

Homework 3- intro to R

Instead of writing down all the answers in a Word document like here, I could have documented everything in an R Markdown file, which I think would have been more organized (I will do that next time).

Assignment 1:

```
14 # Assignment 1
15 library(tidyverse)
16 rooms <- c(1,5,2,1,3,1,NA,3,1,3,2,1,NA,1,8,3,1,4,NA,1,3,1,2,1,7,1,NA)
17
18 # I start by removing NAs to make it more manageable
19 rooms_clean <- rooms[!is.na(rooms)]
20 rooms_clean
21
22 # Then i used the following function, to find all the values higher than two:
23 rooms_above_2 <- rooms_clean[rooms_clean>2]
24 rooms_above_2
25 length(rooms_above_2)
26
27 # I then used the lenght function to find the rooms bigger than 2: and as seen here below, it is 9 rooms
```

13:1 (Top Level) ↕

Console Terminal × Background Jobs ×

R 4.3.1 · ~/Library/CloudStorage/OneDrive-Aarhusuniversitet/Historie 3.semester/digitale metoder/opgave 3/ ↗

```
> rooms_clean <- rooms[!is.na(rooms)]
> rooms_clean
[1] 1 5 2 1 3 1 3 1 3 2 1 1 8 3 1 4 1 3 1 2 1 7 1
> # Then i used the following function:
> rooms_above_2 <- rooms_clean[rooms_clean>2]
> rooms_above_2
[1] 5 3 3 3 8 3 4 3 7
> length(rooms_above_2)
[1] 9
```

Assignment 2:

```
31 # Assignment 2
32
33 # I used the class function
34 class(rooms)
35
```

30:1 (Top Level) ↕

Console Terminal × Background Jobs ×

R 4.3.1 · ~/Library/CloudStorage/OneDrive-Aarhu

```
> class(rooms)
[1] "numeric"
```

Assignment 3:

```
43 # Assignment 3
44 # If you take the median of the entire 'rooms' vector instead of 'rooms_clean', the following happens
45 median(rooms)
46
47 # therefore you have to take the vector with the removed NA values, or you could just alter the code above, but i just do the same with another dataset
48 median(rooms_clean)
49
50 # as you can see under here, the median is 2
```

50:45 (Top Level) ↕

Console Terminal Background Jobs

R 4.3.1 · ~/Library/CloudStorage/OneDrive-Aarhusuniversitet/Historie 3.semester/digitale metoder/opgave 3/ ↗

```
> median(rooms)
[1] NA
> # therefore you have to take the vector with the removed NA values, or you could just write ignore in the code above
> median(rooms_clean)
[1] 2
```

Assignment 4:

Assignment a) I also uploaded it to GitHub: https://github.com/Digital-Methods-HASS/au728030_beuschau_mia

The screenshot shows the RStudio interface with the following components:

- Script Editor:** Contains R code for Assignment 4, starting with `#OPGAVE 4` and `library(tidyverse)`. The code reads a CSV file `data/SAFI_clean.csv` into a data frame named `interviews`.
- Console:** Shows the output of the R code. It displays the number of rows (131) and columns (14) of the `interviews` data frame. It also shows the column specification, including `chr (7): village, respondent_wall_type, memb_assoc, affect_conflicts, i...`, `dbl (6): key_ID, no_membrs, years_liv, rooms, liv_count, no_meals`, and `dtm (1): interview_date`. A message at the bottom suggests using `spec()` to retrieve the full column specification.
- Environment:** Shows the `Global Environment` with 148 MiB of memory used. It lists the `interviews` data frame with 131 observations and 14 variables.
- Files:** Lists the files in the project directory, including `..`, `.Rhistory` (5.6 KB), `data`, `Mia_kong.Rproj` (205 B), `opg1_4.R` (0 B), and `opgave konge.R` (875 B).

