Florian Wickelmaier
Forschungsmethoden und Mathematische Psychologie
Fachbereich Psychologie
Universität Tübingen

# Übungsaufgaben
# Data structures in R (1)

Erzeugen Sie ein lauffähiges und kommentiertes R-Skript, in dem Sie die folgenden Aufgaben bearbeiten.

a) Benutzen Sie die Funktion `mean()`, um den Mittelwert des Vektors $\mathbf{a} = (1\ 3\ 7\ 1\ 2\ 6\ 7\ 8\ 9\ \text{NA})'$ zu berechnen. Wieso ist das Ergebnis `NA`? Wie kann man den Mittelwert ohne die fehlende Zahl berechnen?

b) Machen Sie sich anhand der Funktion `mean()` klar, was auf den Folien mit den Begriffen "positional matching" und "keyword matching" gemeint ist. Welche Argumente nimmt die Funktion (siehe `?mean`)? Was macht das Argument `trim`? Verwenden Sie zuerst alle Argumente, ohne deren Namen explizit anzugeben (positional matching). Verändern Sie dann die Reihenfolge der Argumente (keyword matching).

c) Erzeugen Sie drei Objekte in R: `apfel <- banane <- limone <- 1`. Geben Sie den R-Befehl an, der alle Objekte, die ein `n` im Namen haben, entfernt.

d) Erzeugen Sie einen Vektor $\mathbf{x} = (1\ 5\ 8\ 3\ 7\ 2\ 6)'$. Erzeugen Sie einen zweiten Vektor $\mathbf{y}$ derselben Länge, der die ungeraden Zahlen $1, 3, 5, \ldots$ enthält. Benutzen Sie dafür die Funktion `seq()`. Berechnen Sie einen Vektor $\mathbf{z}$ als Linearkombination von $\mathbf{x}$ und $\mathbf{y}$: $\mathbf{z} = 4\mathbf{x} + 2\mathbf{y}$.

   Erzeugen Sie möglichst effizient eine Matrix $\mathbf{B}$, die aus den Spaltenvektoren $\mathbf{x}$, $\mathbf{y}$, und $\mathbf{z}$ besteht. Erzeugen Sie dann auf zwei Arten eine Matrix $\mathbf{C}$, in der die drei Vektoren als Zeilenvektoren fungieren.

   Lesen Sie durch Indizieren (siehe `?Extract`) den Vektor $\mathbf{y}$ aus $\mathbf{B}$ und $\mathbf{C}$ wieder heraus.

e) Benutzen Sie die Funktion `seq()`, um folgende Zahlenreihen zu erzeugen:

   - $5, 10, 15, 20, \ldots$
   - $0.1, 0.2, 0.3, \ldots$
   - $0.25, 0.5, 0.75, 1.0, \ldots$

   Jede Reihe soll zehn Elemente enthalten. Welche anderen Methoden fallen Ihnen ein, um solche Vektoren zu erzeugen?

f) Erzeugen Sie den Vektor $(a\ a\ b\ b\ c\ c\ d\ d\ e\ e)'$ mit der `rep()`-Funktion. Wieso liefert die Eingabe `rep(letters[1:5], 2)` nicht das gewünschte Ergebnis?

g) Erzeugen Sie folgende Vektoren (ohne Sie "auszuschreiben"):

   - $(13\ 15\ 17\ 19\ 21\ 23\ 25\ 27\ 29\ 31\ 33)'$

- $(-2\ -1\ 0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8)'$. Wie lang ist dieser Vektor?
- Von 5 bis $-5$ in Schritten von 0.5.
- (exp1 exp2 exp3 ... exp15)$'$

h) Erzeugen Sie die Matrix

$$\mathbf{A} = \begin{pmatrix} 11 & 12 & 13 & 14 \\ 15 & 16 & 17 & 18 \end{pmatrix}.$$

Bilden Sie $\mathbf{A}'$ (die transponierte Matrix von $\mathbf{A}$) einmal, indem Sie eine Funktion verwenden, und einmal durch Indizieren.

i) Erzeugen Sie einen Faktor `G` mit `G <- factor(rep(c(0, 2, 5), 10))`. Führen Sie die folgenden R-Befehle aus:

```
mean(G)
mean(as.numeric(G))
mean(as.character(G))
mean(as.numeric(as.character(G)))
```

Vergleichen Sie die Ergebnisse und erklären Sie die Unterschiede.

Verwenden Sie die Funktionen `table()` und `xtabs()`, um zu überprüfen, ob jede Stufe des Faktors gleich häufig vorhanden ist.