



Flächennutzungsmonitoring VI Innenentwicklung – Prognose – Datenschutz

IÖR Schriften Band 65 · 2014

ISBN: 978-3-944101-65-1

Anwendung des SAFE-Verfahrens zur Sicherung der statistischen Geheimhaltung im Zensus 2011

Sarah Giessing

Giessing, Sarah (2014): Anwendung des SAFE-Verfahrens zur Sicherung der statistischen Geheimhaltung im Zensus 2011. In: Gottthard Meinel, Ulrich Schumacher, Martin Behnisch (Hrsg.): Flächennutzungsmonitoring VI. Innenentwicklung – Prognose – Datenschutz. Berlin: Rhombos-Verlag, 2014, (IÖR-Schriften; 65), S. 261-269

Anwendung des SAFE-Verfahrens zur Sicherung der statistischen Geheimhaltung im Zensus 2011

Sarah Giessing

Zusammenfassung

Zur Gewährleistung der statistischen Geheimhaltung bei den Daten des Zensus 2011 wurde ein methodischer Wechsel von den bislang üblichen informationsreduzierenden Zellsperverfahren zu einem datenverändernden Geheimhaltungsverfahren vollzogen. Dazu wurde das im Statistischen Landesamt Berlin entwickelte Verfahren SAFE („Sichere Anonymisierung für Einzeldaten“) eingesetzt. Durch das Verfahren wird die Datenbasis so verändert, dass jede in den Originaldaten existierende Merkmalskombination in der geschützten Datenbasis mindestens dreimal oder gar nicht mehr auftritt. Diese Veränderungen werden so optimiert, dass sie sich bei allen relevanten, ausreichend zusammengefassten Ergebnissen weitgehend untereinander ausgleichen. Diese stimmen also nahezu exakt mit auf Basis der unveränderten Daten errechneten Ergebnissen überein.

Die durch Anwendung des SAFE-Verfahrens in den Ergebnissen erzeugten Abweichungen von den Originalwerten dokumentieren die Statistischen Ämter in Form von Kennziffern. Der vorliegende Beitrag gibt einige zentrale Aussagen aus der Analyse dieser Kennziffern kurz wieder. Abschnitt 2 erläutert Besonderheiten der Anwendung des SAFE-Verfahrens im Umgang mit den speziellen Datenstrukturen des Zensus 2011. Abschließend geht der Beitrag auf verschiedene im Bereich der Ergebnisdarstellung getroffene Maßnahmen ein. Das betrifft insbesondere auch die Darstellung von Verhältniszahlen. Zweck dieser Maßnahmen ist die Vermeidung von Fehlinterpretation der Daten, wenn im Einzelfall bzw. bei vorliegenden Sonderkonstellationen SAFE-Abweichungen Ergebnisse mehr als „üblich“ beeinflussen.

1 Einführung

Die Einzeldaten des Zensus 2011 unterliegen dem Statistikgeheimnis nach § 16 Bundesstatistikgesetz (BMJV 1987). Wegen der großen Bedeutung, die das Vertrauen der Befragten in die Wahrung des Datenschutzes ihrer vertraulichen Angaben hat, bezieht die amtliche Statistik die Verpflichtung zur statistischen Geheimhaltung auch auf den Schutz von in Tabellen zusammengefassten Angaben, wenn die Gefahr besteht, dass dadurch Verhältnisse eines einzelnen Auskunftspflichtigen offenbart werden.

Die klassischen Methoden zur Gewährleistung der statistischen Geheimhaltung in Tabellen beruhen auf Informationsreduktion, wie z. B. dem Unterdrücken einzelner Tabellenfelder. Dabei ist immer zu beachten, dass zusätzlich Sekundärsperungen

gesetzt werden müssen, damit unterdrückte Felder auf Grund von ebenfalls veröffentlichten Summenangaben nicht rückrechenbar sind.

