

# Datenaufbereitung

## Übungsaufgabe 1\*

Auf der Homepage der Universitätsbibliothek finden Sie einen Datensatz <https://data.ub.uni-muenchen.de/13/2/atemwege.asc>, der Daten zu einer Studie von der Universität Innsbruck über Lungenfunktion und Atemwegserkrankungen bei Schulkindern enthält. Weitere Informationen zu dem Datensatz finden Sie unter <https://data.ub.uni-muenchen.de/13/>; die Variablenbeschreibungen finden Sie unter <https://data.ub.uni-muenchen.de/13/1/DETAILS.html>.

- a) Setzen Sie das Arbeitsverzeichnis auf den Ordner, in dem Sie arbeiten möchten.
- b) Lesen Sie den Datensatz direkt in R ein, ohne ihn vorher lokal abzuspeichern. Er soll unter dem Namen `atemwege` eingelesen werden.
- c) Verschaffen Sie sich unter Verwendung geeigneter Funktionen einen Überblick über die Daten.
- d) Betrachten Sie die Variablen `Bildungsstand der Eltern`, `Geschlecht` und `Allergische Erkrankungen der Atemwege`. Von welchen Objektklassen sind diese Variablen? Erscheint Ihnen dies sinnvoll? Falls nein, passen Sie die Objektklassen an.
- e) Speichern Sie den Datensatz `atemwege`, auch für spätere Analysen, in einem passenden R-Format in Ihrem Arbeitsverzeichnis.
- f) Betrachten Sie die Variable `Bildungsstand der Eltern`. Welche und wie viele Ausprägungen hat die Variable? Passen Sie die Namen der Ausprägungen in geeigneter Weise an.
- g) Wenden Sie die Funktion `plot()` auf die Variablen `Bildungsstand der Eltern` und `Körpergewicht` an. Welche Diagrammtypen erhalten Sie? Begründen Sie dies mit geeigneten Aufrufen.
- h) Berechnen Sie das mittlere Körpergewicht getrennt nach Geschlecht.
- i) Berechnen Sie die Kennzahlen des Boxplots der Variable `Vitalkapazität`.
- j) Geben Sie den Teil des Datensatzes `atemwege` aus, der nur im Jahr 1975 geborene Mädchen enthält, deren Körpergewicht unter dem Mittelwert liegt.
- k) Laden Sie das R-Paket `dplyr` und öffnen Sie die Vignette `introduction`. Führen Sie Teilaufgabe j) nochmal mithilfe einer geeigneten Funktion aus dem Paket `dplyr` aus.

\* *Durch Lösen dieser Übungsaufgabe kann ein Teil der Prüfungsleistung erbracht werden.*

**Abgabetermin: 09.05.2017/10.05.2017**