6. Dresdner Flächennutzungssymposium − Dresden, 11./12. Juni 2014

Sarah Gießing
Statistisches Bundesamt

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

Die Gewährleistung der statistischen Geheimhaltung ist seit jeher eine fundamentale Aufgabe der Statistischen Ämter des Bundes und der Länder. 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 Sekundärsperrungen gesetzt werden müssen, damit unterdrückte Felder auf Grund von ebenfalls veröffentlichten Summenangaben nicht rückrechenbar sind.

Die Auswertungssysteme zur Ergebnispräsentation des Zensus 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 – wie auch in vielen anderen Ländern − ein methodischer Wechsel zu einem datenverändernden Geheimhaltungsverfahren vollzogen: Im deutschen Zensus 2011 wird dazu das im Statistischen Landesamt Berlin entwickelte Verfahren SAFE eingesetzt.

Bei SAFE handelt sich um ein sogenanntes Mikroaggregationsverfahren. 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 Ansatz. Er sorgt dafür, dass die Veränderungen so vorgenommen werden, dass sie sich weitgehend untereinander ausgleichen. So wird erreicht, dass die Vielzahl aller für Veröffentlichungen vorgesehen Ergebnisse mit kleinstmöglicher Abweichung gegenüber den entsprechenden Originalhäufigkeiten reproduziert wird.

Kontaktdaten

Sarah Gießing
Statistisches Bundesamt
65180 Wiesbaden

Telefon: 0611 / 752701
Fax: 0611 / 754000
E-Mail: Sarah.Giessing@destatis.de