

STATISTISCHE GEHEIMHALTUNG IN BEVÖLKERUNGSSTATISTIKEN

Einsatz der Fünferrundung in der Ausländerstatistik nach dem Ausländerzentralregister

Dr. Gunter Brückner, Jan Eberle, Florian Götsche

➤ **Schlüsselwörter:** Geheimhaltungsverfahren – Rundungsverfahren – Kriterienkatalog – zentrale Ausländerstatistik – Ausländerzentralregister

ZUSAMMENFASSUNG

Das Statistische Bundesamt hat für das Berichtsjahr 2016 in der Ausländerstatistik ein Rundungsverfahren für die statistische Geheimhaltung eingeführt. Dieser Artikel erläutert die fachliche Abwägung der Vor- und Nachteile des Rundungsverfahrens vor dem Hintergrund der statistikspezifischen Anforderungen. Es zeigt sich, dass das gewählte Geheimhaltungsverfahren Vorteile durch die einfache Implementierung, die große Flexibilität bei Ad-hoc-Auswertungen und seine intuitive Verständlichkeit hat. Nachteile sind die nicht mehr vorhandene Additivität der Ergebnisse und ein zumindest theoretisch weiter existierendes Aufdeckungsrisiko in besonderen Fällen. Auch die Tatsache, dass die Rundung alle Ergebnisse verändert, unabhängig davon, ob eine Geheimhaltung notwendig wäre oder nicht, ist kritisch zu sehen. Dennoch hat sich bisher die Einführung des Verfahrens bewährt.

➤ **Keywords:** *disclosure control method – rounding method – catalogue of criteria – central statistics of foreigners – Central Register of Foreigners*

ABSTRACT

The Federal Statistical Office of Germany introduced a rounding method to ensure statistical confidentiality in the statistics of foreigners for reference year 2016. This article provides a comparison of the advantages and disadvantages of the rounding method in the light of the specific requirements of these statistics. The advantages of the selected disclosure control method are its easy implementation, high degree of flexibility regarding ad hoc analyses and intuitive comprehensibility. Disadvantages are the lost additivity of the results and, at least theoretically, a remaining risk of disclosure in specific cases. Furthermore the fact that all values are modified due to rounding, irrespective of whether statistical confidentiality would be required, must be assessed critically. Nevertheless, the experience gained has so far justified the introduction of the rounding method.

Dr. Gunter Brückner

hat Volks- und Betriebswirtschaftslehre an den Universitäten Gießen und Manhattan (Kansas) studiert und in einem Mikrosimulationsmodell die Beeinflussung der Einkommens- und Vermögensverteilung untersucht. Er leitet das Referat „Ausländer- und Integrationsstatistiken“ des Statistischen Bundesamtes. Er hat dort unter anderem das Konzept der Bevölkerung mit Migrationshintergrund im Mikrozensus entwickelt und etabliert und die Datengrundlage des Projekts „Beschäftigte mit Migrationshintergrund in der Öffentlichen Verwaltung des Bundes“ betreut.

Jan Eberle

ist studierter Volkswirt (M. Sc.) und Referent im Referat „Ausländer- und Integrationsstatistiken“ des Statistischen Bundesamtes. Zu seinen Arbeitsschwerpunkten gehören Auswertungen des Ausländerzentralregisters zu den Themengebieten Ausländerinnen und Ausländer sowie Schutzsuchende.

Florian Götsche

studierte Soziologie an der Universität Mannheim und ist Referent im Referat „Ausländer- und Integrationsstatistiken“ des Statistischen Bundesamtes. Seine Arbeitsschwerpunkte sind die Berechnung des Migrationshintergrunds im Mikrozensus, die Berechnung der Integrationsindikatoren sowie die Betreuung der Einbürgerungsstatistik.

1

Einleitung

Das Statistische Bundesamt veröffentlicht mit der zentralen Ausländerstatistik jährlich Zahlen zur ausländischen Bevölkerung, die jeweils zum Jahresende im Ausländerzentralregister (AZR) erfasst ist. Dafür erhält das Statistische Bundesamt nach § 23 des Gesetzes über das Ausländerzentralregister (AZR-Gesetz) jeweils zum 31. Dezember einen Auszug des Registers. Im Ausländerzentralregister sind personenbezogene Daten aller Ausländerinnen und Ausländer gespeichert, die sich nicht nur vorübergehend, sondern drei Monate oder länger in Deutschland aufhalten. Als eines der größten Register in der öffentlichen Verwaltung in Deutschland dient das AZR allen mit asyl- und aufenthaltsrechtlichen Aufgaben betrauten öffentlichen Stellen als zentrale Datenbank. Neben der registerführenden Behörde, dem Bundesamt für Migration und Flüchtlinge, zählen hierzu maßgeblich die Ausländerbehörden der Länder. Die Daten des AZR werden im Statistischen Bundesamt aufbereitet und anschließend veröffentlicht.

Bei der Veröffentlichung der Ausländerstatistik 2016 hat das Statistische Bundesamt das Verfahren der Fünferrundung angewendet, um die Vorgaben der statistischen Geheimhaltung und des Personendatenschutzes umzusetzen. Dieser Artikel erläutert zunächst, welche Abwägungen zur Auswahl dieses Geheimhaltungsverfahrens führten. Anhand der bisher gesammelten Erfahrungen werden die Vor- und Nachteile der Fünferrundung sowie deren praktische Implementierung dargestellt. Die Evaluierung folgt dabei einem Katalog von Bewertungskriterien für Geheimhaltungsmethoden, der zusammen mit den Statistischen Ämtern der Länder erarbeitet wurde. Ein weiterer Aufsatz in dieser Ausgabe von WISTA stellt den Kriterienkatalog vor (Rohde und andere, 2018). Dieser Kriterienkatalog dient den Fachstatistiken als Rahmen zur Auswahl geeigneter Geheimhaltungsverfahren. Eine Priorisierung der Kriterien in Muss- und Soll-Kriterien wird dabei ausdrücklich nicht vorgegeben. Der folgende Artikel erläutert die fachliche Abwägung der Vor- und Nachteile des Verfahrens der Fünferrundung in der Ausländerstatistik vor dem Hintergrund der statistikspezifischen Anforderungen und liefert damit ein Anwendungsbeispiel dieses Kriterienkatalogs.

1.1 Grundlagen der statistischen Geheimhaltung

Nach § 16 Absatz 1 Bundesstatistikgesetz (BStatG) besteht eine allgemeine Pflicht zur Geheimhaltung von Einzelangaben zum Schutz vor Offenlegung der persönlichen und sachlichen Verhältnisse der Auskunftspflichtigen. Ausnahmen hiervon bestehen unter anderem dann, wenn die Einzelangaben von den Statistischen Ämtern mit den Einzelangaben anderer Befragter zusammengefasst veröffentlicht werden (siehe § 16 Absatz 1 Nr. 3 BStatG) oder den Befragten oder Betroffenen nicht mehr zuzuordnen sind (siehe § 16 Absatz 1 Nr. 4 BStatG). Im Rahmen der statistikspezifischen Operationalisierung dieser gesetzlichen Vorgaben werden in der amtlichen Statistik üblicherweise bestimmte Typen von Ergebnissen als potenziell geheimhaltungsrelevant betrachtet (Gießing und andere, 2018; hier: Kapitel 2.2.1). Hierzu zählen Ergebnisse, die

- › auf den Angaben einer Person beruhen (Eigenschaften dieser Person sind erkennbar),
- › auf den Angaben zweier Personen beruhen (Eigenschaften einer der beiden Personen können von der zweiten Person aus dem Tabellenergebnis „herausgerechnet“ werden),
- › mit der entsprechenden Spalten- oder Zeilensumme übereinstimmen, weil alle Betroffenen die gleiche Eigenschaft aufweisen (wenn beispielsweise alle 25-jährigen Albanerinnen eine Duldung besitzen, dann ist durch die Statistik der Aufenthaltsstatus jeder einzelnen 25-jährigen Albanerin erkennbar).

