

Übung 12

Explorative Datenanalyse und Visualisierung

Wintersemester 2019
S. Döhler (FBMN, h_da)

Name:

Aufgabe 27. Arbeiten Sie das Kapitel 22 in *R for Data Science* durch.

Aufgabe 28. Analysieren Sie die bereinigten Daten `UmfrageBis2019.csv` (s. Aufgabe 21) mit `ggplot2` (arbeiten Sie wieder mit den neuen Spaltennamen...).

a) Erzeugen Sie ein Histogramm der Variablen **Groesse**

- i) Für die Gesamtpopulation
- ii) Getrennt nach Geschlechtern (arbeiten Sie mit facets)

Was fällt Ihnen bei der Default-Klasseneinteilung auf? Verwenden Sie zusätzlich die Aufteilung nach der Diaconis-Friedman-Methode.

b) Stellen Sie Kerndichteschätzer der Variablen **Groesse** dar.

- Getrennt nach Geschlechtern, jedoch in einer gemeinsamen Grafik (mit verschiedenen Farben und Schraffierungen)
- Fügen Sie einen rug-Plot mit entsprechenden Farben hinzu.

c) Stellen Sie die empirische Dichtefunktion der Variablen **Groesse** dar.

- Getrennt nach Geschlechtern, jedoch in einer gemeinsamen Grafik (mit verschiedenen Farben und Schraffierungen)
- Fügen Sie einen rug-Plot mit entsprechenden Farben hinzu.

d) Erzeugen Sie Box- und Violin-Plots der Variablen **Groesse** getrennt nach Geschlechtern, jedoch in jeweils einer gemeinsamen Grafik.

e) Erzeugen Sie einen Scatterplot der Variablen **Schuhgroesse** (y-Achse) und **Groesse** (x-Achse).

- Färben Sie die Datenpunkte nach Geschlecht
- Passen Sie pro Geschlecht jeweils eine lineare Regression an und stellen Sie die resultierenden Regressionsgeraden mit der passenden Farbe zusammen mit den Daten dar.

Lösung

Anmerkungen/Korrektur
