

# Übungsblatt 4

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

Pekka Sagner M.Sc.

20. Mai 2022

## Messy Data $\rightsquigarrow$ Tidy Data

### Aufgabe 1: »Ordnung beginnt am Anfang«

- a) Laden Sie die Datei `Einkommen_Kreise.xlsx` herunter. Öffnen Sie die Datei in Excel oder einer OpenOffice-Variante. Welche Daten und Informationen enthält die Datei?
- b) Lesen Sie den Datensatz `Einkommen_Kreise.xlsx` mit der Funktion `read_excel` aus dem Paket `readxl` in R ein. Betrachten Sie den eingelesenen Datensatz als Tibble. Was fällt Ihnen auf? Was *stimmt nicht* mit dem Datensatz? Notieren Sie die Punkte, die an diesem Datensatz *nicht stimmen*.
- c) Bringen Sie den Datensatz in ein Tidy-Format.
- d) Stellen Sie die zeitliche Entwicklung der Haushaltseinkommen in Bonn, Düsseldorf, Köln und dem Rhein-Sieg-Kreis grafisch dar.
- e) Berechnen Sie die Relation der Haushaltseinkommen in Bonn, Köln und dem Rhein-Sieg-Kreis relativ zur Landeshauptstadt im Zeitverlauf.
- f) Stellen Sie die bestimmte Relation aus e) grafisch dar.