Man bezeichnet Verfahren, die die – nach den für die jeweilige Statistik getroffenen Festlegungen – geheimhaltungsrelevanten Ergebnisse beziehungsweise Tabellenfelder identifizieren und entsprechend schützen, als **primäre Geheimhaltung**. Als **sekundäre Geheimhaltung** werden alle Schritte bezeichnet, die sicherstellen, dass ein geheim gehaltener Wert nicht aus dem Kontext ermittelt werden kann (sogenannte Gegensperrung). Das folgende Beispiel erläutert den Unterschied: In [Tabelle 1](#) sei der Inhalt der Zelle C1 geheim gehalten. Die Tabelle 1 erfüllt jedoch die Anforderungen der Geheimhaltung nicht, weil der nicht dargestellte Wert der Zelle C1 „1“ beträgt, wie man durch Differenzbildung leicht ermitteln kann. Die Geheimhaltung in Zelle C1 lässt sich dadurch rückgängig machen. Zur Sicherstellung der sekundären

Tabelle 1

Darstellung mit primärer Geheimhaltung

| | Spalte A | Spalte B | Spalte C | Spalte D | Insgesamt |
|-----------|----------|----------|----------|----------|-----------|
| Zeile 1 | 4 | 6 | . | 8 | 19 |
| Zeile 2 | 10 | 12 | 14 | 19 | 55 |
| Zeile 3 | 8 | 7 | 5 | 4 | 24 |
| Zeile 4 | 12 | 9 | 5 | 11 | 37 |
| Insgesamt | 34 | 34 | 25 | 42 | 135 |

Geheimhaltung ist es vielmehr nötig, den Inhalt weiterer drei Zellen in der Tabelle geheim zu halten, die nach den Anforderungen der primären Geheimhaltung eigentlich nicht geheimhaltungsbedürftig sind. Das können die Teilsummen sein, in die der geheim zu haltende Wert eingeht. [↪ Tabelle 2](#) Wenn der mit der Geheimhaltung von

Tabelle 2

Darstellung mit primärer Geheimhaltung und sekundärer Geheimhaltung der Summen

| | Spalte A | Spalte B | Spalte C | Spalte D | Insgesamt |
|-----------|----------|----------|----------|----------|-----------|
| Zeile 1 | 4 | 6 | . | 8 | . |
| Zeile 2 | 10 | 12 | 14 | 19 | 55 |
| Zeile 3 | 8 | 7 | 5 | 4 | 24 |
| Zeile 4 | 12 | 9 | 5 | 11 | 37 |
| Insgesamt | 34 | 34 | . | 42 | . |

Zeilen-, Spalten- und Tabellensummen einhergehende Informationsverlust als zu groß erscheint, dann lassen sich ersatzweise drei andere, eigentlich nicht geheimhaltungsbedürftige Zellen sperren, sodass die Summen unverändert ausgegeben werden können. [↪ Tabelle 3](#)

Tabelle 3

Darstellung mit primärer Geheimhaltung und sekundärer Geheimhaltung unbeteiligter Zellen

| | Spalte A | Spalte B | Spalte C | Spalte D | Insgesamt |
|-----------|----------|----------|----------|----------|-----------|
| Zeile 1 | 4 | 6 | . | . | 19 |
| Zeile 2 | 10 | 12 | 14 | 19 | 55 |
| Zeile 3 | 8 | 7 | . | . | 24 |
| Zeile 4 | 12 | 9 | 5 | 11 | 37 |
| Insgesamt | 34 | 34 | 25 | 42 | 135 |

Die Anforderung der sekundären Geheimhaltung erstreckt sich nicht nur auf reine Fallzahlen, sondern darüber hinaus auch auf abgeleitete Größen, wie absolute Abweichungen oder Salden, sowie auf Anteilswerte, Quoten, Verhältniszahlen oder Veränderungsraten. Alle abgeleiteten Größen sind geheimhaltungsbedürftig, um

zu verhindern, dass aus ihnen ein geheim gehaltener Wert indirekt ableitbar wird. Demnach sind Differenzen und Salden geheim zu halten, wenn ein Summand beziehungsweise Subtrahend geheimhaltungsbedürftig ist. Anteilswerte und Veränderungsraten sind geheim zu halten, sofern Zähler und/oder Nenner geheimhaltungsbedürftig sind.

1.2 Die verfügbaren Geheimhaltungsverfahren

Für Fallzahltabellen, wie sie im Rahmen der Ausländerstatistik veröffentlicht werden, diskutieren Fachleute eine Vielzahl alternativer Vorgehensweisen zur Umsetzung der statistischen Geheimhaltung. Die Verfahren lassen sich grundsätzlich in folgende Kategorien einteilen:

- › **Verfahren der Vergrößerung und Zellspernung** identifizieren zunächst Tabellenzellen, die primär geheimhaltungsbedürftig sind. Anschließend wird geprüft, ob die Ausweisung dieser Zellen durch Reduzierung des Detaillierungsgrads vermieden werden kann. Alternativ werden die entsprechenden Zellen gesperrt, wobei die sekundäre Geheimhaltung gegebenenfalls weitere Zellspernungen verlangt („gegenpunkten“, siehe Tabelle 3).
- › **Post-tabular datenverändernde Verfahren** (Gießing und andere, 2018; hier: Kapitel 2.5) verändern die in den Tabellenzellen angegebenen Werte. Hierbei kann weiter unterschieden werden, ob diese Veränderung deterministisch oder zufällig („stochastisch“) erfolgt. Die Fünfferrundung zählt zu den deterministischen Verfahren, bei ihr werden alle Tabellenzellen auf das nächste Vielfache von Fünf gerundet. Primär geheimhaltungsbedürftige Fallzahlen (0, 1 und 2) werden hierbei alle auf den Wert „null“ gerundet. Zur Sicherstellung der sekundären Geheimhaltung werden datenverändernde Verfahren auf alle Tabellenzellen angewendet. Demnach werden auch Tabellenzellen verändert, die nicht primär geheimhaltungsbedürftig sind.
- › **Prä-tabular datenverändernde Verfahren** (Gießing und andere, 2018; hier: Kapitel 2.6) modifizieren die Merkmalsausprägungen auf Ebene des Datenmaterials so, dass für alle zur Tabellierung vorgesehenen Merkmalskombinationen gleichzeitig die Anforderungen der Geheimhaltung erfüllt werden. Dadurch erfüllen alle mit dem veränderten Material erstellten Tabellen die

Geheimhaltungsanforderungen und enthalten tabelleübergreifend konsistente Angaben. Alle Merkmalskombinationen zukünftiger Auswertungen müssen von Anfang an vollständig einbezogen werden, um diese bei der Durchführung des Verfahrens berücksichtigen zu können.

