

**Aufgabe 25**

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

Name	Bildung (Jahre)	mon. Netto- Einkommen (EURO)	mon. Ausgaben für Telekomm. (EURO)	Zufriedenheit (1, ..., 10)	Familienstand
Winzer	10	1 750	50	8	unverheiratet
Schneider	14	4 600	88	9	verheiratet
Hartmann	21	2 000	16	9	unverheiratet
Zwiesel	12	850	36	5	unverheiratet
Dittrich	19	1 200	20	8	verheiratet
Beck	15	2 300	48	7	verheiratet
Fritsche	16	3 250	71	9	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) Betrachten Sie den obigen Datensatz nachdem die direkten Identifikationsmerkmale entfernt wurden.
  - (i) Anonymisieren Sie den Datensatz mit der Methode des einfachen Data-Swapping, wobei als Gruppierungsvariable das Merkmal „Familienstand“ herangezogen wird.
  - (ii) Lassen sich auf Basis des anonymisierten Datensatzes der Erwartungswert und die Varianz eines einzelnen Merkmals erwartungstreu schätzen?
- c) Betrachten Sie den Teildatensatz des obigen Datensatzes, in dem nur die metrischen Merkmale enthalten sind.
  - (i) Anonymisieren Sie diesen Datensatz mit der Methode der deterministischen Mikroaggregation nach der Variable „Zufriedenheit“ .
  - (ii) Lassen sich auf Basis des anonymisierten Datensatzes der Erwartungswert und die Varianz eines einzelnen Merkmals erwartungstreu schätzen?