

Regelungen zur Auswertung von Mikrodaten

(Kurzfassung)

Basisanforderungen an die Syntaxerstellung

Im Folgenden sind zentrale Anforderungen an die Erstellung der Syntax dargelegt. Diese Anforderungen sind einzuhalten, damit die im Rahmen der Kontrollierten Datenfernverarbeitung (KDFV) erstellte Syntaxen ausgeführt bzw. am Gastwissenschaftlerarbeitsplatz (GWAP) erzeugten Ergebnisse einer Prüfung auf Sicherstellung der Geheimhaltung durch die Mitarbeiter/-innen der Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder (FDZ) unterzogen werden. Sollten die Anforderungen nicht berücksichtigt werden, erfolgt keine Ausführung der Syntax bzw. Geheimhaltungsprüfung der erstellten Ergebnisse.

Ausführliche Erläuterungen und Beispiele zu den Anforderungen finden Sie unter dem folgenden Link: www.forschungsdatenzentrum/heimhaltung.asp.

1. Die Syntax ist optisch übersichtlich (ggf. unter Verwendung einer Master-Syntax) zu gestalten und die durch FDZ-Mitarbeiter/-innen bereit gestellte Mustersyntax ist anzuwenden.
2. Die Syntax ist umfassend und für Dritte verständlich zu kommentieren.
3. Alle Variablennamen und Wertelabel sind eindeutig zu benennen.
4. Der Output muss durch die zugehörige Syntax identisch reproduzierbar sein.
5. Alle erstellten tabellarischen und analytischen Auswertungen sind in einem weiter zu verarbeitenden Format (z.B. EXCEL, SAV etc.) zu speichern. Alle grafischen Auswertungen sind in einem *nicht* weiter zu verarbeitenden Format (z.B. PDF, JPEG) zu speichern.
6. In der Syntax ist eindeutig zu kennzeichnen, welche Ergebnisse freigegeben werden sollen und welche Ergebnisse ausschließlich der Geheimhaltungsprüfung durch die FDZ dienen.
7. Zur Durchführung der Geheimhaltungsprüfung sind bei der Erstellung von Tabellen für Teilpopulationen stets auch die Tabellen für die übrige Teilpopulation auszugeben. Die entsprechenden Bezüge sind kenntlich zu machen (vgl. Nr. 6).
8. Zur Durchführung der Geheimhaltungsprüfung sind für alle freizugebenden Ergebnisse die zugrundeliegenden ungewichteten Fallzahlen anzugeben. Die entsprechenden Bezüge sind kenntlich zu machen (siehe auch Nr. 6).
9. Zur Durchführung der Geheimhaltungsprüfung sind für alle freizugebenden grafischen Auswertungen die zugrundeliegenden ungewichteten Fallzahlen auszugeben. Die entsprechenden Bezüge sind kenntlich zu machen (siehe auch Nr. 6).
10. Zur Durchführung der Geheimhaltungsprüfung sind bei freizugebenden Wertetabellen und statistischen Kennzahlen auf Basis von Wirtschafts- und Steuerstatistiken die beiden höchsten Werte, die Fallzahl sowie die Summe auszuweisen. Die entsprechenden Bezüge sind kenntlich zu machen (siehe auch Nr. 6).

Stand: 1. September 2017

11. Identische statistische Ergebnisse sind innerhalb einer Datennutzung lediglich einmal als freizugeben zu kennzeichnen. Falls in begründeten Ausnahmefällen Ergebnisse erneut bereitgestellt werden müssen, ist ein exakter Verweis auf die entsprechend frühere Auswertung erforderlich.
12. Das Anspielen von Merkmalen ist nur zulässig, sofern dies im Vorfeld mit den Mitarbeiter/-innen der FDZ abgestimmt wurde und die anzuspielenden Daten ein bestimmtes Format aufweisen.

Kriterienkatalog zur Zulassung von Output

Im Folgenden sind die wichtigsten Regeln dargestellt, nach denen in der amtlichen Statistik die Sicherstellung der Geheimhaltung der statistischen Ergebnisse erfolgt. Diese Regelungen müssen auch auf die durch wissenschaftliche Analysen erstellten Auswertungen Anwendung finden. Dies wird durch die Mitarbeiter/-innen der Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder geprüft. Werden diese Regelungen durch wissenschaftliche Nutzer/-innen nicht eingehalten, besteht die Möglichkeit, dass die erstellten Auswertungen seitens der FDZ-Mitarbeiter/-innen nicht freigegeben werden.

1. Regeln bei Fallzahl- und Wertetabellen
 - 1.1 Ein Wert ist geheim zu halten, wenn zu diesem Wert nur 1 oder 2 Fälle beitragen. Dies gilt auch für weiterführende Analysen (sog. Mindestfallzahlregel).
 - 1.2 Ein Tabellenfeld ist geheim zu halten, wenn sich die Häufigkeit in einem Tabelleninnenfeld um höchstens 1 von der Häufigkeit des entsprechenden übergeordneten Randfeldes unterscheidet (sog. Randwertregel).
 - 1.3 Ein Wert ist geheim zu halten, wenn der Beitrag des größten Einzelwerts oder der beiden größten Einzelwerte einen festgelegten Anteil am Gesamtwert übersteigt (sog. Dominanzregel).
2. Regeln bei weiterführenden Analysen
 - 2.1 Einzelwerte sind prinzipiell geheim zu halten. Darunter können fallen u.a. Auflistungen von Einzelwerten (list-Befehl) sowie ggf. Minima und Maxima.
 - 2.2 Ein Ergebnis einer Quantilberechnung ist geheim zu halten, wenn zu einem Quantil-Abschnitt weniger als 3 Fälle beitragen.
 - 2.3 Eine Grafik ist geheim zu halten, wenn in den ihr zugrundeliegenden Fallzahlen oder Werten mindestens ein Geheimhaltungsfall auftritt (vgl. Nr. 1-3).
 - 2.4 Ergebnisse von Teilpopulationen sind geheim zu halten, wenn durch die Kenntnis der Ergebnisse der Teilpopulation und vorhandene Kenntnisse über die Gesamtpopulation geheimhaltungsrelevante Rückschlüsse auf die Restpopulation gezogen werden können.

Bitte beachten Sie, dass nicht alle Statistiken mit den gleichen Regeln geheim gehalten werden. Für einzelne Statistiken gelten nicht alle der oben beschriebenen Regeln, für andere Statistiken gibt es darüber hinausgehende Maßnahmen. Ausführliche Erläuterungen und Beispiele zu den Regeln sowie Hinweise zu ihrer Gültigkeit finden Sie unter dem folgenden Link: www.forschungsdatenzentrum/geheimhaltung.