Für die Ausländerstatistik war das Statistische Bundesamt an einer kurzfristigen Umsetzung der Geheimhaltung interessiert. Dadurch war die Auswahl potenziell nutzbarer Verfahren stark eingeschränkt. Der Einsatz eines prä-tabular datenverändernden Verfahrens schied aufgrund des erwarteten hohen Implementierungsaufwandes aus. Geheimhaltung durch Vergrößerung und Zellspernung ist in umfangreichen Publikationen fehleranfällig und erfordert aufwendige Koordination zwischen Bund und Ländern, um die Geheimhaltung auch tabelleübergreifend gewährleisten zu können. Für die Anwendung stochastisch datenverändernder Verfahren (zum Beispiel stochastische Überlagerung) standen bislang keine IT-Werkzeuge für eine dezentrale Anwendung zur Verfügung. Die Geheimhaltung hätte damit nur an zentraler Stelle und nur für ein festgelegtes Programm an Auswertungstabellen angewandt werden können. Die damit einhergehende fehlende Flexibilität war angesichts der Vielzahl der nicht vorhersehbaren Sonderauswertungen in der Ausländerstatistik de facto ein Ausschlusskriterium.

In Anbetracht der gegebenen Restriktionen fiel trotz gewisser Nachteile im Hinblick auf die Sicherstellung der sekundären Geheimhaltung (siehe Abschnitt 2.1) beziehungsweise Genauigkeit (siehe Abschnitt 2.2) die Entscheidung letztlich im Ausschlussverfahren zugunsten der Fünferrundung. Im Folgenden wird der bisherige Einsatz der Fünferrundung in der Ausländerstatistik anhand der Kriterien

- › Gewährleistung der Statistischen Geheimhaltung (Schutz),
- › Informationsverlust (Qualität),
- › Praktikabilität und
- › Wirtschaftlichkeit

bewertet.

2

Evaluierung der Fünferrundung in der Ausländerstatistik

2.1 Gewährleistung der Geheimhaltung

Die Erfüllung der Anforderungen des § 16 Bundesstatistikgesetz wird anhand von vier Kriterien bewertet:

Sicherstellung der primären Geheimhaltung

Durch die Rundung entsteht eine generelle, aber beschränkte Unsicherheit bezüglich der tatsächlichen Fallzahlen. Letztlich weisen Tabellen nach der Rundung keine sensiblen Fallzahlen (0, 1 oder 2) mehr auf, da diese Tabellenzellen gesperrt werden.

Sicherstellung der sekundären Geheimhaltung

Durch die konsequente Veränderung aller Fallzahlen werden Rückschlüsse durch Differenzbildung verhindert. Aufdeckungsrisiken bestehen allerdings in Zusammenhang mit bestimmten Rundungsmustern. Bei Anwendung der Fünferrundung kann beispielsweise das gerundete Muster $5 + 5 + 5 + 5 = 10$ nur aus den tatsächlichen Fallzahlen $3 + 3 + 3 + 3 = 12$ resultieren. Das Auftreten solcher Muster könnte im Prinzip in einem aufwendigen Verfahren automatisiert geprüft werden. Mit dem Auftreten einer solchen Konstellation¹ ist zwar grundsätzlich ein Enthüllungsrisiko verbunden. Das Risiko, dass eine dadurch theoretisch mögliche Enthüllung tatsächlich zur Identifizierung einer Person und Aufdeckung ihrer Angaben führt, wird in der Ausländerstatistik aber als zu gering eingeschätzt, um diesen Aufwand zu rechtfertigen.

Einheitliche Anwendung über das gesamte Auswertungsprogramm

Eine nicht einheitliche Geheimhaltung birgt die Gefahr, dass durch Zuhilfenahme verschiedener Veröffentlichungen geheimhaltungsbedürftige Ergebnisse rückrechnen-

¹ Im Vergleich zu einer entsprechenden stochastischen Überlagerung treten aufdeckbare Konstellationen bei Fünferrundung verhältnismäßig häufig auf. Der Einsatz eines stochastischen Verfahrens anstelle der Fünferrundung würde die Enthüllungsrisiken deutlich reduzieren. Bei prä-tabularen Verfahren treten in dieser Weise aufdeckbare Konstellationen sogar grundsätzlich überhaupt nicht auf (Gießing und andere, 2018; hier: Kapitel 2.5.1.2).

bar werden. Durch die vergleichsweise unkomplizierte praktische Implementierung in Excel und SAS erfolgt die Geheimhaltung der Ausländerstatistik einheitlich bei allen Veröffentlichungen, Sonderauswertungen und Datenbankangeboten des Statistischen Bundesamtes und der Statistischen Ämter der Länder.

Geheimhaltung bei Nutzung von Einzeldaten durch Dritte

Die Einzeldaten des Ausländerzentralregisters dürfen nach § 23 AZR-Gesetz durch das Statistische Bundesamt ausschließlich an die Statistischen Ämter der Länder weitergegeben und ausgewertet werden. Eine Nutzung der Daten durch Dritte außerhalb des Statistischen Verbunds von Bund und Ländern findet demnach nicht statt.

2.2 Informationsverlust

Die Gewährleistung der Statistischen Geheimhaltung ist eine notwendige, aber noch keine hinreichende Bedingung bei der Entscheidung für ein geeignetes Verfahren. Grundsätzlich muss darüber hinaus der Umfang des Schutzes mit dem Anspruch einer bestmöglichen Qualität der statistischen Produkte abgewogen werden. Das Ausmaß des Informationsverlustes durch die Geheimhaltung wird anhand von sechs Kriterien bewertet:

Relevanz

Dieses Bewertungskriterium verlangt, die Nutzbarkeit der Daten durch das angewendete Verfahren nicht derart zu beeinträchtigen, dass die Statistik für die Nutzerinnen und Nutzer an Relevanz verliert. Das Ausländerzentralregister ist eine administrative Datenquelle, in der die lokalen Datenbestände der Ausländerbehörden zentral zusammengeführt werden. Durch die Zuordnung der Datensätze zu den jeweils zuständigen Ausländerbehörden ermöglicht es Auswertungen zur ausländischen Bevölkerung in regional tiefer Gliederung (mit wenigen Ausnahmen bis auf die Kreisebene). Bei aktuell über 10 Millionen registrierten Ausländerinnen und Ausländern treten seltene Ereignisse auch auf der Kreisebene nur bei gering besetzten Staatsangehörigkeiten und/oder einzelnen Altersjahren auf. Das Statistische Bundesamt rät Kundinnen und Kunden bei entsprechenden Datenanfragen aber generell von der Interpretation

solch seltener Ereignisse im AZR ab, da die Qualität des Registers keine hinreichende Genauigkeit gewährleistet.

Qualitätsverluste treten nur dann auf, wenn das Geheimhaltungsverfahren die Nutzbarkeit der Daten über die bereits bestehenden Einschränkungen hinaus beeinträchtigt. Die Implementierung der Fünferrundung in der Ausländerstatistik führt zu keinen zusätzlichen Einschränkungen des standardmäßigen Auswertungsprogramms und es müssen keine Anpassungen bei der Gliederungstiefe vorgenommen werden. Kundinnen und Kunden finden weiterhin alle Auswertungen in der Fachveröffentlichung und in der Verbreitungsdatenbank GENESIS-Online in dem gewohnten Tabellenaufbau.

Genauigkeit

Ein geeignetes Geheimhaltungsverfahren sollte einen minimalinvasiven Eingriff darstellen, das heißt die Genauigkeit der Ergebnisse insoweit erhalten, dass falsche Rückschlüsse verhindert werden. Durch das Geheimhaltungsverfahren auftretende Genauigkeitsverluste können anhand geeigneter Kennzahlen bewertet werden. Bei der Fünferrundung beträgt die maximale absolute Abweichung zwischen gerundeten und originalen Fallzahlen genau 2. In einer exemplarischen Auswertung über 394 Kreise liegt die mittlere absolute Abweichung über alle Tabellenzellen bei 1,23. Die empirische Bleibewahrscheinlichkeit liegt in der Beispielauswertung bei 19%². Anteilswerte, die aufgrund der Rundung stark verzerrt sind, werden nicht ausgegeben (siehe Abschnitt 3.2).

