

Konzept zur faktischen Anonymisierung statistischer Einzelangaben (§ 16 Abs. 6 BStatG) – Scientific-Use-File (SUF) der Statistik des Bauabgangs –

Produkt: Scientific-Use-File
Statistik: Statistik des Bauabgangs (EVAS: 31141)

Vorbemerkungen

In Form faktisch anonymisierter Mikrodaten bieten die Forschungsdatenzentren einem wissenschaftlichen Nutzerkreis die Einzelangaben stark nachgefragter Statistiken als standardisierte Scientific-Use-Files (SUF) an. Diese Daten sind so weit anonymisiert, dass sie außerhalb der geschützten Räume der amtlichen Statistik verwendet werden dürfen. Sie bieten daher im Vergleich zu den On-Site-Zugangswegen ein geringeres Analysepotenzial, sind jedoch so konzipiert, dass sie sich für einen großen Teil der wissenschaftlichen Forschungsvorhaben eignen. Der Nutzerkreis ist auf Hochschulen und Einrichtungen mit der Aufgabe unabhängiger wissenschaftlicher Forschung beschränkt. Scientific-Use-Files dürfen ausschließlich in den Räumen der Einrichtungen, die die Nutzung der SUF beantragt und genehmigt bekommen haben, genutzt werden. Die Datennutzung wird weiterhin durch einen Nutzungsvertrag abgesichert, in dem ein Reidentifikationsverbot sowie eine Strafe bei Vertragsbruch festgeschrieben ist.

Die Vorbereitung, Bearbeitung und Betreuung von Nutzungsanfragen im Rahmen des Forschungsdatenzentrums umfassen je nach Anfrage, Datenmaterial und Nutzungsweg in der Regel umfangreiche Abstimmungs- und Prüfarbeiten seitens der amtlichen Statistik sowie der Datennutzer. Vor allem bei der On-Site-Nutzung via Gastwissenschaftlerarbeitsplätzen oder der kontrollierten Datenfernverarbeitung ist der organisatorische, administrative sowie inhaltlich-praktische Arbeitsaufwand besonders hoch. Deshalb bietet das Forschungsdatenzentrum der Statistischen Ämter des Bundes und der Länder u.a. Scientific-Use-Files für einen wissenschaftlichen Nutzerkreis an.

Das vorliegende Anonymisierungskonzept, zur Erstellung von Scientific-Use-Files der Bauabgangsstatistik hilft dabei, den Prüf- und Abstimmungsaufwand bei der Bearbeitung von Nutzungsanfragen in Zukunft zu verringern, da es sich hierbei um ein einmalig zu prüfendes und abzustimmendes Dokument handelt. Das Konzept stellt den Merkmalsumfang und die notwendigen Anonymisierungsmaßnahmen zur Erstellung der Scientific-Use-Files der Bauabgangsstatistik dar und kann, einmal abgestimmt, in Zukunft für den Großteil der Datennutzungsanfragen herangezogen werden.

Reidentifikation(en) und deren Nutzen

Einzeldaten für die Weitergabe an wissenschaftliche Einrichtungen gelten gemäß § 16 Abs. 6 Bundesstatistikgesetz (BStatG) dann als ausreichend anonymisiert, wenn sie nur mit einem „unverhältnismäßig hohen Aufwand“ reidentifiziert werden können. Das Unverhältnismäßigkeitsgebot des BStatG impliziert dabei, dass eine Verletzung der faktischen Anonymität nur im Falle einer als nutzbringend einzustufenden Reidentifikation gegeben sein kann. Der Nutzen, der durch die Zuordnung

eines Merkmals entsteht, hängt stark vom Informationsgehalt und somit von der Sensibilität des zu reidentifizierenden Merkmals ab. Ein Merkmal, das dem Datenangreifer keine verwertbare Information bringt, liefert kein plausibles Motiv für einen Datenangriff und gilt daher als unsensibel.

Laut des Leitfadens zur Sicherstellung der statistischen Geheimhaltung im Rahmen der Bautätigkeitsstatistiken vom 07.12.2010, welcher bei der Geheimhaltungsprüfung zu veröffentlichender Auswertungen Anwendung findet, können die Daten der Bauabgangsstatistik bis auf Gemeindeebene veröffentlicht werden. Aus diesem Grund werden die Scientific-Use-Files der Bauabgangsstatistik nach dem vorliegenden Anonymisierungskonzept zur Off-Site-Nutzung bis auf Gemeindeebene als tiefste regionale Gliederung bereitgestellt. Somit bleibt nahezu das gesamte Analysepotential der Mikrodaten erhalten, weshalb sich die Scientific-Use-Files der Bauabgangsstatistik je nach Bedarf auch für die On-Site-Nutzung an Gastwissenschaftsarbeitsplätzen oder per kontrollierter Datenfernverarbeitung anbieten.