Assignment 5:

a) At få kongerækken ind i R

My file seems to have a mind of its own, so even though I had converted my Excel file into a comma-separated file, all my data ended up in one column, separated by semicolons. I noted it like this to get the data in

```
library(tidyverse)
kong <- read_csv("data/kongerokken.csv", na="NULL")
```

And this was how the data was organized in R:

	konge;fodsel;regering_start;regering_slut;dod;
1	gorm_den_gamle;NA;NA;958;NA;

However, I found a solution after a long struggle with CSV formats, and it is to write:

```
library(tidyverse)
kong <- read_delim("data/kongerokken2.csv", delim = ";")
```

And then the spreadsheet seems to be fine and separated correct:

```
> library(tidyverse)
> kong <- read_delim("data/kongerokken2.csv", delim = ";")
New names:
• `` -> `...6`
Rows: 54 Columns: 6
— Column specification —————
Delimiter: ";"
chr (1): konge
dbl (4): fodsel, regering_start, regering_slut, dod
lgl (1): ...6
```

	konge	fodsel	regering_start	regering_slut	dod
1	gorm_den_gamle	NA	NA	958	NA

b) Beregning af median og gennemsnit

Her var jeg igennem en del skridt, som kan ses nedenfor:

```
2 # I begin by loading my data,
3 # Here, I have made sure to place the CSV file in a subfolder of the project
4
5 library(tidyverse)
6 kong <- read_delim("data/kongerokken2.csv", delim = ";")
7 # I use the code above because R assumes that I intend to separate using ;
8
9 # I then maked the NA recognizable
10 kong <- read_delim("data/kongerokken2.csv", delim = ";", na="NA")
11 kong
12
13 # To calculate the mean and median of the reign duration,
14 # I need to calculate the number of years each king has spent on the throne.
15 # so i work towards making a new column
16
17 regstart<- kong[["regering_start"]]
18 regstart
19 regslut<- kong[["regering_slut"]]
20 regslut
21
22 # now i make a new column, by subtracting the two vectors above (the start of their reign and the end of their reign)
23 regper <- regslut-regstart
24 regper
25
26 #In order to calculate the median and mean, I will first remove the NA.
27 regper_clean <- regper[!is.na(regper)]
28 regper_clean
29
30 # Now I can calculate the median and mean
31 mean(regper_clean)
32 median(regper_clean)
33
46.1 (Top Level) :
```

Console

Terminal

Background Jobs

```
R 4.3.1 · ~/Library/CloudStorage/OneDrive-Aarhusuniversitet/Historie 3.semester/digitale metoder/opgave 3/ ↗
> regstart
[1] NA 958 987 1014 1018 1035 1042 1047 1074 1080 1086 1095 1104 1134 1137 1146 1146 1157 1182 1202 1241 1250 1252 1259 1286 1320 1330 1326 1340 1387 1376 1412 1440 1448 1481
[36] 1523 1513 1534 1559 1588 1648 1670 1699 1730 1746 1766 1808 1839 1848 1863 1906 1912 1947 1972
> regslut<- kong[["regering_slut"]]
> regslut
[1] 958 987 1014 1018 1035 1042 1047 1074 1080 1086 1095 1103 1134 1137 1146 1157 1157 1182 1202 1241 1250 1252 1259 1286 1319 1326 1332 1330 1375 1412 1387 1439 1448 1481 1513
[36] 1533 1523 1559 1588 1648 1670 1699 1730 1746 1766 1808 1839 1848 1863 1906 1912 1947 1972 NA
> # now i make a new column, by subtracting the two vectors above (the start of their reign and the end of their reign)
> regper <- regslut-regstart
> regper
[1] NA 29 27 4 17 7 5 27 6 6 9 8 30 3 9 11 11 25 20 39 9 2 7 27 33 6 2 4 35 25 11 27 8 33 32 10 10 25 29 60 22 29 31 16 20 42 31 9 15 43 6 35 25 NA
> #In order to calculate the median and mean, I will first remove the NA.
> regper_clean <- regper[!is.na(regper)]
> regper_clean
[1] 29 27 4 17 7 5 27 6 6 9 8 30 3 9 11 11 25 20 39 9 2 7 27 33 6 2 4 35 25 11 27 8 33 32 10 10 25 29 60 22 29 31 16 20 42 31 9 15 43 6 35 25
> # Now I can calculate the median and mean
> mean(regper_clean)
[1] 19.46154
> median(regper_clean)
[1] 18.5
```