.statistik-berlin-brandenburg.de)

Lindenstruth, Thomas/Claußen, Sven. *Metadatenmanagement als neue Integrationsarchitektur*. In: WISTA Wirtschaft und Statistik. Ausgabe 5/2017, Seite 76 ff.

Rothe, Patrick. *Statistische Geheimhaltung – Der Schutz vertraulicher Daten in der amtlichen Statistik, Teil 1: Rechtliche und methodische Grundlagen*. In: Bayern in Zahlen. Ausgabe 5/2015, Seite 294 ff. [Zugriff am 19. September 2022]. Verfügbar unter: [www.statistik-berlin-brandenburg.de](http://www.statistik-berlin-brandenburg.de)

Statistisches Bundesamt. *Geheimhaltungs-Informationsplattform (GHIP) Benutzerhandbuch*. (Internes Dokument).

Statistische Ämter des Bundes und der Länder. *Fachkonzept für eine Informationsplattform für den Teilprozess Geheimhaltung*. (Internes Dokument).

Statistische Ämter des Bundes und der Länder. *Einheitliche Geheimhaltung in allen Statistiken mit Geheimhaltungsbedarf – Vorgehenskonzept –*. (Internes Dokument).

Statistische Ämter des Bundes und der Länder. *Geschäftsprozessmodell Amtliche Statistik (GMAS)*. (Internes Dokument).

UNECE. *Generic Statistical Business Process Model*. Zuletzt aktualisiert im Juli 2022. [Zugriff am 25. August 2022]. Verfügbar unter: [statswiki.unece.org](https://statswiki.unece.org)

### RECHTSGRUNDLAGEN

---

Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) in der Fassung der Bekanntmachung vom 20. Oktober 2016 (BGBl. I Seite 2394), das zuletzt durch Artikel 2 des Gesetzes vom 14. Juni 2021 (BGBl. I Seite 1751) geändert worden ist.



**Prof. Dr. Göran Kauermann**

ist Inhaber des Lehrstuhls für Statistik und ihre Anwendungen in den Sozial- und Wirtschaftswissenschaften an der Ludwig-Maximilians-Universität München.



**Giacomo De Nicola**

hat an der Universität Florenz und in Mailand studiert und promoviert derzeit an der Ludwig-Maximilians-Universität München. Er befasst sich im Rahmen seiner Doktorarbeit mit der statistischen Analyse komplexer Netzwerke, um Fragen, die in den Wirtschaftswissenschaften, der Soziologie und der Politikwissenschaft gestellt werden, zu beantworten.

Für ihre Arbeiten „On assessing excess mortality in Germany during the COVID-19 pandemic“ (De Nicola und andere, 2022) und „An update on excess mortality in the second year of the COVID-19 pandemic in Germany“ (De Nicola/Kauermann, 2022) sind die beiden Autoren zusammen mit Prof. Dr. Michael Höhle vom Robert Koch-Institut im Jahr 2022 mit dem [Sonderpreis „Corona – Auswirkungen der Pandemie auf Wirtschaft und Gesellschaft“](#) des Statistischen Bundesamtes ausgezeichnet worden.

# ÜBERSTERBLICHKEIT DURCH CORONA?

Göran Kauermann, Giacomo De Nicola

➤ **Schlüsselwörter:** COVID-19 – Altersadjustierung – Übersterblichkeit – Sterbefallzahlen

## ZUSAMMENFASSUNG

Der Beitrag stellt zunächst die grundlegenden Ideen der mit dem Corona-Sonderpreis des Statistischen Bundesamtes ausgezeichneten Arbeiten vor, die sich mit der Schätzung der Übersterblichkeit in den Jahren 2020 und 2021 befassen. Dabei wurde eine altersadjustierte Schätzung der jährlichen Übersterblichkeit basierend auf aktuellen Sterbetafeln vorgeschlagen und gezeigt, dass sich der Anstieg der Sterbezahlen in den Jahren 2020 und 2021 größtenteils durch die alternde Bevölkerung erklären lässt. Der Beitrag präsentiert zudem aktuelle Analysen, die die bereits veröffentlichten Ergebnisse erweitern.

