

28. November 2022

## Übungen 5

**Bitte geben Sie nicht nur die Antwort an sich ab, sondern auch den R-Code, der Sie zur Lösung geleitet hat. Das gilt nur, falls die jeweilige Antwort R-Code zur Lösung benötigt.**

### 01 Der Datensatz `lexdec` aus dem Paket `{languageR}`

- Laden Sie das Paket `{languageR}` und schauen Sie sich die Hilfe zum Datensatz `lexdec` an.
- Berechnen Sie die Spearman-Korrelation aus Frequenz (Frequency) und Reaktionszeit (RT). Bilden Sie vor der Berechnung eine Hypothese zum Zusammenhang der beiden Variablen heraus.
- Was würde eine negative Korrelation zwischen den Variablen RT und `Trial` inhaltlich bedeuten? Worauf könnte solch eine Korrelation zurückzuführen sein?
- Berechnen Sie die Spearman-Korrelation zwischen den beiden Variablen und interpretieren Sie das Ergebnis.

### 02 Pearson- und Spearman-Korrelation

- Laden Sie sich die Datei `cor-df.Rds` aus ADAM herunter und lesen Sie die Datei in R ein. Wenn Sie Probleme mit dem Download haben, melden Sie sich bitte per E-Mail.
- Berechnen Sie den Pearson-Korrelationskoeffizienten zwischen den Spalten A und B. Berechnen Sie außerdem den Spearman-Korrelationskoeffizienten.
- Vergleichen Sie die beiden Koeffizienten und finden Sie heraus, warum der Wert für den Pearson-Korrelationskoeffizienten so hoch ist ( $> 0,4$ ) während der Wert für den Spearman-Korrelationskoeffizienten so niedrig ist ( $< 0,04$ ). Tipp: Mit `plot(<x-Vektor>, <y-Vektor>)` können Sie ein Streudiagramm zweier Variablen erstellen.
- Bonusaufgabe: Indizieren Sie den Dataframe so, dass Sie Ihre Hypothese zum Ursprung des hohen Unterschieds überprüfen können.

### 03 Größeneinschätzungen von Pflanzen und Tieren: `ratings`

- Schauen Sie sich die Hilfe zum Datensatz `ratings` aus dem Paket `{languageR}` an. Achten Sie insbesondere auf die Variablen `meanSizeRating`, `Frequency` und `Class`.
- Erstellen Sie ein Regressionsmodell, mit dem Sie `meanSizeRating` aus `Frequency` und `Class` vorhersagen (nur *single effects*, ohne Interaktion).

- c. Lassen Sie sich die `summary()` dieses Modells ausgeben.
- d. Interpretieren Sie die beiden Effekte: Welchen Einfluss haben Wortfrequenz und semantische Klasse auf die eingeschätzte Größe der Pflanzen/Tiere? Wie viel Prozent der Varianz der eingeschätzten Größe wird von den beiden Prädiktoren gemeinsam vorhergesagt?
- e. Bonusaufgabe: Betreiben Sie Modelldiagnostik und geben Sie eine kurze vorläufige Einschätzung des *model fits* ab.