Höhere Genauigkeitsverluste treten auf, wenn Nutzerinnen und Nutzer gerundete Fallzahlen eigenständig summieren (siehe Kriterium Additivität). In der Ausländerstatistik wird in allen Veröffentlichungen auf diese Problematik hingewiesen. Über GENESIS-Online steht außerdem eine Vielzahl von relevanten Zwischensummen (zum Beispiel unterschiedliche Ländergruppierungen) zur Verfügung, um eine eigenständige Summenbildung zu vermeiden.

2 Die empirische Bleibewahrscheinlichkeit gibt den Anteil der Tabellenzellen an, die durch das Geheimhaltungsverfahren nicht verändert wurden. Bei der Fünferrundung werden alle Fallzahlen, bei denen die letzte Ziffer eine 0 oder eine 5 ist, nicht verändert. Demgegenüber bleiben bei einer stochastischen Überlagerung Fallzahlen oberhalb von etwa 5 mit deutlich höheren Bleibewahrscheinlichkeiten erhalten (je nach Parametrisierung: 40 bis 80 %).

Additivität

Datenverändernde Geheimhaltungsverfahren, wie die deterministische Rundung und die stochastische Überlagerung, führen zu nicht additiven Ergebnissen. Die Akzeptanz von nicht additiven Ergebnissen muss daher antizipiert und bewertet werden. Nach Rundung kommt es zu Differenzen zwischen der Summe der gerundeten Fallzahlen und der gerundeten Randsumme. Bilden Nutzerinnen und Nutzer selbstständig Summen aus gerundeten Fallzahlen, kann es außerdem zu größeren Genauigkeitsverlusten kommen.

In der exemplarischen Auswertung über 394 Kreise ist die theoretisch maximale Abweichung zwischen der Summe der gerundeten Fallzahlen und der gerundeten Summe gleich 788 ($394 \cdot 2$). Bei einer großen Anzahl an Summanden sollten sich Auf- und Abrundungen im Erwartungswert jedoch ausgleichen. In der exemplarischen Auswertung nach Kreisen liegt die beobachtete Differenz zwischen gerundeter Randsumme und der Summe der gerundeten Einzelwerte letztlich bei 10 (10 039 090 – 10 039 080). [↗ Tabelle 4](#)

Tabelle 4

Abweichungen gerundeter und nicht gerundeter Fallzahlen nach Kreisen

| | Fallzahl | | |
|------------------------|------------|------------|------------|
| | gerundet | original | Abweichung |
| 01001 (Flensburg) | 11 235 | 11 235 | – |
| 01002 (Kiel) | 26 885 | 26 884 | 1 |
| ... | ... | ... | ... |
| Summe über alle Kreise | 10 039 090 | 10 039 080 | 10 |

Der Genauigkeitsverlust vergrößert sich mit der Zahl der Summanden. Zur Veranschaulichung wird die exemplarische Auswertung der ausländischen Bevölkerung auf der Kreisebene um eine Klassifizierung nach 20 Altersgruppen erweitert. Bilden Nutzerinnen und Nutzer auf

Tabelle 5

Abweichungen gerundeter und nicht gerundeter Fallzahlen nach Kreisen

| | 0 bis unter 5 Jahren | | | ... | 95 Jahre und älter | | | Summe über alle Alterskategorien | | |
|------------------------|----------------------|----------|------------|-----|--------------------|----------|------------|----------------------------------|------------|------------|
| | gerundet | original | Abweichung | | gerundet | original | Abweichung | gerundet | original | Abweichung |
| 01001 (Flensburg) | 385 | 384 | 1 | ... | 10 | 8 | 2 | 11 235 | 11 235 | – |
| 01002 (Kiel) | 1 175 | 1 174 | 1 | ... | 20 | 21 | 1 | 26 875 | 26 884 | 9 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| Summe über alle Kreise | 405 945 | 405 979 | 34 | ... | 5 260 | 5 355 | 95 | 10 039 025 | 10 039 080 | 55 |

Basis der gerundeten Zwischensummen für jeden Kreis die Summe der ausländischen Bevölkerung, kommt es zu einer maximalen Abweichung von 55 Fällen (10 039 025 – 10 039 080). Die mittlere absolute Abweichung (1,20) und die empirische Bleibewahrscheinlichkeit (21 %) bleiben weitgehend unverändert. Bei den Zwischensummen über alle Kreise in den Altersgruppen zeigt sich der größte Genauigkeitsverlust mit einer absoluten Abweichung von 95 (5 355 – 5 260) in der Gruppe mit der niedrigsten Fallzahl (95 Jahre und älter). [↗ Tabelle 5](#)

Die maximale Verzerrung tritt auf, wenn Nutzerinnen und Nutzer die gerundeten Fallzahlen für alle einzelnen Staatsangehörigkeiten über alle Einzelaltersjahre aufsummieren. Hierbei steigt die Gefahr von Genauigkeitsverlusten mit sinkender Fallzahl bei der Staatsangehörigkeit. Innerhalb der seltenen Staatsangehörigkeiten kommt es zu größeren Genauigkeitsverlusten, wenn sich die ausländischen Personen auf viele Altersjahre verteilen. In diesem Fall treten vermehrt Tabellenzellen mit nur einem oder zwei Fällen auf, die abgerundet werden (zum Beispiel São Tomé und Príncipe). Konzentrieren sich die Ausländerinnen und Ausländer auf nur wenige Altersjahre, fällt der Genauigkeitsverlust geringer aus.

[↗ Tabelle 6, Tabelle 7](#)

Tabelle 6

Abweichungen gerundeter und nicht gerundeter Fallzahlen nach Staatsangehörigkeit

| | Summe über alle Altersjahre | | Abweichung | Relative Abweichung |
|---------------------------------------|-----------------------------|------------|------------|---------------------|
| | gerundet | original | | |
| | Anzahl | % | | |
| São Tomé und Príncipe | 0 | 21 | 21 | 100 |
| Äquatorialguinea | 95 | 131 | 36 | 27 |
| Botsuana | 100 | 118 | 18 | 15 |
| ... | ... | ... | ... | ... |
| Summe über alle Staatsangehörigkeiten | 10 037 035 | 10 039 080 | 2 045 | 0,02 |

Tabelle 7

Relative Abweichung nach Gesamtfallzahl der jeweiligen Staatsangehörigkeit

| | Relative Abweichung | | |
|--|---------------------|-----|--------------|
| | max | min | Durchschnitt |
| | % | | |
| Fallzahlen von 1 492 000 bis 35 000 (1. Quintil) | 0,1 | 0,0 | 0,0 |
| Fallzahlen kleiner 35 000 bis 9 000 (2. Quintil) | 0,2 | 0,0 | 0,1 |
| Fallzahlen kleiner 9 000 bis 1 600 (3. Quintil) | 1,3 | 0,0 | 0,4 |
| Fallzahlen kleiner 1 600 bis 160 (4. Quintil) | 17,1 | 0,0 | 3,4 |
| Fallzahlen kleiner 160 bis 1 (5. Quintil) | 100 ¹ | 7,3 | 68,6 |

¹ Zu einer relativen Abweichung von 100% kommt es bei den 18 Staatsangehörigkeiten mit den geringsten Fallzahlen (alle mit Fallzahlen kleiner 3).

