Aufgabe 1 Wiederholungsfragen zu Kap. 4 – Datenproduktion: Amtliche Statistik und Anonymisierung von Einzeldaten

1. Welche Gesetze bilden die rechtliche Grundlage der amtlichen Statistik in Deutschland?
2. Nennen Sie Beispiele für die Notwendigkeit amtlicher Statistik.
3. Nennen Sie jeweils drei Beispiele für die Träger der amtlichen bzw. der nichtamtlichen Statistik in Deutschland.
4. Finden sich im Bundesstatistikgesetz (BStatG) Elemente des Datenschutzes?
5. Inwieweit hat das Volkszählungsurteil Einfluss auf die Gesetzgebung in Deutschland gehabt? Was ist die Kernaussage dieses Urteils?
6. Welchem Ministerium untersteht das Statistische Bundesamt und wo hat es seinen Sitz?
7. Welche Aufgaben hat das Statistische Bundesamt?
8. Dürfen Meldebehörden die Adresse einer gemeldeten Bürgerin an Dritte weitergeben?
9. Welche der folgenden Institutionen dürfen statistische Erhebungen ohne Gesetzesgrundlage machen:
   - Statistisches Bundesamt
   - Industrie- und Handelskammern
   - Sie als Bürger
   - das Statistikamt der Stadt München
   - die Gesellschaft für Konsumforschung in Nürnberg
   - Statistische Landesämter
10. Unter welcher Bedingung dürfen Einzelangaben aus amtlichen statistischen Erhebungen an Hochschulen und unabhängige Forschungseinrichtungen weitergegeben werden?
11. Inwieweit schadet bzw. nutzt Datenschutz der Statistik?
12. Geben Sie für die folgenden Anonymisierungsverfahren jeweils an, ob es zu den informationsreduzierenden oder zu den datenverändernden Verfahren gehört:
    - Ersetzen der Werte eines metrischen Merkmals durch eine Indexzahl zu einer bestimmten Basis
    - (Sub-)Stichprobenziehung
    - einfaches Data-Swapping
    - Zusammenfassung der Kategorien eines kategorialen Merkmals
13. Lassen sich auf Basis eines mithilfe der Mikroaggregation (nach einer Variable) anonymisierten Datensatzes der Erwartungswert und die Varianz eines einzelnen Merkmals erwartungstreu schätzen?

Aufgabe 2 Betrachten Sie den folgenden fiktiven Mikrodatensatz, der bei einer Stichprobenbefragung von 10% der Mieter eines großen Wohngebäudes erhoben wurde:

Name	Bildung (Jahre)	Beschäftigung	Netto-Einkommen (EURO)	Zufriedenheit $(1, \ldots, 10)$	Familienstand
Schneider	12	beschäftigt	1 750	8	unverheiratet
Hartmann	14	sonstiges	4 600	10	verheiratet
Wichmann	21	beschäftigt	2 000	9	unverheiratet
Zwiesel	13	arbeitslos	850	6	unverheiratet
Dittrich	20	arbeitslos	1 200	8	verheiratet
Beck	15	beschäftigt	2 300	7	verheiratet

a) Welche 3 Gruppen von informationsreduzierenden Verfahren zur Anonymisierung haben Sie in der Vorlesung kennengelernt? Geben Sie jeweils ein Beispiel, wie der obige Datensatz damit anonymisiert werden könnte.

b) Anonymisieren Sie den obigen Datensatz mit der Methode des einfachen Data-Swapping, wobei als Gruppierungsvariable das Merkmal „Familienstand“ herangezogen wird.

c) Betrachten Sie den Teildatensatz des obigen Datensatzes, in dem nur die metrischen Merkmale enthalten sind. Anonymisieren Sie diesen Datensatz mit der Methode der deterministischen Mikroaggregation nach der Variable „Zufriedenheit“ .