➤ **Keywords:** Covid-19 – age adjustment – excess mortality – death figures

## ABSTRACT

*The article first presents the basic ideas of the papers which have been awarded the Covid-19 special prize of the Federal Statistical Office and deal with estimating excess mortality in 2020 and 2021. An age adjusted estimation of annual excess mortality based on up-to-date life tables was proposed there. It was also shown that the increase in death figures in 2020 and 2021 can largely be explained by population ageing. In addition, the article presents current analyses that enrich the results published.*

## 1

### Einleitung

Das Thema Übersterblichkeit ist seit der COVID-19-Pandemie in aller Munde und während der Hochphasen der Pandemie haben das Statistische Bundesamt, aber auch zahlreiche Medien sowie Wissenschaftlerinnen und Wissenschaftler, regelmäßige Grafiken und Tabellen veröffentlicht, die wochenweise Übersterblichkeit zeigten. So präsentiert beispielsweise das Statistische Bundesamt unter [www.destatis.de](http://www.destatis.de) die aktuellen Werte für Deutschland und fügt üblicherweise die letzten vier Jahre zu Vergleichszwecken hinzu. [Grafik 1](#) zeigt eine aktuelle Version dieser Darstellung. Die Auswertungen werden auch auf [Ebene der Bundesländer](#) bereitgestellt. Zudem enthält die Webseite Verlinkungen auf in ähnlicher Form verfügbare Sterbedaten anderer Länder und zu weiteren internationalen Quellen zu Sterbefallzahlen.

Neben diesen wochenweisen Vergleichen ist die Berechnung von jahresweisen Übersterblichkeiten von Interesse. Hier steht eher der internationale Vergleich im Vordergrund: Wie haben unterschiedliche Länder in Bezug auf erhöhte Todeszahlen die COVID-19-Pandemie überstanden? Herausstechend ist hier die Arbeit von Wang und anderen (2022), die für Deutschland für die Jahre 2020 und 2021 mehr als 200 000 zusätzliche Todesfälle auflisten. Die Weltgesundheitsorganisation kommt in einer Analyse (basierend auf Knutson und

andere, 2022) zu dem Ergebnis, dass in Deutschland fast 100 000 zusätzliche Todesfälle zu verzeichnen sind. Levitt und andere (2022) vergleichen unterschiedliche Berechnungsmethoden von Übersterblichkeit. Die sich ergebenden Übersterblichkeiten für Deutschland bewegen sich demnach zwischen 50 000 und 200 000 zusätzlichen Todesfällen. Die Autoren zeigen dabei auch, dass sich die für Deutschland berechnete Übersterblichkeit beim Verfolgen einer Altersadjustierung mehr als halbiert.

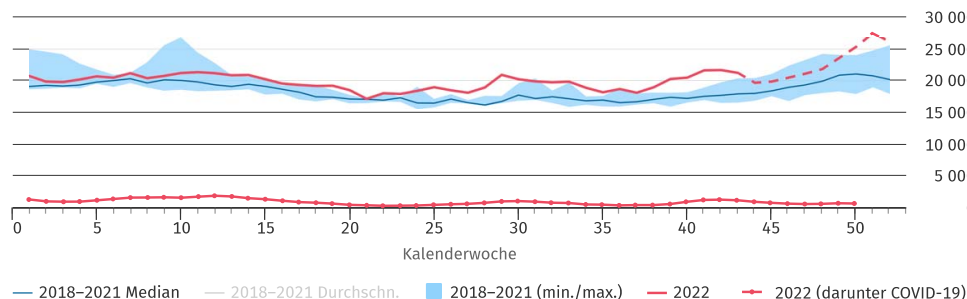
Um Übersterblichkeiten zu berechnen ist es notwendig, eine mittlere Todeszahl zu berechnen, also eine erwartete Anzahl von Todesfällen. Wird diese überschritten, spricht man von Übersterblichkeit, die dann absolut oder auch relativ angegeben werden kann. Die methodische Herausforderung liegt dabei auf der Berechnung der erwarteten Anzahl von Todesfällen. Der Ad-hoc-Ansatz ist, einfach den Mittelwert der Todeszahlen aus vorherigen Jahren als erwartete Todeszahl heranzuziehen, etwa den Mittelwert der letzten vier Jahre. Das ist plausibel, solange sich die Lebenserwartung nicht substantiell ändert und solange die Altersstruktur der Bevölkerung annähernd gleichbleibt. Auch wenn anzunehmen ist, dass die erste Bedingung noch zutrifft, so hat sich in Deutschland die Altersstruktur gerade der Hochbetagten in den letzten Jahren doch erheblich verändert. Das bedeutet aber, dass zur Berechnung der erwarteten Todeszahlen die Altersstruktur der Bevölkerung herangezogen werden muss. Genau hier setzt die Arbeit von De Nicola, Kauermann und Höhle (2022) an (ebenso

