

Übung: Bestellungsdaten erkunden

Vorbereitung

Ladet den Datensatz der Campusbier-Bestellungen mit den folgenden Zeilen:

```
library(tidyverse)
orders <- read_csv("orders.csv")
```

Achtung: Der Code oben geht davon aus, dass die Datei `orders.csv` in eurem Arbeitsverzeichnis im Unterordner `data/` liegt. Ist das nicht Fall, muss die Pfadangabe angepasst werden.

Speichert die Zeilen in einem neuen R-Skript und tragt dort auch eure Lösungen für die folgenden Aufgaben ein. Vergesst nicht, die Datei zwischendurch zu speichern.

Aufgabe 1: Dimensionen ermitteln

Ermittelt die Anzahl Spalten und Zeilen des Datensatzes. Ihr werdet herausfinden, dass es dafür verschiedene Möglichkeiten gibt. Findet auch eine Möglichkeit, mit der ihr beide Größen auf jeweils einer Variable speichern könnt. Das kann zum Beispiel wichtig sein, wenn ihr mit den Größen im weiteren Skriptverlauf weiter arbeiten wollt.

Aufgabe 2: Nützliche Funktionen

Recherchiert und testet die folgenden Funktionen. Notiert euch in eigenen Worten, was jede Funktion zurückgibt:

- `head()`
- `tail()`
- `print()`
- `sample_n()`

- `glimpse()`
- `summary()`
- `colnames()`

Welche Parameter könnt ihr den Funktionen jeweils übergeben? Wozu würdet ihr die Funktionen verwenden?

Aufgabe 3: Datentypen ermitteln

Findet einen Weg, um den Datentyp für jede Spalte auf der Konsole auszugeben!

- Welche Datentypen sind vertreten?
- Sind eurer Meinung nach alle Spalten korrekt erkannt worden?

Aufgabe 4: Das Paket {skimr}

Installiert das Paket {skimr}, indem ihr die folgende Zeile ausführt:

```
install.packages("skimr")
```

Betrachtet nun die Funktion `skim()`.

- Was ermittelt die Funktion, wenn wir ihr einen Datensatz übergeben?
- Gebt mithilfe der Funktion nur den prozentualen Anteil fehlender Werte (NA) für jede Spalte auf der Konsole aus!