In den Veröffentlichungen des Statistischen Bundesamtes zur Ausländerstatistik werden die Nutzerinnen und Nutzer auf die Möglichkeit von Genauigkeitsverlusten durch eigenständige Summierungen hingewiesen. Sowohl die Fachserienveröffentlichung als auch das Datenangebot in GENESIS-Online sind mit entsprechenden Hinweisen versehen.

Aktualität

Im Hinblick auf die Aktualität darf der Anspruch einer termingetreuen Veröffentlichung der Ausländerstatistik durch das angewendete Geheimhaltungsverfahren nicht dauerhaft gefährdet werden. Die erstmalige Implementierung der Fünferrundung in der Ausländerstatistik hat zwangsläufig zu Verzögerungen bei der Veröffentlichung der Fachserie geführt. Sowohl für die Erstellung der Fachserie als auch für die Befüllung von GENESIS-Online und für die Bearbeitung von Sonderauswertungen konnten bereits in der Implementierungsphase vollständig automatisierte Verfahren etabliert werden. Mit Aktualitätsverlusten bei Veröffentlichungen ist aufgrund der Durchführung der Fünferrundung daher künftig nicht mehr zu rechnen.

Vergleichbarkeit

Geheimhaltungsverfahren können Auswirkungen auf die externe Vergleichbarkeit (zum Beispiel andere Veröffentlichungen nach dem AZR) oder die intertemporale Vergleichbarkeit (zum Beispiel Zeitreihenbrüche) haben. In der Ausländerstatistik liegen Einzeldaten des AZR seit

dem Jahr 1998 vor. Zeitreihenbetrachtungen gehören zu den Standarddatenangeboten und werden darüber hinaus häufig angefragt. Die Rundung auf ein Vielfaches von Fünf führt in der Ausländerstatistik zu keinen relevanten Auswirkungen auf die Gliederungstiefe oder auf Zeitreihenvergleiche.

Konsistenz

Durch das Geheimhaltungsverfahren sollen in der Ausländerstatistik keine widersprüchlichen Angaben in unterschiedlichen Auswertungen entstehen. Logisch identische Angaben sollen über alle Auswertungen hinweg auch konsistente Ergebnisse liefern. In der Ausländerstatistik müssen im gesamten Statistischen Verbund bei einer Vielzahl von Veröffentlichungen und über das Standarddatenangebot hinausgehenden Auswertungen konsistente Ergebnisse produziert werden. Durch die einheitliche Anwendung des Verfahrens der Fünferrundung ist sichergestellt, dass die in den Statistischen Landesämtern ermittelten Ergebnisse den veröffentlichten Zahlen des Statistischen Bundesamtes entsprechen. Das Statistische Bundesamt stellt mithilfe von Anleitungen und IT-Werkzeugen eine einheitliche Umsetzung der Geheimhaltung in der Ausländerstatistik im Statistischen Verbund mit geringem Aufwand und ohne zusätzliche Koordination sicher.

2.3 Praktikabilität

Komplexität und Transparenz

Statistische Produkte müssen für deren Nutzerinnen und Nutzer verständlich sein. Demnach müssen auch die Auswirkungen des Geheimhaltungsverfahrens auf eine Statistik nachvollziehbar sein. Die Nachvollziehbarkeit des Verfahrens spielt unter dem Gesichtspunkt der Praktikabilität des Geheimhaltungsverfahrens für die Ausländerstatistik eine zentrale Rolle. Wegen des großen Interesses einer breiten Öffentlichkeit an Zahlen aus dem AZR zur ausländischen Bevölkerung und zu Schutzsuchenden³ muss das angewendete

³ Als Schutzsuchende bezeichnet das Statistische Bundesamt Ausländerinnen und Ausländer, die sich unter Berufung auf humanitäre Gründe in Deutschland aufhalten. Darunter befinden sich unter anderem Personen im laufenden Asylverfahren (Asylbewerber/-innen), anerkannte Flüchtlinge nach der Genfer Konvention, Personen mit eingeschränktem Schutzstatus (subsidiär Schutzberechtigte) und Personen, deren Asylantrag abgelehnt wurde.

Geheimhaltungsverfahren für alle Nutzerinnen und Nutzer transparent nachvollziehbar sein. Komplexere Geheimhaltungsverfahren (zum Beispiel SAFE⁴ oder auch eine stochastische Überlagerung⁵), die gegenüber der Fünferrundung schwerer nachvollzogen werden können, bergen das Risiko, dass die Öffentlichkeit die Ergebnisse ablehnt, weil das Vorgehen für sie nicht nachvollziehbar ist. Gelingt die Kommunikation des Verfahrens gegenüber den Nutzerinnen und Nutzern nicht, besteht womöglich sogar die Gefahr einer Fehlinterpretation als willkürliche Veränderung.

Die Zahlen zu Ausländerinnen und Ausländern sowie zu Schutzsuchenden in Deutschland sollen nicht allein deshalb angreifbar sein, weil durch das Geheimhaltungsverfahren deren Werte auf für Nutzerinnen und Nutzer nicht einfach nachvollziehbare Weise verändert wurden. Ein nachvollziehbares Geheimhaltungsverfahren sorgt für mehr Transparenz und stellt deshalb eine vertrauensbildende Maßnahme gegenüber kritischen Nutzerinnen und Nutzern dar.

Replizierbarkeit

Analysen auf einer identischen Datenbasis sollen zu identischen Ergebnissen führen. Durch die Rundung erfolgt eine Veränderung der Daten erst nach der Tabellierung. Der originale Datensatz bleibt vom Geheimhaltungsverfahren unverändert. Alle Auswertungen im Rahmen der Ausländerstatistik sind damit jederzeit replizierbar.

Anwendungsbreite des Verfahrens

Ein geeignetes Verfahren zur Geheimhaltung soll sich für alle relevanten statistischen Auswertungen eignen. Für die Auswertungen im Rahmen der Ausländerstatistik werden hauptsächlich Fallzahlen, Anteile und Durchschnittswerte aus den Einzeldaten ausgewertet. Bei der Auswertung von Einzeldaten sind die durch Fünferrundung auftretenden Genauigkeitsverluste kontrollierbar. Bilden Nutzerinnen und Nutzer aus den Daten eigenständig Summen, Mittelwerte oder Varianzen auf Basis

geheim gehaltener Tabellen, dann kann es zu größeren Genauigkeitsverlusten kommen.

Flexibilität bei Veröffentlichungen

Ein geeignetes Verfahren zur Geheimhaltung in der Ausländerstatistik muss ohne größeren Aufwand flexibel für alle Auswertungs- und Veröffentlichungsbedarfe anwendbar sein. Aufgrund des großen und vielfältigen Bedarfs an unterschiedlichsten Auswertungen ist eine größtmögliche Flexibilität bei Veröffentlichungen ein zentrales Kriterium. Die Anwendung eines Verfahrens, das nicht zentral auf Einzeldatenebene erfolgt, erlaubt eine leichtere Nutzung von historischen Datenbeständen für Auswertungen, die nicht im Vorhinein antizipiert wurden (zum Beispiel Auswertungen nach dem Schutzstatus). Regelmäßige Änderungen im Aufenthalts- und Asylrecht erschweren weiterhin die Festlegung auf ein längerfristig feststehendes Tabellenprogramm.

