

Übung 12

Explorative Datenanalyse und Visualisierung

Wintersemester 2019
S. Döhler (FBMN, h_da)

Name:

Aufgabe 29. Arbeiten Sie weiter an dem Datensatz `UmfrageBis2019.csv` (s. Aufgabe 21).

- a) Erzeugen Sie eine Scatterplotmatrix des gesamten Datensatzes (entfernen Sie ggf. irrelevante Merkmale). Experimentieren Sie auch mit der Funktion `'gpairs'` aus dem gleichnamigen R-Paket.
- b) Stellen Sie nun den 3-dimensionalen Datensatz der Merkmale Schuhgrösse, Grösse und Anzahl Schuhe als 3-dimensionalen scatterplot dar (Sie können dazu das Paket `'scatterplot3d'` verwenden). Welche Einstellung der Parameter liefert eine aufschlussreiche Darstellung? Interpretieren Sie die Daten.
- c) Visualisieren Sie die Correlationsmatrix durch ein Correlogramm der Daten aus b) z.B. mithilfe der Pakete `'corrplot'` und `'psych'`. Interpretieren Sie die Ergebnisse.
- d) Stellen Sie die Daten aus b) durch einen parallelen Koordinaten-Plot dar. Interpretieren Sie diesen.
- e) Betrachten Sie nun die gemeinsame Verteilung von `'Grösse'` und `'Schuhgrösse'`. Sie sollen einen 2-dimensionalen Kerndichteschätzer der Daten visualisieren. Ein solcher ist beispielsweise im Paket `'MASS'` implementiert.
 - i) Stellen Sie einen Höhenlinienplot des Schätzers zusammen mit den Originaldaten dar.
 - ii) Stellen Sie einen Graphen des Schätzers dar.
 - iii) Stellen Sie eine heatmap des Schätzers dar.

Interpretieren Sie die Ergebnisse.

Denken Sie bitte wie immer daran, die Parameter der Grafiken sorgfältig zu wählen sowie sie sinnvoll zu beschriften (Titel, Achsen, etc)!

Anmerkungen/Korrektur
