

# Übungsblatt 3

Datenanalyse und -visualisierung mit R  
Hochschule Bonn-Rhein-Sieg

Pekka Sagner M.Sc.

5. November 2021

## Einführung in ggplot2

### Aufgabe 1: »Always plot your data«

- a) Laden Sie den Datensatz **datasaurus.tsv** aus dem LEA-Kursordner herunter. Speichern Sie diesen in einem Unterordner (z. B. *data*) Ihres R-Projektordners.
- b) Der Datensatz enthält die Variable **gruppe**. Wie viele Gruppen gibt es? Wie viele Beobachtungen je Gruppe sind in den Daten enthalten? (*Hinweis: Nutzen Sie hierfür die Funktion `count()`.*)
- c) Bestimmen Sie den Mittelwert der Variablen **x** und **y** für die Gruppen **01**, **04**, **06** und **10**. Was fällt Ihnen auf?
- d) Erstellen Sie jeweils ein Punktediagramm (auch: Scatterplot) der Variablen **x** und **y** für die vier Gruppen **01**, **04**, **06** und **10**. Was fällt Ihnen auf?
- e) Erstellen Sie ein Punktediagramm der Variablen **x** und **y** für alle Gruppen. (*Hinweis: Nutzen Sie hierfür die Funktion `facet_wrap()` aus ggplot2.*)

### Aufgabe 2: Olympische Spiele 2

- a) Erstellen Sie den finalen Plot aus der Vorlesung, der den Zusammenhang zwischen dem Gewicht und der Größe der GoldmedaillengewinnerInnen bei den Olympischen Sommerspielen 2016 zeigt.
- b) Wie viele MedaillengewinnerInnen bei den Olympischen Sommerspielen 2016 waren männlich, wie viele weiblich? Erstellen Sie ein einfaches Säulendiagramm, das diese Häufigkeiten darstellt.
- c) Welches Team gewann bei den Olympischen Sommerspielen 2016 die meisten Medaillen? Wie viele davon waren Gold, Silber, Bronze? Stellen Sie ihr Ergebnis in einem Säulendiagramm dar.
- d) Stellen Sie die Entwicklung der Teilnehmerzahlen im Zeitverlauf grafisch dar.
- e) Stellen Sie die Altersverteilung der TeilnehmerInnen bei den Sommerspielen 2016 dar: (i) zuerst für alle TeilnehmerInnen und (ii) dann nach Geschlecht. Nutzen Sie hierfür Boxplots. (*Hinweis: Boxplots sind prinzipiell wie unten dargestellt aufgebaut. Mehr Informationen finden Sie z. B. in [Introduction to Modern Statistics, Kapitel 5.5](#).*)

## Exemplarischer Aufbau eines Boxplots