Für die praktische Anwendung von datenverändernden Verfahren müssen IT-Werkzeuge bereitstehen. Diese sollten sowohl in die bestehenden Programmstrukturen zur Erstellung von Standarddatenangeboten eingebunden, als auch flexibel bei Ad-hoc-Auswertungen zur Beantwortung spezieller Kundenanfragen angewandt werden können.

Datenbanken und Web-Zugriff

Nutzerinnen und Nutzer sollen zunehmend flexibel auf die Daten von Statistiken zugreifen können. Datenbanken wie GENESIS-Online spielen hierbei eine wichtige Rolle. Ein geeignetes Geheimhaltungsverfahren soll demnach auf Datenbankangebote anwendbar sein. Das Angebot eines individualisierten Tabellenabrufs über GENESIS-Online wird von den Kundinnen und Kunden der Ausländerstatistik intensiv genutzt. Außerdem verweist der Zentrale Auskunftsdienst des Statistischen Bundesamtes bei entsprechenden Anfragen häufig auf dieses Angebot. Mit Blick auf den für individuelle Kundenanfragen anfallenden Aufwand ist es ausschlaggebend, dass dieses Angebot uneingeschränkt aufrechterhalten wird. Eine Anpassung des Datenangebotes in GENESIS-Online musste durch die Implementierung der Fünferrundung nicht erfolgen. In den Prozess zur Befüllung von GENESIS-Online konnte die Rundung ohne größeren Aufwand integriert werden. In einem zusätzlichen Bearbeitungs-

4 Das Geheimhaltungsverfahren SAFE kam bei der Geheimhaltung der Ergebnisse des Zensus 2011 zum Einsatz. Dieses Verfahren verändert in einem aufwendigen algorithmischen Verfahren Häufigkeiten von Merkmalsausprägungen.

5 Für die Geheimhaltung der Ergebnisse des Zensus 2021 vorgesehenes Verfahren (Kleber/Gießing, 2018).

schrift werden die Daten gerundet und Fallzahlen von 0, 1 und 2 gesperrt. Die Daten werden anschließend wie gewohnt in die Quader importiert. Die automatische Berechnung von Randsummen wird unterdrückt, um vor allem bei Auswertungen nach gering besetzten Staatsangehörigkeiten oder einzelnen Altersjahren größeren Genauigkeitsverlusten vorzubeugen (siehe Kriterium Nicht-Additivität). Die nötigen Randsummen müssen bei diesem Verfahren zusätzlich gespeichert werden.

Wissenschaftliche Nutzung

Eine Weitergabe der Einzeldaten zur wissenschaftlichen Nutzung findet nicht statt.

2.4 Wirtschaftlichkeit

Implementierungsaufwand

Für die Ausländerstatistik war das Statistische Bundesamt an einer kurzfristigen Umsetzung der Geheimhaltung interessiert. Dadurch war die Zahl potenziell nutzbarer Verfahren eingeschränkt. Die Ausländerstatistik ist eine Zentralstatistik, bei der die gesamte Aufbereitung beim Statistischen Bundesamt stattfindet. Das Statistische Bundesamt ist jedoch berechtigt, den Statistischen Landesämtern die jeweiligen landesspezifischen Daten für eigene Auswertungen bereitzustellen. Daher musste auch der Implementierungsaufwand in den Statistischen Landesämtern in die Überlegungen einbezogen werden. Die Implementierung des Rundungsverfahrens wurde dazu genutzt, die Ausländerstatistik vollständig zu automatisieren. Alle notwendigen Arbeitsschritte – von der Berechnung von Anteilswerten über die Rundung bis zur Tabellierung – werden in SAS⁶ automatisch durchgeführt. Die Statistischen Landesämter erhalten ab dem Berichtsjahr 2017 daher ein bereits der Geheimhaltung unterzogenes veröffentlichungsfähiges Tabellenprogramm für ihr jeweiliges Land, dessen Aufbau identisch mit der Fachserienveröffentlichung (Statistisches Bundesamt, 2018) ist. Zusätzlich erhalten sie diese Fachserie mit ungerundeten Werten für interne Berechnungen.

⁶ SAS ist eine Software zur Datenaufbereitung und -analyse, die das amerikanische Unternehmen SAS Institute entwickelt und vertreibt.

Laufender Aufwand

Durch die vollständig automatisierte Implementierung der Fünferrundung ist der laufende Aufwand für das Erstellen von Standardprodukten minimal. Jedoch steigt der laufende Aufwand bei Ad-hoc-Auswertungen durch die fehlende Additivität von Fallzahlen geringfügig an. So müssen in einigen Anfragen zusätzliche Zwischensummen gebildet werden (siehe Kriterium Additivität). Der zusätzliche Aufwand bleibt aber begrenzt, da bereits in der Implementierungsphase in den verwendeten Auswertungssystemen automatisierte Arbeitsabläufe etabliert werden konnten.

Akzeptanzprobleme bei Kundinnen und Kunden wurden bislang nicht beobachtet. Weder der Zentrale Auskunftsdienst noch das für die Erstellung der Ausländerstatistik im Statistischen Bundesamt zuständige Fachreferat haben bisher einen erhöhten Beratungsaufwand festgestellt. Eine Evaluierung des laufenden Aufwands in den Statistischen Landesämtern bleibt abzuwarten.

Potenzial für standardisierte IT-Lösungen zur Durchführung der Geheimhaltung

Für die Fachveröffentlichungen zur Ausländerstatistik und für die Befüllung der Datenbank GENESIS wurde die Fünferrundung bereits vollständig automatisiert. Das Verfahren ist relativ einfach durchführbar und bietet daher ein großes Potenzial für standardisierte IT-Lösungen. Auch für Ad-hoc-Auswertungen gibt es einfache und flexible Möglichkeiten zur Standardisierung, die teils in SAS, teils in VBA⁷ programmiert und als Add-Ins (Zusatzprogramme) verbreitet werden können.

Nutzbarkeit des Potenzials der Informations- und Kommunikationstechnologie bei der Datenverbreitung

Die Verbreitung von Daten mittels moderner Informations- und Kommunikationstechnologie soll Nutzerinnen und Nutzern eine individualisierte Auswertung des Datenangebots gewährleisten. Hierbei ist zu unterscheiden, ob vorgehaltene Tabellen bedarfsorientiert abgerufen werden können (wie zum Beispiel in GENESIS-Online) oder ob Nutzerinnen und Nutzern

⁷ Visual Basic for Applications (VBA) ist eine Programmiersprache, die zur Steuerung von Abläufen in Microsoft Office Produkten wie Word oder Excel genutzt wird.

über Web-Applikationen die Möglichkeit gegeben wird, eigenständig Auswertungen des Einzelmaterials durchzuführen.

In den Prozess zur Befüllung von GENESIS-Online konnte die Rundung ohne größeren Aufwand integriert werden (siehe Kriterium Datenbanken und Web-Zugriff).

Prinzipiell sind auch Webanwendungen vorstellbar, bei denen individuelle Tabellen durch Nutzerinnen und Nutzer selbst erstellt werden können und bei denen die Geheimhaltung mittels der Rundung vor der Ausgabe der Tabellen automatisch umgesetzt wird.

3

Umsetzung der Fünferrundung in der Praxis

3.1 Umsetzung der Geheimhaltung für Fallzahlen

Durch die Rundung aller Tabellenzellen auf ein Vielfaches von 5 werden alle geheimhaltungsbedürftigen Fallzahlen (0, 1, 2) zusammengefasst (Rundung zu 0) und können einheitlich zum Beispiel mit einem Punkt („.“) gesperrt werden. Durch die konsequente Rundung aller Fallzahlen ist außerdem die sekundäre Geheimhaltung – bis auf die in Abschnitt 2.1 erwähnten, in Kauf genommenen Risiken – gewährleistet.

