

Checkliste Lernziele:

- Wie öffnet man Daten mit R?
- Wie ruft man bestimmte in R geladene Daten auf (bspw. bestimmte Variablen) und wie kann man diese verändern?
- Wie lassen sich Daten mit R einfach beschreiben?
- Wie lassen sich einfache Abstände zwischen Beobachtungen bzw. Verteilungen berechnen?

Aufgaben:

1. Schauen Sie sich die Hilfe zum Befehl `read.table` an und versuchen Sie, die Datei `klaus1.dat` zu laden (Box 2.2-1 im Skript).
2. Verwenden Sie den Befehl `str`, um eine Übersicht über die Daten zu erhalten. Welche Informationen liefert der Output? Vergleichen Sie diese mit den Angaben im Skript!
3. Lassen Sie sich für alle Variablen einfache deskriptive Kennwerte ausgeben (über den Befehl `summary`). Lassen sich die Ergebnisse für die einzelnen Variablen vergleichen?
4. Schauen Sie sich die Hilfe für den Befehl `dist` an. Berechnen Sie Abstandsmatrizen sowohl mit dem euklidischen Abstand, als auch mit der City-Block-Metrik. Vergleichen Sie die Ergebnisse mittels der Befehle `table`, `summary` und `plot`. Wie unterscheiden sich die beiden Abstandsmaße?
5. Verwenden Sie den Befehl `unique`, um mehrfach auftretende Datenzeilen auszuschließen. Worauf müssen Sie dabei achten?
6. Wiederholen Sie Aufgabe 4 mit dem eingegrenzten Datensatz. Welche Unterschiede zeigen sich?
7. Verwenden Sie den Befehl `cor` um Tabelle 2.2-2 zu replizieren. Benutzen Sie dabei wieder den ursprünglichen Datensatz.

8. Laden Sie den Datensatz `bs1.dat`. Benennen Sie die Variablen wie im Skript (S. 12) und verwenden Sie den Befehl `xtabs` um Tabelle 2.3-1 zu reproduzieren.

9. Tabelle 2.3-3 lässt sich wie folgt reproduzieren. Versuchen Sie die entsprechenden Befehle nachzuvollziehen:

```
tab1 <- xtabs(h~, dat)
tab2 <- prop.table(tab1,1)
tab3 <- ftable(tab2, col.vars=c("Y","Z"),row.vars="X")
dist(tab3,method="manhattan")/2
```