Die Auswertungssysteme zur Ergebnispräsentation des Zensus 2011 eröffnen die Möglichkeit flexibler Auswertungen in vielfältigen Kombinationen. Die Risiken der Rückrechenbarkeit von gesperrten Feldern in den so generierbaren Auswertungen wären jedoch wegen der Vielzahl an Summenbeziehungen zwischen den generierbaren Tabellenfeldern nicht mehr kontrollierbar. Aus diesem Grund wurde in vielen Ländern ein methodischer Wechsel zu einem datenverändernden Geheimhaltungsverfahren vollzogen. Datenverändernde Verfahren der statistischen Geheimhaltung unterscheiden sich von informationsreduzierenden Verfahren dadurch, dass sie geheim zu haltende Informationen nicht unterdrücken, sondern durch veränderte Ergebnisse ersetzen. Den Nutzern werden der Realität sehr ähnliche Ergebnisse präsentiert. Diese sollen die Eigenschaften der Gesamtheit sehr gut widerspiegeln, gleichzeitig aber auch verhindern, dass sicher Rückschlüsse über Einzelangaben gezogen werden können.

1.1 Die Entscheidung für das Verfahren SAFE

Es gibt eine Reihe unterschiedlicher Verfahrensansätze. Verwiesen sei an dieser Stelle auf datenverändernde Geheimhaltungsmethoden im britischen, neuseeländischen und australischen Zensus 2011 (Forbes et al. 2009) sowie auf Techniken, die vom US Census Bureau zur Geheimhaltung des amerikanischen Zensus 2010 verwendet wurden (Zayatz et al. 2009).

Im deutschen Zensus 2011 wird das Verfahren SAFE (Höhne 2003) eingesetzt. Bei SAFE handelt sich um ein sogenanntes Mikroaggregationsverfahren (s. Domingo-Ferrer und Mateo-Sanz 2002). Die Datenbasis wird so verändert, dass jede in den Originaldaten existierende Merkmalskombination (bei Personen z. B. aus Alter, Geschlecht, Familienstand, Religion usw.) in der geschützten Datenbasis mindestens dreimal oder gar nicht mehr auftritt. SAFE benutzt zur Durchführung der dafür nötigen Veränderungen in den Daten einen intelligenten Optimierungs-Ansatz. Dieser sorgt dafür, dass die Veränderungen so vorgenommen werden, dass sie sich weitgehend untereinander ausgleichen. So wird erreicht, dass die Vielzahl aller für Veröffentlichungen vorgesehenen Ergebnisse mit kleinstmöglicher Abweichung gegenüber den entsprechenden Originalhäufigkeiten reproduziert wird. Gleichzeitig wird bei allen relevanten, ausreichend zusammengefassten Ergebnissen gewährleistet, dass diese mehr oder weniger exakt mit Ergebnissen übereinstimmen, die auf Basis der unveränderten Daten errechnet werden. Deshalb wurden die gegenüber einem im Rahmen einer methodischen Voruntersuchung (s. Giessing, Höhne 2010) näher geprüften alternativen Ansatz (Fraser, Wooton 2005) festgestellten Nachteile in Bezug auf die Genauigkeit der geschützten Ergebnisse als nicht so gravierend beurteilt. Letztlich ausschlaggebend für die Wahl von SAFE wa-

ren zum einen seine praktischen Vorzüge (es handelt sich um eine im Verbund der deutschen Statistikämter erprobte Software, die nicht in die Online-Anwendung der Zensus-Auswertungsdatenbank integriert werden musste). Zum anderen gehören beim geprüften Alternativverfahren geringfügige Abweichungen zwischen logisch identischen Ergebnissen zum Konzept, was als sehr nachteilig bewertet wurde. Untersuchungen zu weiteren Alternativansätzen (etwa dem Verfahren des US Census Bureau, s. o.) wären relativ aufwändig gewesen und aus Kapazitätsgründen nicht möglich.