#### Grafik 1

#### Übersterblichkeiten, aufbereitet vom Statistischen Bundesamt

##### Wöchentliche Sterbefallzahlen in Deutschland

(gestrichelte Werte enthalten Schätzanteil)



Quellen: Sterbefallzahlen insgesamt: Statistisches Bundesamt (Stand 09.01.2023), COVID-19-Todesfälle: Robert Koch-Institut (Stand 05.01.2023)

© Statistisches Bundesamt (Destatis), 2023

2023 - 020

De Nicola/Kauermann, 2022). Die Arbeiten zeigen eine einfache Korrekturmethode für ein Problem. Dabei ist aus Sicht der Autoren erstaunlich, dass sich bis heute kein Standardverfahren für die Fragestellung etabliert hat. Der folgende Beitrag stellt die Notwendigkeit und die grundlegenden Ideen der Altersadjustierung kurz vor und zeigt einige aktuelle Ergebnisse.

## 2

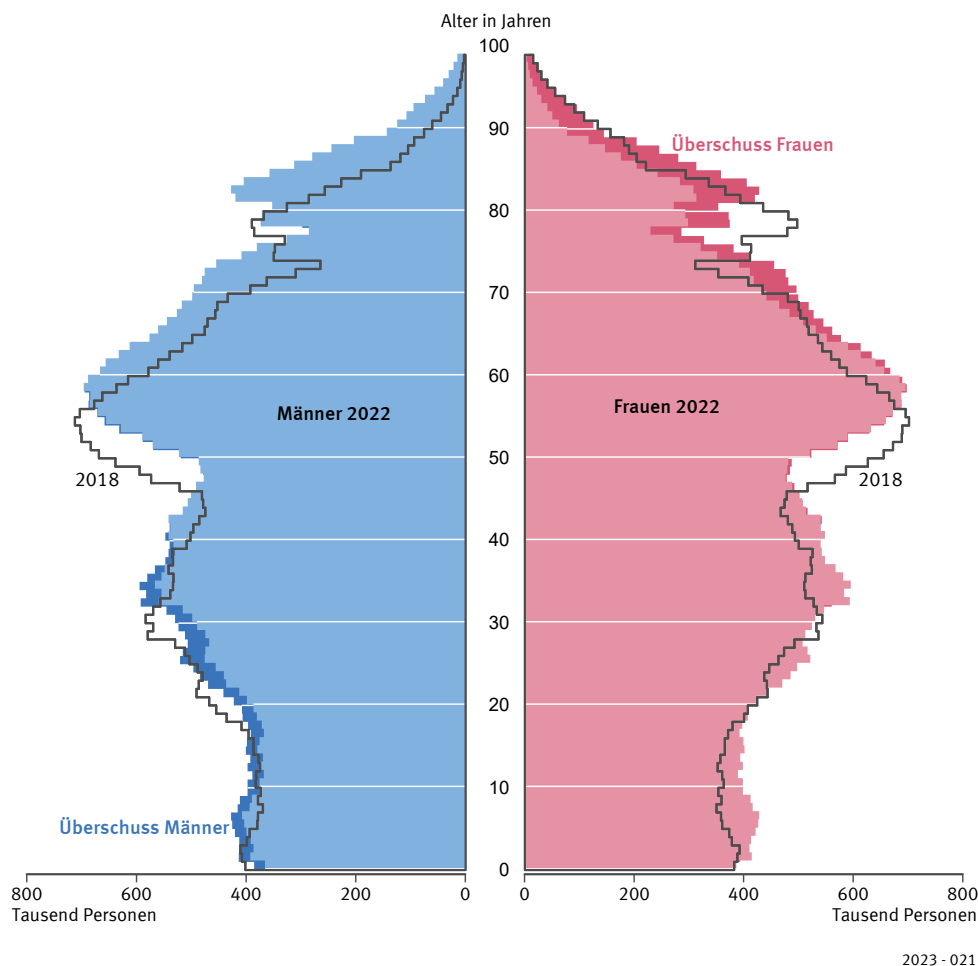
### Altersadjustierung

Das Statistische Bundesamt stellt den Altersaufbau der Bevölkerung in Deutschland in Form von Bevölkerungspyramiden zur Verfügung. Für die Altersadjustierung

