

Konzept zur Anonymisierung des Mikrozensus 1976 zur Verwendung als CAMPUS-File

I. Vorbemerkungen

Das CAMPUS-File ist ein absolut anonymisiertes Einzeldatenmaterial, das für Lehrende und Studierende erstellt wird. Seine Funktion besteht darin, die praktische Statistikausbildung mit amtlichen Einzeldaten anzureichern und damit den Hochschulen ein effektives Werkzeug für eine qualitativ hochwertige Lehre zu liefern.

Das vorliegende Anonymisierungskonzept befasst sich mit der absoluten Anonymisierung des Mikrozensus 1976. Das Konzept sieht vor, dass eine 10%-Stichprobe aus dem Originalmaterial gezogen und die für das Scientific-Use-File vorgesehenen Anonymisierungsmaßnahmen übernommen werden.

II. Basismaterial

Das Einzelmateriale des Mikrozensus 1976 dient als Ausgangsmaterial bei der Erstellung des Campus-Files. Das Erhebungsprogramm des Mikrozensus 1976 umfasst insgesamt 146 Variablen und 632.602 Datensätze (ohne Gewichtung und Selektion).

III. Anonymisierungsmaßnahmen

Im nachfolgenden Abschnitt werden die Maßnahmen beschrieben, die zur Einhaltung der Maßgabe der absoluten Anonymität der Einzeldaten eingesetzt wurden.

1. Direkte Identifikatoren (wie Personen- und sonstige Ordnungsnummern), soweit sie im Mikrozensus vorhanden sind, wurden gelöscht und durch systemfreie Ordnungsnummern ersetzt.
2. Die Anordnung der Datensätze der anonymisierten Einzelangaben ist durch die Vergabe von systemfreien Ordnungsnummern selbst systemfrei.
3. Die anonymisierten Einzelangaben beschränken sich auf eine zufällig gezogene Unterstichprobe der Haushalte von 10 % des Mikrozensus.
4. Als unmittelbare Regionalangabe wird nur das Bundesland weitergegeben.
5. Das Merkmal "Gemeindegrößenklassen" wurde so zusammengefasst, damit keine einzelne Gemeinde mit weniger als 500 000 Einwohnern identifizierbar ist und in jeder Gemeindegrößenklasse mindestens 400 000 Einwohner vertreten sind. Darüber hinaus wurden weitere Zusammenfassungen in einzelnen Ländern vorgenommen.

6. Staatsangehörigkeiten oder Gruppen von Nationalitäten, denen weniger als ca. 50 000 Personen angehören, sind zu größeren Gruppen zusammengefasst und daher nicht einzeln identifizierbar.
7. Jede ausgewiesene Merkmalsausprägung umfasst in der univariaten Verteilung der Grundgesamtheit hochgerechnet mindestens 5 000 Fälle. Gegebenenfalls wurde dies durch sachgerechte Zusammenfassung von Merkmalsausprägungen erreicht.

IV. Fazit

Die in dem vorliegenden Konzept beschriebenen Anonymisierungsmaßnahmen führen zur absoluten Anonymität des CAMPUS-Files des Mikrozensus 1976.