Die dezentrale Anwendung des Geheimhaltungsverfahrens in allen Auswertungen in den Statistischen Landesämtern und dem Statistischen Bundesamt wird durch den geringen Implementierungsaufwand sowohl in SAS als auch in Excel erleichtert.

Umsetzung in SAS

Mit der Umstellung von SAS auf die Version 9.4 besteht die Möglichkeit, benutzerdefinierte Funktionen in SAS-Formate einzubauen. Zuvor konnten nur numerische oder alphanumerische Konstanten in Formaten verwendet werden. Demnach kann eine Rundungsfunktion definiert werden, die anschließend einem Format übergeben wird. Dieses Format kann bei Auswertungen, zum Beispiel Proc Tabulate, auf alle Fallzahlen angewendet werden. [↗ Grafik 1](#)

Grafik 1

Schritt 1: Definieren der Rundungsfunktion

```
proc fcmp outlib=work.functions.round;  
  function round5(value);  
    x = (round(value, 5));  
    if x = 0 then x = .;  
    return(x);  
  endsub;  
run;
```

Schritt 2: Übergabe der Rundungsfunktion an ein Format

```
proc format;  
  value round (multilabel notsorted)  
    low-high =[round5()];  
run;
```

Schritt 3: Anwendung des Formats in der Auswertung

```
proc tabulate data=azr.azr_schutzsuchende;  
  
  class schutzstatus / mlf order=data preloadfmt;  
  class kreise / mlf order=data preloadfmt;  
  
  table kreise, schutzstatus * F=round. / printmiss miss-  
    text=‘.’;  
  
run;
```

2018 - 01 - 0500

Umsetzung in Microsoft Excel mit der Funktion VRUNDEN(Zahl;5)

Für alle Auswertungssysteme, die eine solche Funktionalität nicht bereitstellen, können die Fallzahlen wie gewohnt ohne Rundung ausgegeben und anschließend in Excel gerundet werden. Hierzu kann die Funktion VRUNDEN(Zahl;5) verwendet werden, wobei nach der Rundung in einem weiteren Schritt der Wert 0 durch ein geeignetes Symbol für die Zellsperrung aus Geheimhaltungsgründen (zum Beispiel „.“) ersetzt werden muss.

Umsetzung in Microsoft Excel mit einem VBA-Makro

Die beiden Schritte Runden und Ersetzen des Wertes „0“ durch ein Symbol für geheim gehaltene Werte können in einem VBA-Makro automatisiert werden. Der folgende VBA-Code erlaubt Nutzerinnen und Nutzern, beide Schritte gemeinsam auf einen zuvor markierten Bereich anzuwenden. Bei häufiger Anwendung kann das Makro auch einer Schaltfläche im Excel-Menüband zugewiesen werden. [↗ Grafik 2](#)

Grafik 2

```

Sub Fünfferrundung()
    Dim cell As Range
    Dim selectedRange As Range
    Dim N As Byte
    Dim X As Long

    Set selectedRange = Application.Selection

    For Each cell In selectedRange.Cells
        If IsNumeric(cell.Value) = False Then
            GoTo Skip
        Else
            X = cell.Value
            End If

            If X = 0 Or X = 1 Or X = 2 Then
                cell.Value = „.“
            Else
                cell.Value = round(X / 5) * 5
            End If

        Skip:
        Next cell
    End Sub

```

2018 - 01 - 0501

3.2 Umsetzung der Geheimhaltung bei Anteilswerten

Anteilswerte müssen – ebenso wie Fallzahlen – uneingeschränkt der Geheimhaltung unterworfen werden, da aus Anteilswerten auf die zugrundeliegenden Fallzahlen zurückgeschlossen werden kann. Anteilswerte werden in der Ausländerstatistik immer auf Grundlage von geheim gehaltenen Fallzahlen errechnet. Daraus ergibt sich, dass Anteilswerte immer geheim gehalten werden, wenn ihnen ein geheim zu haltender Nenner oder Zähler zugrunde liegt.

In der Praxis wird die Umsetzung der Geheimhaltung bei Anteilswerten dadurch erschwert, dass sich die Rundung direkt auf die Berechnung des Anteilswertes auswirkt. Bei großen Fallzahlen ist die Differenz zwischen Anteilswerten basierend auf gerundeten oder nicht gerundeten Fallzahlen zu vernachlässigen. Bei kleinen Fallzahlen

Tabelle 8

Anteile bei gerundeten und nicht gerundeten Fallzahlen

| | Insgesamt | Darunter: | |
|--------------|-----------|-------------|--------------------|
| | | verheiratet | Anteil verheiratet |
| Ohne Rundung | 6 | 3 | 50% |
| Mit Rundung | 5 | 5 | 100% |

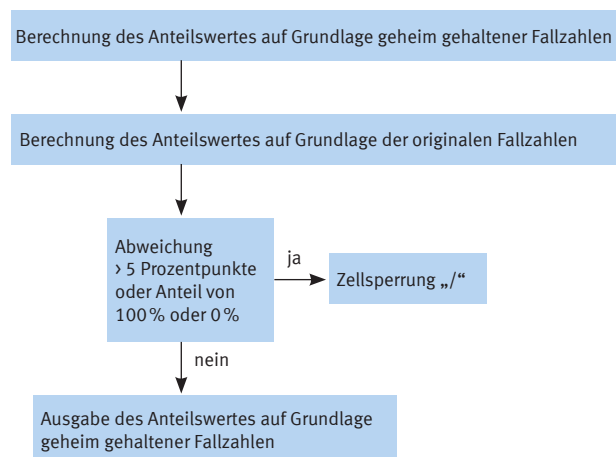
hingegen kann es zu größeren Verzerrungen kommen.

➔ **Tabelle 8**

In den Publikationen und Sonderauswertungen zur Ausländerstatistik werden regelmäßig Anteile nachgefragt und veröffentlicht. Bei der praktischen Anwendung der Geheimhaltung mittels Fünfferrundung muss demnach sichergestellt werden, dass keine zu stark verzerrten Anteilswerte veröffentlicht werden. Anteilswerte mit einer rundungsbedingten Verzerrung von mehr als 5 Prozentpunkten werden nicht ausgegeben, sondern mit Hinweis auf die fehlende Belastbarkeit der Angabe gesperrt. Hierzu wird das in ➔ **Grafik 3** dargestellte Vorgehen angewendet.

Grafik 3

Anwendung der Geheimhaltung mittels Fünfferrundung



2018 - 01 - 0502

Das Statistische Bundesamt hat eine VBA-basierte Lösung in Excel bereitgestellt, die einen einheitlichen Umgang mit Anteilswerten im Statistischen Verbund gewährleistet. Die Vorgehensweise eignet sich gleichermaßen zur Geheimhaltung von Fallzahlen und von Anteilswerten und soll eine unkomplizierte Umsetzung der Geheimhaltung ermöglichen. Zur Berechnung von Anteilswerten müssen die zugrunde liegenden Nenner und Zähler sowohl gerundet als auch nicht gerundet vorliegen. Über eine Benutzeroberfläche (User-Form) können die entsprechenden Bereiche ausgewählt werden. Die Berechnung und Ausgabe der Anteilswerte unter Berücksichtigung der Geheimhaltungsregeln (keine Ausgabe bei zugrundeliegenden Fallzahlen kleiner „3“, Zellspernung bei Abweichung zwischen Anteilswerten basierend auf gerundeten beziehungsweise unveränderten

Fallzahlen größer als 5 Prozentpunkte und Geheimhaltung von 0% und 100% Anteilswerten) erfolgt im Anschluss automatisiert.