werden die Jahre 2018 (als Kontur) und 2022 genutzt. [➤ Grafik 2](#) Zu erkennen ist eine Ausbuchtung bei den Jahrgängen, die am Beginn ihres neunten Lebensjahrzehnts stehen. Diese Ausbuchtung geht auf den geburtenreichen Jahrgang 1940 zurück. Durch ihn ist in den Jahren 2020 und 2021 – den untersuchten Jahren der COVID-19-Pandemie – der Anteil der über 80-Jährigen rein bedingt durch die alternde Bevölkerungsstruktur deutlich angestiegen. Da die Sterbewahrscheinlichkeiten in höherem Alter stark zunehmen, ist in den genannten Jahren in Deutschland jährlich mit etwa 2% mehr Todesfällen zu rechnen, was das Statistische Bundesamt in den Jahresergebnissen berücksichtigt. Um dies schlüssig zu quantifizieren, sind die vom Statistischen Bundesamt bereitgestellten aktuellen Sterbetafeln heranzuziehen. Hier setzt unsere Arbeit (De Nicola und

Grafik 2

Alterspyramide in Deutschland in den Jahren 2018 und 2022



andere, 2022) an: Wir schlagen eine einfache, wenn gleich robuste altersadjustierte Berechnung von erwarteten Todesfällen vor. Es zeigt sich dabei, dass ein einfacher gewichteter Mittelwert aus Bevölkerungsstruktur zum Jahreswechsel gewichtet mit den Sterbetafeln eine verzerrte erwartete Todesanzahl liefert. Kurz gesprochen liegt dies daran, dass ein zum ersten Januar eines Jahres 79-Jähriger, der im Laufe desselben Jahres verstirbt, noch mit 79, aber eben auch mit 80 Jahren versterben kann. So einfach dies klingt, so wichtig ist es, dies bei der Berechnung der erwarteten Todeszahlen zu berücksichtigen, wie unsere Arbeit zeigt. Eine einfache Mittelwertbildung der Sterbetafeln führt zu einem korrigierten Schätzer der Todeszahlen. In gleicher Weise ist eine Mittelwertbildung zu berücksichtigen bei der Berechnung von Übersterblichkeiten für unterschiedliche Altersgruppen. Auch hier zeigt die zitierte Arbeit einen einfachen, aber hinreichend exakten Ansatz (für Details siehe De Nicola und andere, 2022).

Es folgen einige reproduzierte Ergebnisse mit aktuellen Zahlen: [↗ Grafik 3](#) auf Seite 84 zeigt die berechneten Übersterblichkeiten für die Jahre 2016 bis 2022. Basierend auf De Nicola und andere (2022) werden dabei zunächst die erwarteten Todeszahlen in den genannten Jahren berechnet. Es folgt eine Aufteilung proportional auf die Wochen des Jahres, sodass es möglich ist, Saisonalität zu visualisieren. Die Abbildungen zeigen in Grau und Hellblau die (altersadjustierten) Sterblichkeiten in den Jahren vor Beginn der COVID-19-Pandemie. Der Mittelwert dieser vier Jahre ist als schwarze Kurve abgebildet. Die Jahre mit COVID-19 sind farblich rot (2020), hellrot (2021) und dunkelblau (2022) dargestellt. Deutlich sichtbar ist die Grippewelle im Jahr 2018, die zu einer stark erhöhten Sterblichkeit bei den über 80-Jährigen geführt hat. Ebenso sind die COVID-19-Wellen zum Jahreswechsel 2020/2021 und die Herbstwelle 2021 ausgeprägt zu sehen. Jenseits dessen ist die kurzfristige Übersterblichkeit im Oktober 2022 zu erwähnen, die zunächst schnell wieder abgeflacht ist. Zum Ende des Jahres 2022 steigt die Übersterblichkeit bei den über 60-Jährigen deutlich an. In der Altersgruppe der 60- bis 79-Jährigen liegt sie im Dezember 2022 sogar auf ihrem Maximum und ist vergleichbar mit den Werten der Grippewelle im Frühjahr 2018.

Im Vergleich zu Grafik 1 enthält Grafik 3 deutlich mehr Information: Zum einen sind die Daten des Statistischen Bundesamtes in Grafik 1 nicht altersadjustiert.