1.2 Mit SAFE behandelte Datenbestände

Vor Anwendung des Verfahrens wurden die statistischen Einheiten zunächst in zwei Datenbestände aufgeteilt. Der eine Datenbestand umfasst alle Merkmale der statistischen Einheit Person und der andere Datenbestand umfasst alle anderen statistischen Einheiten des Zensus 2011 (Haushalte, Familien, Wohnungen und Gebäude). Die Geheimhaltung der Daten mit SAFE erfolgt nicht bei allen ausgewiesenen Tabellen. Bei Auswertungen, die aus der Haushaltsstichprobe hochgerechnet wurden, wird ein Rückschluss auf Einzelfälle bereits durch die Hochrechnung und anschließende Rundung verhindert. Die Einwohnerzahlen werden dagegen stets als unveränderter Originalwert ausgewiesen. Bei allen anderen Auswertungen werden die Daten mit SAFE geheim gehalten. Zu Verhältniszahlen sei auf Kapitel 4.2 („Verhältniszahlen“) verwiesen.

2 Wie stark verändert SAFE die Daten?

Kennzahlen zu den durch die Anwendung des SAFE-Verfahrens in den Daten erzeugte Abweichungen von den Originalwerten werden im Anhang der von den Statistischen Ämtern herausgegebenen Nutzer-Kurzinformation zu SAFE (Statistische Ämter des Bundes und der Länder 2014) gezeigt. Die zentralen Aussagen seien hier kurz wiedergegeben. Sie beziehen sich auf die Gesamtheit der Tabellenfelder der Auswertungstabellen innerhalb der beiden Datenbestände, bei denen eine Geheimhaltung mit SAFE durchgeführt wurde. Es handelt sich um ca. 214 Millionen Tabellenfelder, die Ergebnisse zur Bevölkerung ausweisen bzw. ca. 164 Millionen Tabellenfelder mit Ergebnissen zu Gebäude-, Wohnungs-, Haushalts- und Familiendaten. Die feinste Regionalebene in diesem Auswertungstabellenbestand ist die Gemeinde bzw. in Berlin und Hamburg der Stadtbezirk. Ebenfalls enthalten sind höhere Regionalebenen (Verbandsgemeinden, Kreise, Regierungsbezirke, Bundesländer und Bund). Nicht enthalten sind kleinräumigere Gliederungen, wie z. B. kirchenregionale Gebietsgliederungen oder Ortsteile.

Abbildung 1 zeigt die Verteilung der Tabellenfelder nach der durch SAFE erzeugten Abweichung von ihren Originalwerten.

Insgesamt wirkt sich die Geheimhaltung durch SAFE auf die Ergebnisse zu Gebäuden, Wohnungen, Haushalten und Familien etwas stärker aus – die mittlere absolute durch

SAFE bewirkte Veränderung der Originalhäufigkeiten beträgt hier 3,8. Bei den Bevölkerungsdaten hingegen liegen die absoluten Abweichungen im Mittel bei 2,5 Personen. Bei der Mehrheit der Tabellenfelder liegt die durch SAFE bewirkte Veränderung der Originalhäufigkeiten bei bis zu +/-2 Einheiten¹. Abweichungen von +/-3 treten bei 11 % der Tabellenfelder auf. Abweichungen von +/-7 sind schon deutlich seltener. Sie finden sich bei nur noch 4,3 % der Tabellenfelder mit Bevölkerungsergebnissen bzw. bei weniger als 15 % der Tabellenfelder mit Ergebnissen zu Gebäuden, Wohnungen, Haushalten und Familien.

Bei den Bevölkerungsdaten treten Abweichungen in den ausgewiesenen Ergebnissen von mehr als +/-12 sehr selten auf. Abweichungen in dieser Größenordnung kommen nur bei etwa 2 von 1 000 Ergebnissen vor. Bei den ausgewiesenen Ergebnissen zu Gebäuden, Wohnungen, Haushalten und Familien kommen – äußerst selten – auch Abweichungen von 20 und mehr vor. Weniger als 8 von 10 000 Ergebnissen weisen eine Abweichung dieser Größenordnung auf.

Abbildung 1 zeigt die Verteilung der Tabellenfelder nach der durch SAFE erzeugten Abweichung von ihren Originalwerten.

