Homework 3

WIBE

8/30/2022

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

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

# Counting the sum of numbers greater than two excluding the NA's
length(na.omit(rooms) > 2)
```

```
## [1] 26
```

```
sum(na.omit(rooms) > 2)
```

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

Answer:

2) What type of data is in the 'rooms' vector?

'Class' function tells us the class of a vector:

```
class(rooms)
```

```
## [1] "numeric"
```

3) Installing tidyverse, loading data using read_csv()

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.

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.7     v dplyr   1.0.9
## v tidyr   1.2.0     v stringr 1.4.0
## v readr   2.1.2     v forcats 0.5.1

## -- Conflicts ---------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
interviews <- read_csv("/Users/WIBE/Desktop/CogSci/Cultural data science/Introduction course/Day 2/data/
```

```
## 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
## dttm (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, you 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?)

```
kings_and_queens <- read_csv(as_tibble("/Users/WIBE/Desktop/CogSci/Cultural data science/Introduction co
```

```
## Rows: 55 Columns: 7
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (3): Monarch, birth_date, death_date
## dbl (4): birth_year, death_year, start_reign, end_reign
##
## 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.
```

Calculating mean and median duration of rule over time

```
median((kings_and_queens$end_reign)-(kings_and_queens$start_reign), na.rm = T)
```

```
## [1] 16.5
```

```
mean((kings_and_queens$end_reign)-(kings_and_queens$start_reign), na.rm = T)
```

```
## [1] 19.33333
```

Mean reign length is: **47 years** Median reign length is: **43 years**

Finding the three monarchs ruling the longest

```
kings_and_queens %>%
  mutate(reign_length = end_reign-start_reign) %>%
  filter(reign_length > 42) %>%
  select(Monarch,reign_length)
```

```
## # A tibble: 3 x 2
##   Monarch            reign_length
##   <chr>                     <dbl>
## 1 Erik 7. af Pommern           43
## 2 Christian 4.                 60
## 3 Christian 9.                 43
```

The three longest reigning kings were

Monarch reign_length

1 Erik 7. af Pommern 43 2 Christian 4. 60 3 Christian 9. 43