Das bedeutet, dass die Sterbezahlen erwartungsgemäß durch die sich ändernde Altersstruktur steigen, was Grafik 1 nicht berücksichtigt. Zum anderen zeigt Grafik 3 die Sterbezahlen für unterschiedliche Altersgruppen und lässt insbesondere für die Hochbetagten deutliche Auffälligkeiten erkennen. Dazu zählt auch die unterdurchschnittliche Sterblichkeit in den ersten Wochen des Jahres 2022. Hier verläuft die dunkelblaue Kurve des Jahres 2022 bei den Jahrgängen, die 80 Jahre und älter sind, und noch mehr bei der Altersgruppe der 60- bis 79-Jährigen deutlich über dem Durchschnitt und ist mit der Kurve des ersten COVID-19-Winters (2020/2021) vergleichbar.

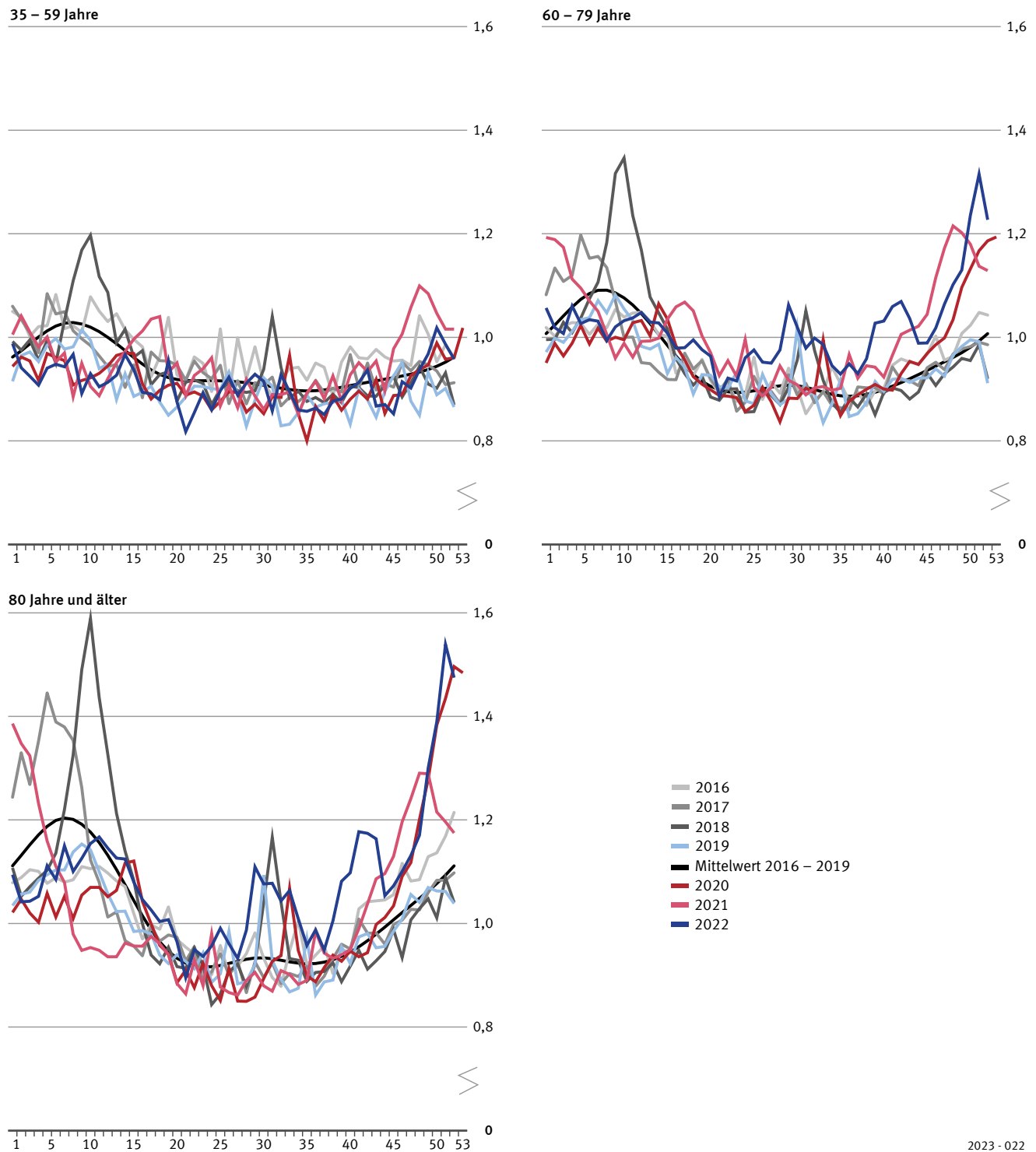
### 3

## Jahresweise Übersterblichkeiten

Außer wochenweise können Sterbezahlen über einzelne Jahre verglichen werden. Hier zeigt sich die Notwendigkeit einer Altersadjustierung noch nachdrücklicher, wie die Ergebnisse für Deutschland und im Vergleich exemplarisch für die Vereinigten Staaten für die Jahre 2016 bis 2021 verdeutlichen. [↗ Grafik 4](#) auf Seite 85 Die erwarteten Todesfälle