3.3 Umsetzung in der Verbreitungsdatenbank GENESIS-Online

Zeitgleich mit der Veröffentlichung der Fachserie zur Ausländerstatistik erhalten die Nutzerinnen und Nutzer außerdem die Möglichkeit, sich über die Verbreitungsdatenbank GENESIS-Online individuelle Informationen zur Ausländerstatistik zusammenzustellen. Ein elementares Entscheidungskriterium für die Auswahl eines Geheimhaltungsverfahrens war der Anspruch, dieses Angebot ohne Verzögerung für unsere Kundinnen und Kunden aufrechterhalten zu können.

Für die Implementierung einer Rundung in einer Verbreitungsdatenbank sind generell zwei Umsetzungsmöglichkeiten denkbar: In der ersten Variante werden die den Auswertungen zugrunde liegenden Datenquader mit den – nicht geheim gehaltenen – Fallzahlen befüllt. Die Ergebnistabellen werden automatisiert auf Grundlage dieser Fallzahlen ermittelt und abschließend automatisiert dem Geheimhaltungsverfahren unterworfen. Die technischen Voraussetzungen für diese Variante stehen in GENESIS-Online gegenwärtig nicht zur Verfügung. Daher wird in der Ausländerstatistik derzeit die zweite Variante angewendet: Die Datenquader werden mit bereits geheim gehaltenen Fallzahlen befüllt und die Randsummen werden zusätzlich in speziellen Randsummenquadern vorgehalten und von dort ausgegeben, weil eine automatisierte Berechnung aus den gerundeten Fallzahlen aufgrund der Nicht-Additivität zu ungewünschten Verzerrungen führen würde.


4

Kundenreaktionen und deren Bewertung

In allen Verbreitungsformen der Ausländerzahlen 2016 wurde darauf hingewiesen, dass diese Zahlen den Anforderungen der Geheimhaltung entsprechen, und welche Auswirkungen die Verwendung des speziellen Geheimhaltungsverfahrens auf die Eigenschaften der veröf-

fentlichten Daten hat. Man kann also davon ausgehen, dass die überwiegende Mehrheit der Nutzerinnen und Nutzer über die Änderungen informiert ist, die seit dem Berichtsjahr 2016 gelten. Deshalb sollte die Auswertung des Nutzerfeedbacks seit dem Veröffentlichungstag, das heißt im Zeitraum vom 1. Juli bis 30. September 2017, verlässlich Aufschluss darüber geben, ob und inwieweit Nutzerinnen und Nutzer Probleme mit den geheim gehaltenen Daten der Ausländerstatistik haben.

Der größte Unterschied zwischen den geheim gehaltenen und den nicht geheim gehaltenen Daten der Ausländerstatistik besteht bei der Nicht-Additivität. Bei der Fünferrundung kann die rechnerische Summe der (gerundeten) Zahl der Ausländerinnen und der Ausländer mit einer gegebenen Staatsangehörigkeit von der (gerundeten Summe) für die ausländische Bevölkerung mit dieser Staatsangehörigkeit insgesamt abweichen. Das Gleiche ist bei allen Variablen möglich, nach denen Tabellen gegliedert sind, das heißt bei Alter, Familienstand, Aufenthaltsdauer oder Aufenthaltsstatus. Der zweite Unterschied besteht darin, dass die letzte Stelle der Fallzahlen nur den Wert 0 und 5 und das Zeichen für „Wert geheim gehalten“ annehmen kann. Ansonsten stimmen die Tabellen der Berichtsjahre bis 2015 vollständig mit denen von 2016 überein; es gibt keine Unterschiede in der Zahl und Bezeichnung der Spalten oder Zeilen.

Deshalb lag die Vermutung nahe, dass Kundinnen und Kunden vor allem Fragen zur Nicht-Additivität haben oder Kritik an ihr äußern würden. Dies ist bislang jedoch nicht geschehen. Sowohl der Zentrale Auskunftsdienst des Statistischen Bundesamtes als auch die Mitarbeiterinnen und Mitarbeiter in der Fachabteilung berichten einvernehmlich, dass es keine Rückfragen zu oder Kritik an Daten der Ausländerstatistik gab, die sich mit dem Geheimhaltungsverfahren oder den geheim gehaltenen Daten in Verbindung bringen lassen. Die Kundinnen und Kunden scheinen die Daten vielmehr so anzunehmen und zu nutzen, wie sie ihnen präsentiert wurden. 

LITERATURVERZEICHNIS

Gießing, Sarah/Habla, Heike/Höninger, Julia/Hoffmeister, Rita/Merz, Franz-Josef/Richter, Alexander/Scharnhorst, Sebastian/Schmidtke, Kerstin/Spies, Lydia/Tonte, Andreas/Uhrich, Stefanie. *Handbuch zur Statistischen Geheimhaltung. Stand 12. April 2018*. Internes Dokument der Statistischen Ämter des Bundes und der Länder.

Kleber, Birgit/Gießing, Sarah. [*Geheimhaltung beim Zensus 2021*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 3/2018. Kurznachrichten, Seite 6 ff., hier: Seite 7 f.

Rohde, Johannes/Seifert, Christiane/Gießing, Sarah. [*Entscheidungskriterien für die Auswahl eines Geheimhaltungsverfahrens*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 3/2018, Seite 90 ff.

Statistisches Bundesamt. *Fachserie 1 Bevölkerung und Erwerbstätigkeit, Reihe 2 Ausländische Bevölkerung – Ergebnisse des Ausländerzentralregisters 2017*. Wiesbaden 2018. Verfügbar unter: www.destatis.de

RECHTSGRUNDLAGEN

Gesetz über das Ausländerzentralregister (AZR-Gesetz) vom 2. September 1994 (BGBl. I Seite 2265), das zuletzt durch Artikel 4 des Gesetzes vom 17. Juli 2017 (BGBl. I Seite 2615) geändert worden ist.

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 10 Absatz 5 des Gesetzes vom 30. Oktober 2017 (BGBl. I Seite 3618) geändert worden ist.

Herausgeber

Statistisches Bundesamt (Destatis), Wiesbaden

Schriftleitung

Dr. Sabine Bechtold

Redaktionsleitung: Juliane Gude

Redaktion: Ellen Römer

Ihr Kontakt zu uns

www.destatis.de/kontakt

Erscheinungsfolge

zweimonatlich, erschienen im Juni 2018

Das Archiv aller Ausgaben ab Januar 2001 finden Sie unter www.destatis.de/publikationen

Print

Einzelpreis: EUR 18,- (zzgl. Versand)

Jahresbezugspreis: EUR 108,- (zzgl. Versand)

Bestellnummer: 1010200-18003-1

ISSN 0043-6143

ISBN 978-3-8246-1070-9

Download (PDF)

Artikelnummer: 1010200-18003-4, ISSN 1619-2907

Vertriebspartner

IBRo Versandservice GmbH

Bereich Statistisches Bundesamt

Kastanienweg 1

D-18184 Roggentin

Telefon: +49 (0) 382 04 / 6 65 43

Telefax: +49 (0) 382 04 / 6 69 19

destatis@ibro.de

Papier: Metapaper Smooth, FSC-zertifiziert, klimaneutral, zu 61% aus regenerativen Energien

© Statistisches Bundesamt (Destatis), 2018

Vervielfältigung und Verbreitung, auch auszugsweise, mit Quellenangabe gestattet.