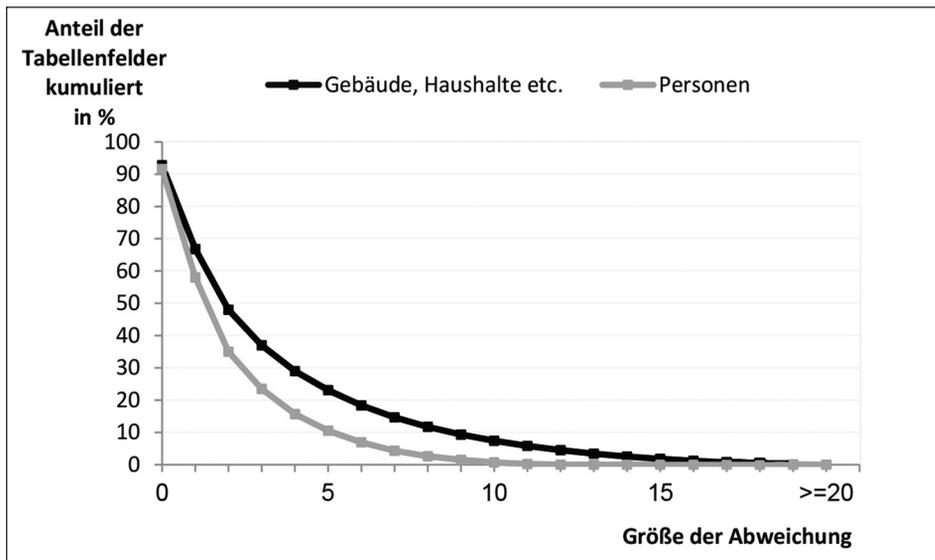


Abb. 1: Anteil der Tabellenfelder kumuliert (mit SAFE-Abweichungen von über ... in %)
(Quelle: eigene Bearbeitung der Statistiken aus Statistische Ämter des Bundes und der Länder 2014)

¹ Je nach Tabelle handelt es sich bei den „Einheiten“ um Personen, Gebäuden, Wohnungen, Haushalte oder Familien.

3 Besondere Datenstrukturen im Zensus 2011

Im Zensus 2011 werden Daten für unterschiedliche Einheiten erhoben. Es handelt sich bei diesen Einheiten entweder um Personen oder um Haushalte, um Familien, um Gebäude oder um Wohnungen. Wie eingangs erwähnt, zielt das SAFE-Verfahren darauf ab, Veränderungen in den Daten so vorzunehmen, dass sie sich weitgehend untereinander ausgleichen und Ergebnisse mit „kleinstmöglicher“ Abweichung gegenüber den entsprechenden Originalhäufigkeiten produziert werden.

Beziehungen zwischen unterschiedlichen Erhebungseinheiten (z. B. zwischen Wohnungen und Gebäuden) können dabei nur zu einem gewissen Grad berücksichtigt werden. So wird beispielsweise jeweils die beste Lösung für die Auszählungen „Gebäude mit zwei Wohnungen“ und „Wohnungen in Gebäuden mit zwei Wohnungen“ ermittelt. Während im Originaldatenbestand zwischen den betreffenden Ergebnissen normalerweise ein logischer Zusammenhang besteht (die Zahl der Wohnungen in Gebäuden mit zwei Wohnungen entspricht dem Doppelten der Zahl der betreffenden Gebäude) wird dieser in den veränderten Daten nur noch „grob“ eingehalten. Mit den veränderten Daten gerechnet werden beide Zahlen eine Abweichung von um ± 2 oder sogar etwas mehr zum Originalergebnis beinhalten. Diese Abweichung kann bei beiden Zahlen unterschiedlich groß und sogar unterschiedlich gerichtet ausfallen. Eine Änderung auf 2 Gebäude weniger und 2 Wohnungen mehr würde sich beispielsweise in der logischen Beziehung zwischen der Zahl der „Gebäude mit 2 Wohnungen“ und der „Anzahl an Wohnungen“ in diesen Gebäuden als 6 Wohnungen „zu viel“ darstellen.