sind als blaue Linien dargestellt. Die Berechnung dieser Kurven erfolgte anhand der Methode in De Nicola und andere (2022). Die tatsächlich beobachteten Todeszahlen sind als Punkte dargestellt. Es sind deutliche Unterschiede zwischen Deutschland und den Vereinigten Staaten zu erkennen: Die Übersterblichkeit in den Vereinigten Staaten ist um ein Vielfaches stärker ausgeprägt als in Deutschland. Interessant ist dabei auch, dass die Linie der erwarteten Todeszahlen für die Vereinigten Staaten im Jahr 2020 einen leichten Knick aufweist. Konkret weist dies darauf hin, dass im Jahr 2020 primär ältere Personen verstorben sind, sodass die erwartete Anzahl bedingt durch eine Abnahme des Anteils der Hochbetagten an der Bevölkerung abgenommen hat.

### Grafik 3

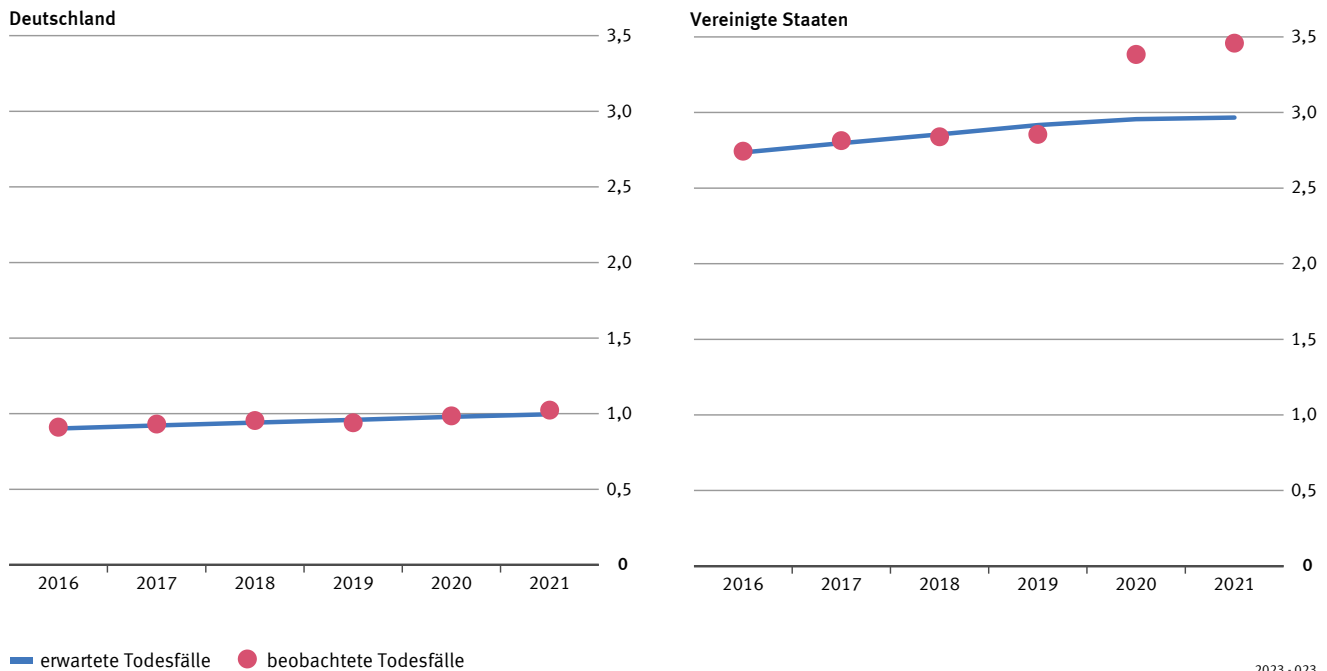
Altersadjustierte wöchentliche Sterbefallzahlen der Jahre 2016 bis 2022 in Deutschland nach Kalenderwochen und Altersgruppen



2023 - 022

## Grafik 4

Jahresweise Übersterblichkeiten in Deutschland und den Vereinigten Staaten  
Erwartete und beobachtete jährliche Todesfälle, in Mill.



