

Self Assessment in R

An Introduction to Data Collection with R - Workshop

Philipp Behrendt & Jan Dix

14. Juni 2019

Mit diesem kleinen Self Assessment könnt ihr testen, wie gut es um eure R-Skills steht und euch schonmal ein wenig für den Workshop eingrooven. Die Lösungen zu den Aufgaben stellen wir in einer separaten Datei zur Verfügung. Bei der Bearbeitung könnt ihr natürlich alle möglichen Hilfsmittel benutzen (Google, StackOverflow, rHelp, etc . . .). Kein Mensch kann alle Befehle auswendig und die Hälfte unserer Zeit verbringen wir selbst damit, Lösungen/Beispiele und Dokumentation aus dem Internet zusammen zu suchen.

Aufgabenteil 1

- Erstelle einen Vektor **x** mit den Zahlen von 1 bis 10
- Berechne die Summe von **x**
- Erstelle einen Vektor **y**, der die Quadratzahlen von **x** enthält
- Füge die letzten 5 Elemente von **x** und die ersten 3 von **y** in einen neuen Vektor. Nenne diesen **z**.
- Bestimme den höchsten Wert von **z**.

Aufgabenteil 2

- Schreibe eine for-loop, die die Wurzeln der Zahlen 1 bis 10 ausgibt!
- Modifiziere deine for-loop so, dass du die Wurzel nicht printest, sondern dem Vektor **a** hinzufügst.
- Erstelle einen Vektor **b**, der sich nach folgender Formel berechnet. Wenn **a** kleiner als 2.5 ist, nimmt er 0 an, andernfalls 1.
- Erstelle einen plot mit **a** auf der x-Achse und **b** auf der y-Achse

Aufgabenteil 3

- Erstelle die Variable **vname** mit deinem Vornamen.
- Erstelle die Variable **nname** mit deinem Nachnamen.
- Wie viele Buchstaben enthält dein Vorname.
- Extrahiere den ersten Buchstaben deines Vornames und speicher ihn unter **acro** (Tipp: substr)
- Extrahiere die letzten 3 Buchstaben deines Nachnamens und speicher sie unter **nym**.
- Füge **acro** und **nym** zusammen speicher es in **kuerzel**.
- Transformiere **kuerzel** in Großbuchstaben.

Aufgabenteil 4

Lade den Iris-Datensatz mit folgendem Befehl!

```
iris <- iris
```

- Inspiziere den Datensatz !
- Lass dir zusammenfassende Statistiken für alle Variablen ausgeben (e.g. Quartile, Mix, Max). Es bedarf nur einer Funktion!
- Erstelle einen neuen Dataframe **setosa**, der nur die Einträge für setosa beinhaltet
- Erstelle einen neuen Dataframe **virginica**, der nur die Einträge für virginica beinhaltet
- Füge **setosa** und **virginica** nun zu einem neuen Dataframe **combined** *zeilenweise* zusammen.
- Welchen Wert hat der 54. Eintrag der 2. Spalte ? Welchen der 107. Eintrag der 9. Spalte ?