Die eingangs erwähnte, vom übrigen Datenbestand getrennte SAFE-Verarbeitung des Datenbestands zur statistischen Einheit „Person“ hat zunächst zwei separate SAFE-Datenbestände zur Folge. Diese enthalten nur die jeweiligen Auswertungsmerkmale und keine Hilfsmerkmale (sog. Identifikatoren). Aus den beiden getrennten Datenbeständen lassen sich keine gemeinsamen Auswertungen vornehmen, wie beispielsweise eine Auszählung von Personen nach bestimmten Haushaltstypen.

Um dies zu ermöglichen, werden in einem Zwischenschritt die Datenbestände nach SAFE über eine Ähnlichkeitssuche eins zu eins mit den Originalbeständen zusammengeführt, um so Identifikatoren aus den Originaldaten an den SAFE-Datenbestand anzufügen. Der Identifikator, z. B. einer Original-Person, wird dabei mit Merkmalsausprägungen aus einem SAFE-Personen-Datensatz verknüpft. Diese Merkmalsausprägungen können bei allen Merkmalen mit denen der Original-Person übereinstimmen. In vielen Fällen werden jedoch in einzelnen Merkmalen gewisse Abweichungen vorliegen. Es wird so vorgegangen, dass Merkmale mit vielen Ausprägungen (z. B. Alter in Einzeljahren) zwar relativ häufig im Vergleich zu einem Merkmal wie beispielsweise „Geschlecht“ aber möglichst geringfügig (also z. B. von 33 auf 34) verändert werden.

Mithilfe der Original-Identifikatoren lassen sich zum einen Auswertungen mit nicht im SAFE-Auswertungsdatenbestand enthaltenen regionalen Gliederungen (z. B. nach Ortsteilen) erstellen. Zum anderen können dann die beiden getrennten Datenbestände nach SAFE nach dem gleichen Verfahren miteinander verknüpft werden, das auch bei den Originaldaten benutzt wird, um z. B. Personen aus dem Melderegister Wohnungen aus der Gebäude- und Wohnungszählung zuzuweisen. Dass es dabei im Einzelfall zu Kombinationen kommt, die in den echten Daten nicht vorkommen können, lässt sich nicht vermeiden. Beispielsweise kann eine Person, deren Originalalter 55 Jahre ist, mit einem Datensatz mit der Altersangabe 70 Jahre aus dem Datenbestand nach SAFE verknüpft worden sein. Die Originalperson lebt im Haushaltstyp „Haushalt ohne Senioren“. Bei einer Auszählung der Personen in diesem Haushaltstyp nach Altersklassen wird sie in der Klasse 70-74 Jahre gezählt.

4 Besonderheiten in der Ergebnisdarstellung

4.1 Ergebnisse mit größeren Abweichungen

Um Fehlinterpretationen durch Nutzer von Zensusergebnissen vorzubeugen, werden in den Ergebnistabellen Zahlen, bei denen sowohl die absolute als auch die relative Abweichung des veränderten Zahlenwerts vom Original-Zahlenwert deutlich erhöht ist, geklammert ausgewiesen. Werte mit ungewöhnlich großen Abweichungen werden gesperrt. Derartig große Abweichungen kommen vor allem in kombinierten Auswertungen aus beiden Datenbeständen oder bei Darstellungen nach nicht im SAFE-Auswertungsdatenbestand enthaltenen regionalen Gliederungen vor, denn hier gelten die Qualitätsaussagen aus Abschnitt 2 („Wie stark verändert SAFE die Daten?“) zur Häufigkeitsverteilung der durch SAFE erzeugten Abweichungen der Tabellenfelder von ihren Originalwerten naturgemäß nicht. Das Sperr-Symbol „.“ wird zudem bei Fallkonstellationen eingesetzt, die beim Zusammenlegen von getrennt mit SAFE behandelten Datenbeständen entstanden sind, aber in den Originaldaten nicht vorkommen (z. B. 70-jährige Personen im Haushaltstyp „Haushalt ohne Senioren“).

4.2 Verhältniszahlen

Zur Berechnung von Verhältniszahlen – gebildet als Quotient aus Zähler und Nenner – (z. B. durchschnittliche Wohnungsgröße), werden in der Auswertungsdatenbank die Originaldaten benutzt. Grund dafür ist, dass ein Quotient von durch SAFE veränderten Zahlen in bestimmten Konstellationen erheblich vom Originalverhältniswert abweichen kann. Es kann sein, dass die Breite dieses Wertespektrums in keinem sinnvollen Verhältnis zur Darstellungsgenauigkeit steht (z. B. als Prozentzahl mit einer Nachkommastelle bei Anteilswerten).