2023 - 023

## 4

### Diskussion

Die Diskussion über die durch COVID-19 verursachten Todesfälle geht weiter. In diesem Beitrag wurde dafür die Notwendigkeit einer Altersadjustierung deutlich herausgearbeitet. Reine Zahlenvergleiche zwischen Jahren, zwischen einzelnen Ländern, ja selbst zwischen einzelnen Bundesländern können zu verzerrten Ergebnissen führen. Dies gilt insbesondere, wenn die Alterspyramide

keine „klassische“ Form hat, also geburtenstarke Jahrgänge alt werden. Dieser Effekt tritt insbesondere in europäischen Ländern auf, bedingt durch die weiterhin spürbaren Auswirkungen des Zweiten Weltkriegs auf den Altersaufbau der Bevölkerung. Zu sehen ist aber auch, dass die Übersterblichkeiten in Deutschland – sofern altersadjustiert – sich nicht abhoben von denen in anderen Ländern, wie in diesem Beitrag am Beispiel der Vereinigten Staaten gezeigt wurde. Dies gibt Anlass zu weiterer Forschung, auch um zu untersuchen, ob die getroffenen Maßnahmen wie Kontaktbeschränkungen einen Effekt auf Übersterblichkeit hatten. [\[1\]](#)

## LITERATURVERZEICHNIS

---

De Nicola, Giacomo/Kauermann, Göran/Höhle, Michael. *On assessing excess mortality in Germany during the COVID-19 pandemic*. In: AStA Wirtschafts- und Sozialstatistisches Archiv. Jahrgang 16. Ausgabe 1/2022, Seite 5 ff. [Zugriff am 25. Januar 2023]. Verfügbar unter: [link.springer.com](https://link.springer.com)

De Nicola, Giacomo/Kauermann, Göran. *An update on excess mortality in the second year of the COVID-19 pandemic in Germany*. In: AStA Wirtschafts- und Sozialstatistisches Archiv. Jahrgang 16. Ausgabe 1/2022, Seite 21 ff. [Zugriff am 25. Januar 2023]. Verfügbar unter: [link.springer.com](https://link.springer.com)

Knutson, Victoria/Aleshin-Guendel, Serge/Karlinsky, Ariel/Msemburi, William/Wakefield, Jon. *Estimating Global and Country-Specific Excess Mortality During the COVID-19 Pandemic*. arXiv preprint arXiv:2205.09081 (2022). [Zugriff am 25. Januar 2023]. Verfügbar unter: <https://arxiv.org/abs/2205.09081>

Levitt, Michael/Zonta, Francesco/Ioannidis, John P. A. *Comparison of pandemic excess mortality in 2020-2021 across different empirical calculations*. In: Environmental Research. Band 213: 113754. Oktober 2022. doi: 10.1016/j.envres.2022.113754. [Zugriff am 25. Januar 2023]. Verfügbar unter: [www.sciencedirect.com](https://www.sciencedirect.com)

Wang, Haidong und andere. *Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020–21*. In: The Lancet 399.10334. 2022. Seite 1513 ff. [Zugriff am 25. Januar 2023]. Verfügbar unter: [www.thelancet.com](https://www.thelancet.com)

**Herausgeber**  
Statistisches Bundesamt (Destatis), Wiesbaden

---

**Schriftleitung**  
Dr. Daniel Vorgrimler  
Redaktion: Ellen Römer

---

**Ihr Kontakt zu uns**  
[www.destatis.de/kontakt](http://www.destatis.de/kontakt)

---

**Erscheinungsfolge**  
zweimonatlich, erschienen im Februar 2023  
Ältere Ausgaben finden Sie unter [www.destatis.de](http://www.destatis.de) sowie in der [Statistischen Bibliothek](#).

---

Artikelnummer: 1010200-23001-4, ISSN 1619-2907

---

© Statistisches Bundesamt (Destatis), 2023  
Vervielfältigung und Verbreitung, auch auszugsweise, mit Quellenangabe gestattet.