

Homework2 - Start with R

Instructions: For this assignment, you need to answer a couple questions with code and then take a screenshot of your working environment.

Submit the solutions including the URL to the screenshot in a doc/pdf to Brightspace.

```
pacman::p_load(tidyverse)
```

- 1) Use R to figure out how many elements in the vector below are **greater than 2** and then tell me what their **sum** (of the larger than 2 elements) is.

```
rooms <- c(1, 2, 4, 5, 1, 3, 1, NA, 3, 1, 3, 2, 1, NA, 1, 8, 3, 1, 4, NA, 1, 3, 1, 2, 1, 7, 1, 9, 3, NA)
```

```
# define the vector
rooms <- c(1, 2, 4, 5, 1, 3, 1, NA, 3, 1, 3, 2, 1, NA, 1, 8, 3, 1, 4, NA, 1, 3, 1, 2, 1, 7, 1, 9, 3, NA)
# getting rid of missing values
rooms <- na.omit(rooms)
# define vector with values > 2
rooms_large <- rooms[rooms > 2]
# finding the length of the vector
length(rooms_large)
```

```
## [1] 12
```

```
# taking the sum of this vector
sum(rooms_large)
```

```
## [1] 55
```

There are 12 number with a value > 2 and their sum is in total 55

- 2) What **type** of data is in the ‘rooms’ vector?

```
str(rooms)
```

```
## num [1:26] 1 2 4 5 1 3 1 3 1 3 ...
## - attr(*, "na.action")= 'omit' int [1:4] 8 14 20 30
```

It’s “numeric” (in fact it’s “integer”)

- 3) Submit the following image to Github: Inside your R Project (.Rproj), install the ‘tidyverse’ package and use the download.file() and read_csv() function to read the SAFI_clean.csv dataset into your R project as ‘interviews’ digital object (see instructions in <https://datacarpentry.org/r-socialsci/setup.html> and ‘Starting with Data’ section). Take a screenshot of your RStudio interface showing

- a) the line of code you used to create the object,
- b) the 'interviews' object in the Environment, and
- c) the file structure of your **R project** in the bottom right "Files" pane.

Save the screenshot as an image and put it in your **AUID_lastname_firstname** repository inside our Github organisation (github.com/Digital-Methods-HASS) or equivalent. Place **here** the URL leading to the screenshot in your repository.

```
SAFI_clean <- read_csv("SAFI_clean.csv")
```

```
## Rows: 131 Columns: 14
## -- Column specification -----
## Delimiter: ","
## chr (7): village, respondent_wall_type, memb_assoc, affect_conflicts, items...
## dbl (6): key_ID, no_membrs, years_liv, rooms, liv_count, no_meals
## dtm (1): interview_date
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

- 4) Challenge: If you managed to create your own Danish king dataset, use it. If not, use the one attached to this assignment (it might need to be cleaned up a bit). Load the dataset into R as a tibble. Calculate the mean() and median() duration of rule over time and find the three monarchs ruling the longest. How many days did they rule (accounting for transition year?)

```
# load the data
df_kings <- read_delim("kings.csv")
```

```
## Rows: 47 Columns: 4
## -- Column specification -----
## Delimiter: ";"
## chr (2): Kings, Yearasruler
## dbl (2): Start_date, End_date
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
# data cleaning + adding days of rule
df_kings <- na.omit(df_kings)
df_kings$`Yearasruler` <- as.integer(df_kings$`Yearasruler`)
df_kings$yearasruler <- df_kings$`Yearasruler`
df_kings$daysofrule <- df_kings$yearasruler*365
```

```
# finding mean
mean(df_kings$yearasruler)
```

```
## [1] 18.68182
```

```
# finding median
median(df_kings$yearasruler)
```

```
## [1] 14
```

```
# finding 3 longest rulers
df_kings %>%
  arrange(desc(yearasruler)) %>%
  slice(1:3)
```

```
## # A tibble: 3 x 6
##   Kings          Start_date End_date 'Yearasruler ' yearasruler daysofrule
##   <chr>          <dbl>    <dbl>          <int>         <int>         <dbl>
## 1 "Christian 4. "    1588      1648             60           60          21900
## 2 "Erik 7. af Pommern" 1396      1439             43           43          15695
## 3 "Christian 7. "    1766      1808             42           42          15330
```

The mean ruling time is 18.68 years and the median time is 14 years. The monarchs ruling the most time are Christian 4., Erik 7. af Pommern and Christian 7., who ruled 21.900 days, 15.695 days and 15.330 days respectively.