Ein Beispiel dazu: Angenommen, es gibt in einer kleineren Gemeinde 20 Gebäude eines bestimmten Typs, von denen genau 50 % (also 10 Gebäude) in eine bestimmte Altersklasse des Baujahrs fallen. Wird diese Verhältniszahl aus den SAFE-Werten berechnet, können sich – je nach Ausprägung der SAFE-Veränderung – enorme Unterschiede ergeben. So könnte der Nenner durch SAFE beispielsweise um 4 auf 24 vergrößert und der Zähler um 5 auf 5 verkleinert worden sein – ausgewiesen würde ein Prozentsatz von 20,8 %. Umgekehrt, bei Verkleinerung des Nenners um 4 Fälle und Vergrößerung des Zählers um 5 Fälle durch SAFE, würde das Ergebnis als 93,8 % ausgewiesen.

Werden die Originaldaten benutzt, muss jedoch verhindert werden, dass aus den Verhältniswerten auf Originalwerte von Zählern oder Nennern zurückgeschlossen werden kann. Verhältniswerte werden deshalb nur dann ausgewiesen, wenn sie (a) für ausreichend große Gruppen statistischer Einheiten gebildet werden und (b) nicht im direkten Widerspruch zu den durch SAFE geänderten Zählern oder Nennern stehen.

Bei der Bewertung der Gruppengröße (Punkt (a)) spielt die Darstellungsgenauigkeit eine Rolle: Wenn beispielsweise in einer Gemeinde mit ca. 1 000 Einwohnern nur eine Person mit einer bestimmten Staatsangehörigkeit lebt, kann der Anteil dieser Staatsangehörigengruppe für diese Gemeinde nicht als Prozentzahl mit einer Nachkommastelle als 0,1 % ausgewiesen werden. Denn multipliziert man diesen Anteil mit der Einwohnerzahl wird offensichtlich, dass es sich um genau eine Person handelt ($0,001 * 1000 = 1$). In diesem Falle würde keine Verhältniszahl ausgewiesen werden. Dagegen wäre bei einem Ergebnisausweis ohne Nachkommastellen kein exakter Rückschluss möglich, denn das dann dargestellte Ergebnis 0 % kommt sowohl bei 1 als auch bei 2, 3 oder 4 Personen mit der betreffenden Staatsangehörigkeit zustande. Bei einer größeren Gruppe statistischer Einheiten kann meist auch mit einer oder sogar zwei Nachkommastellen ausgewiesen werden. In einer Gemeinde mit 10 000 Einwohnern beispielsweise kommt das Ergebnis 0,1 % bei jeder Staatsangehörigenanzahl zwischen 5 und 14 Personen zustande. Ein exakter Rückschluss auf die Originalzahl ist nicht möglich.

Mit Rücksicht auf Punkt (b) werden im Zuge der Rundung auf das vorgesehene Darstellungsformat (z. B. als Prozentwert mit einer Nachkommastelle) Ergebnisse gelegentlich aufgerundet, obwohl nach kaufmännischer Rundungsvorschrift abzurunden wäre und umgekehrt.

5 Fazit

Zur Sicherung der Geheimhaltung in den Ergebnistabellen des Zensus 2011 wurde das datenverändernde Verfahren SAFE eingesetzt. Auf dieser Grundlage konnte ein Datenangebot geschaffen werden, das flexible Auswertungen in vielfältigen Kombinationen ermöglicht. Gleichzeitig ist aber garantiert, dass Rückschlüsse auf Einzelangaben nicht möglich sind.