Basismaterial

Die Daten der Statistik des Bauabgangs können grundsätzlich für die Berichtsjahre ab 2000 bereitgestellt werden. Das Datenmaterial zum Berichtsjahr 2000 enthält keine Angaben für Schleswig-Holstein und Hamburg. Das Berichtsjahr 2001 wird ohne Schleswig-Holstein und Bremen bereitgestellt. Die Daten des Jahres 2002 beinhalten keine Angaben aus Bremen.

Merkmale

Die Mikrodatenfiles der Bauabgangsstatistik werden durch Löschung um die Regionalkennung unterhalb der Gemeindeebene bereinigt (EF2U5 „Gemeindeteil“). Die erhebungstechnisch relevanten Merkmale „Identifikationsnummer“ (EF1) sowie „Straßenschlüssel“ (im alten Datensatz EF27 bzw. im neuen Datensatz EF25) werden ebenfalls aus den bereitzustellenden Daten entfernt. Aufgrund von Statistikänderungen weicht die Datensatzstruktur der Jahre 2000-2004 von der Struktur der Jahre ab 2005 ab. Siehe nachfolgende Merkmalslisten:

Merkmale des anonymisierten Einzeldatenmaterials der Statistik des Bauabgangs von 2000 bis 2004

Feldbezeichnung	Merkmal
EF1	Identifikationsnummer
EF2U1	Land
EF2U2	Regierungsbezirk (mit EF2U1)
EF2U3	Kreis (mit EF2U2)
EF2U4	Gemeinde (mit EF2U3)
EF2U5	Gemeindeteil
EF3	Berichtszeitraum des Bauabgangs
EF3U1	Berichtszeitraum des Bauabgangs – Jahr
EF3U2	Berichtszeitraum des Bauabgangs – Monat
EF4	Datum des Bauabgangs
EF4U1	Datum des Bauabgangs – Jahr
EF4U2	Datum des Bauabgangs – Monat
EF5	Eigentümer
EF7	Wohngebäude/ Wohnheim
EF8	Art des Nichtwohngebäudes
EF9	Baualter - klassifiziert

Feld- bezeichnung	Merkmal
EF10	Umfang des Abgangs
EF11	Art und Ursache des Abgangs
EF15	Nutzfläche in qm
EF16	Wohnfläche in qm
EF17	Abgegangene Wohnungen mit 1 Raum
EF18	Abgegangene Wohnungen mit 2 Räumen
EF19	Abgegangene Wohnungen mit 3 Räumen
EF20	Abgegangene Wohnungen mit 4 Räumen
EF21	Abgegangene Wohnungen mit 5 Räumen
EF22	Abgegangene Wohnungen mit 6 Räumen
EF23	Abgegangene Wohnungen mit 7 und mehr Räumen
EF24	Zahl der Räume in Wohnungen mit 7 oder mehr Räumen
EF25	Sonstige abgegangene Wohneinheiten
EF26	Räume in sonstigen abgegangene Wohneinheiten
EF27	Straßenschlüssel
EF29	Kreisangehörige Gemeinde mit 20.000 und mehr Einwohnern
EF30	Gemeindegrößenklassen

Merkmale des anonymisierten Einzeldatenmaterials der Statistik des Bauabgangs ab 2005

Feld- bezeichnung	Merkmal
EF1	Identifikationsnummer
EF2U1	Land
EF2U2	Regierungsbezirk (mit EF2U1)
EF2U3	Kreis (mit EF2U2)
EF2U4	Gemeinde (mit EF2U3)
EF2U5	Gemeindeteil
EF3	Gemeindegrößenklassen
EF4	Kreisangehörige Gemeinde mit 20.000 und mehr Einwohnern
EF5	Berichtszeitraum des Bauabgangs – Jahr
EF6	Datum des Bauabgangs
EF6U1	Datum des Bauabgangs – Monat
EF6U2	Datum des Bauabgangs – Jahr
EF7	Eigentümer
EF8	Wohngebäude/ Wohnheim
EF9	Art des Nichtwohngebäudes
EF10	Baualter - klassifiziert (<i>bis 2015 erhoben</i>)
EF11	Umfang des Abgangs
EF12	Art und Ursache des Abgangs
EF13	Nutzfläche in qm
EF14	Wohnfläche in qm
EF15	Abgegangene Wohnungen mit 1 Raum
EF16	Abgegangene Wohnungen mit 2 Räumen
EF17	Abgegangene Wohnungen mit 3 Räumen
EF18	Abgegangene Wohnungen mit 4 Räumen
EF19	Abgegangene Wohnungen mit 5 Räumen
EF20	Abgegangene Wohnungen mit 6 Räumen
EF21	Abgegangene Wohnungen mit 7 und mehr Räumen
EF22	Zahl der Räume in Wohnungen mit 7 oder mehr Räumen
EF23	Sonstige abgegangene Wohneinheiten (<i>bis 2011 erhoben</i>)
EF24	Räume in sonstigen abgegangene Wohneinheiten (<i>bis 2011 erhoben</i>)
EF25	Straßenschlüssel
EF27	Baualter - neu klassifiziert (<i>ab 2016 erhoben</i>)