

## 2.2 Hands-On: R Markdown

Marco Wähler

2022-03-25

### Aufgabenstellung - Lösung

#### Aufgabe 1

Erstellen Sie ein R Markdown File im Arbeitsverzeichnis des R Projects. Die Output-Datei soll das Format “HTML” haben. Geben Sie dem Dokument ein Titel, einen Untertitel und fügen Sie Ihren Namen als Autor\*in ein.

```
---
title: "Titel des Dokuments"
subtitle: "Der Untertitel des Dokuments"
author: "Marco Waehner"
date: "`r Sys.Date()`"
output:
  html_document
---
```

#### Aufgabe 2 (Optional)

Spezifizieren Sie nun die `knitr` chunk Optionen, sodass alle Chunks im Dokument gleich interpretiert werden. Der Code in den chunks soll stets angezeigt werden, “messages” sollen nicht angezeigt werden. Hinweise zu den key-values finden Sie im entsprechenden Cheat Sheet

Lösung: `{r setup, include=FALSE} knitr::opts_chunk$set(echo = TRUE, message = FALSE)`

#### Aufgabe 3

Importieren Sie nun die Daten aus ALLBUS 2018.

```
library(rio)
```

```
df <- import("data/allbus_2018_gesamt.sav")
```

#### Aufgabe 4

Wie hoch ist der Anteil der Befragten, die an vier oder mehr Tagen in der Woche TV gucken? Die entsprechende Variable lautet `lm01`. Rekodieren Sie hierzu eine Dummy-Variable, die den Wert 1 annimmt,

wenn die Befragten an vier oder mehr Tagen Fernsehen schauen und den Wert 0 annimmt, wenn die Befragten weniger als vier Tage (auch "nie" oder "seltener") Fernsehen schauen. Erstellen Sie anschließend eine Häufigkeitstabelle mit absoluten, relativen und prozentualen Häufigkeiten. Optional: Vergeben Sie Labels für die entsprechenden Zeilen mit `row.names()`.

```
attributes(df$lm01)
```

```
## $label
## [1] "HAEUFIGKEIT VON FERNSEHEN PRO WOCH"
##
## $format.spss
## [1] "F3.1"
##
## $display_width
## [1] 6
##
## $labels
##           NIE           SELTENER       AN EINEM TAG       AN 2 TAGEN
##           0.0           0.5           1.0           2.0
##       AN 3 TAGEN       AN 4 TAGEN       AN 5 TAGEN       AN 6 TAGEN
##           3.0           4.0           5.0           6.0
## AN ALLEN 7 TAGEN
##           7.0
```

```
#Häufigkeitsauszählung
table(df$lm01)
```

```
##
##    0  0.5    1    2    3    4    5    6    7
## 174 108 115 202 220 226 272 191 1969
```

```
#recode dummy-variable
df$tv <- ifelse(df$lm01 >= 4, 1,0)
table(df$tv)
```

```
##
##    0    1
## 819 2658
```

```
#Häufigkeitstabelle
absolut <- table(df$tv)
relativ <- prop.table(absolut)
prozentual <- 100 * relativ

#optional auch kumuliert
kumuliert <- cumsum(prozentual)

haeufigkeiten <- cbind(absolut, relativ, prozentual, kumuliert)
print(haeufigkeiten)
```

```
##   absolut   relativ prozentual kumuliert
## 0      819 0.2355479  23.55479 23.55479
## 1     2658 0.7644521  76.44521 100.00000
```

```
row.names(haeufigkeiten) <- c("weniger/gleich 3 Tage", "mehr/gleich 4 Tage")  
print(haeufigkeiten)
```

```
##               absolut   relativ prozentual kumuliert  
## weniger/gleich 3 Tage      819 0.2355479   23.55479  23.55479  
## mehr/gleich 4 Tage       2658 0.7644521   76.44521 100.00000
```