Bei der Mehrheit der Tabellenfelder liegt die durch SAFE bewirkte Veränderung der Originalhäufigkeiten bei bis zu +/-2 Einheiten (Einheiten sind je nach Thematik der Tabelle Personen, Haushalte usw.). Abweichungen nach oben und unten gleichen sich tendenziell aus, größere Abweichungen sind eher selten. Hier gilt: Je größer die Abweichung, desto seltener kommt diese vor. Bei den Bevölkerungsdaten treten Abweichungen in der Größenordnung ab 12 Personen sehr selten auf. Bei den übrigen ausgewiesenen Ergebnissen kommen – äußerst selten – auch Abweichungen von 20 und etwas mehr Einheiten vor.

Tabellenfelder mit „deutlich erhöhter“ Abweichung werden „markiert“. Das gilt insbesondere für Auswertungen, die Personenmerkmale und Merkmale anderer Einheiten kombinieren oder kleinräumige Auswertungen, denn auf diese Auswertungen treffen die o. a. Qualitätsaussagen allgemein nicht zu.

Bei Verhältniswerten werden Originaldaten ausgewiesen. Dabei wird teils auf den Nachweis von Nachkommastellen verzichtet, um zu vermeiden, dass Rückschlüsse auf Originalwerte von Zähler oder Nenner möglich sind.

6 Literatur

- BMJV – Bundesministeriums der Justiz und für Verbraucherschutz (1987): Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG).
http://www.gesetze-im-internet.de/bundesrecht/bstatg_1987/gesamt.pdf
(Zugriff: 10.08.2014).
- Domingo-Ferrer, J.; Mateo-Sanz, J. M. (2002): Practical data-oriented microaggregation for statistical disclosure control. In: IEEE Transactions on Knowledge and Data Engineering 14(1)/2002, 189-201.
- Forbes, A.; Naylor, J.; Leaver, V.; Gare, M.; Hawkes, T.; Camden, M. (2009): Confidentiality Plans for the 2011 Censuses in the United Kingdom, Australia and New Zealand: A Comparison, Beitrag zur Joint UNECE/Eurostat work session on statistical data confidentiality, 2-4 December 2009, Bilbao.
<http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2009/wp.28.e.pdf> (Zugriff: 10.08.2014).
- Fraser, B.; Wooton, J. (2006): A proposed method for confidentialising tabular output to protect against differencing. In: Monographs of Official Statistics. Work session on Statistical Data Confidentiality, Eurostat-Office for Official Publications of the European Communities, Luxembourg, 299-302.
- Giessing, S.; Höhne, J. (2010): Eliminating Small Cells from Census Count Tables: Some Considerations on Transition Probabilities. In: J. Domingo-Ferrer J.; Magkos, E. (Eds.), Privacy in Statistical Databases, 52-65. New York: Springer-Verlag. LNCS 6344.

- Höhne, J. (2003): SAFE – Ein Verfahren zur Anonymisierung statistischer Einzelangaben. Statistisches Landesamt Berlin (Hrsg.). In: Statistische Monatsschrift. Nr. 3. https://www.statistik-berlin-brandenburg.de/Publikationen/Aufsaeetze/2003/MS-BE_200303-01.pdf (Zugriff: 10.08.2014).
- Forbes, A.; Naylor, J.; Leaver, V.; Gare, M.; Hawkes, T.; Camden, M. (2009): Confidentiality Plans for the 2011 Censuses in the United Kingdom, Australia and New Zealand: A Comparison, Beitrag zur Joint UNECE/Eurostat work session on statistical data confidentiality, 2-4 December 2009, Bilbao. <http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2009/wp.28.e.pdf> (Zugriff: 10.08.2014).
- Statistische Ämter des Bundes und der Länder (2014): Nutzer-Kurzinformation zu SAFE. https://www.zensus2011.de/SharedDocs/Downloads/DE/Merkmale/Nutzerhinweise_safe.pdf?__blob=publicationFile&v=16 (Zugriff: 10.08.2014).
- Zayatz, L.; Lucero, J.; Massell, P.; Ramanayake, A. (2009): Disclosure Avoidance for Census 2010 and American Community Survey Five-year Tabular Data Products, Beitrag zur Joint UNECE/Eurostat work session on statistical data confidentiality, 2-4 December 2009, Bilbao. <http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2009/wp.29.e.pdf> (Zugriff: 10.08.2014).