

# Assignment 11:

*By Cecilie Lykke Knudsen, edits to question 3 made by Thorbjørn Rød and Oliver Boisen*

For this assignment, you need to

- a) present the R code you used to answer a few questions and then
- b) take a screenshot of your working environment.

## Questions:

1) Use R to figure out how many elements in the vector below are greater than 2.

```
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, 4, 3, 1, 7, 8, 2, 1, NA, 1, 1, 3)
```

After you've written the code we need to figure out how to remove NA.

How to remove NA:

**median(rooms)** → Here you get the answer NA

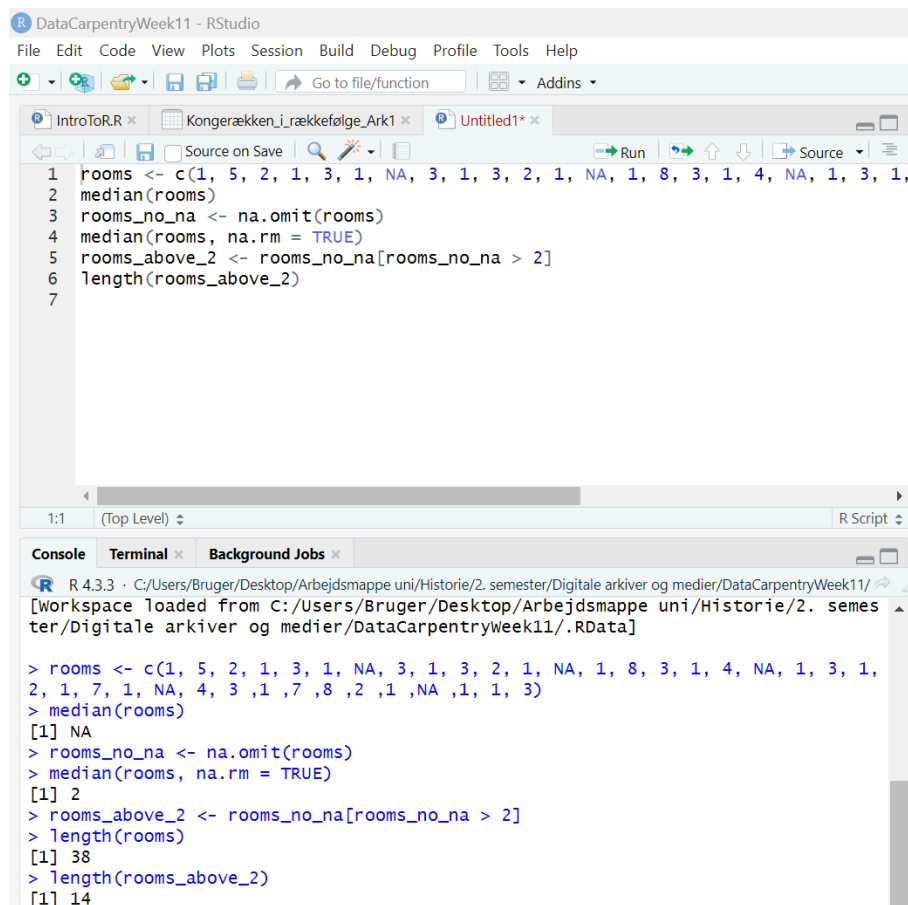
**rooms\_no\_na <- na.omit(rooms)** → To tell what needs to happen with →

**NAmedian(rooms, na.rm = TRUE)** → To turn off the NA-setting which is automatically applied in R

**rooms\_above\_2 <- rooms\_no\_na[rooms\_no\_na > 2]** → to tell → *rooms\_above\_2* is the same as → *rooms\_no\_na* which needs to be bigger than 2

**length(rooms\_above\_2)** → To find out how many rooms that are bigger than 2

The answer is therefore that there's 14 rooms above 2



The screenshot shows the RStudio interface. The script editor contains the following code:

```
1 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 median(rooms)
3 rooms_no_na <- na.omit(rooms)
4 median(rooms, na.rm = TRUE)
5 rooms_above_2 <- rooms_no_na[rooms_no_na > 2]
6 length(rooms_above_2)
7
```

The console shows the execution of the code:

```
R 4.3.3 · C:/Users/Bruger/Desktop/Arbejdsmappe uni/Historie/2. semester/Digitale arkiver og medier/DataCarpentryWeek11/
[workspace loaded from C:/Users/Bruger/Desktop/Arbejdsmappe uni/Historie/2. semester/Digitale arkiver og medier/DataCarpentryWeek11/.RData]

> 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, 4, 3, 1, 7, 8, 2, 1, NA, 1, 1, 3)
> median(rooms)
[1] NA
> rooms_no_na <- na.omit(rooms)
> median(rooms, na.rm = TRUE)
[1] 2
> rooms_above_2 <- rooms_no_na[rooms_no_na > 2]
> length(rooms)
[1] 38
> length(rooms_above_2)
[1] 14
```

2) Which function tells you the **type of data** the 'rooms' vector above contains?

the vector `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, 4, 3, 1, 7, 8, 2, 1, NA, 1, 1, 3)` is the type double. which I concluded by writing `class(rooms)`

```
1 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, 4, 3, 1)
2
3 rooms_no_na <- na.omit(rooms)
4 # 2.
5 median(rooms, na.rm = TRUE)
6
7 # 3.
8 rooms_above_2 <- rooms_no_na[rooms_no_na > 2]
9 length(rooms_above_2)
10
11 class(rooms)
12 |
```

R 4.3.3 · C:/Users/Asus PC/OneDrive - Aarhus universitet/Uni AU/2 semester/DAM\_dok/R og Rstudio/Week\_11\_assignment/Rstudio\_assignment/

```
> rooms_above_2 <- rooms_no_na[rooms_no_na > 2]
> length(rooms_above_2)
[1] 14
> # 2.
> median(rooms, na.rm = TRUE)
[1] 2
> # or
> rooms_no_na <- na.omit(rooms)
> rooms_no_na <- na.omit(rooms)
> rooms_above_2 <- rooms_no_na[rooms_no_na > 2]
> length(rooms_above_2)
[1] 14
> typeof(rooms)
[1] "double"
> class(rooms)
[1] "numeric"
> |
```

Which allowed me to find the type of the vector which was numeric

3) What is the result of running the **median()** function on the above 'rooms' vector?

Running the **median(rooms)** function simply provides us with the answer NA, as it cannot process the median of the sequence of numbers, due to the missing values. Running the **median(rooms\_no\_nas)** instead, provides us with the following answer

```
> # Median
> median(rooms)
[1] NA
> median(rooms_no_nas)
[1] 2
```

4) 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`

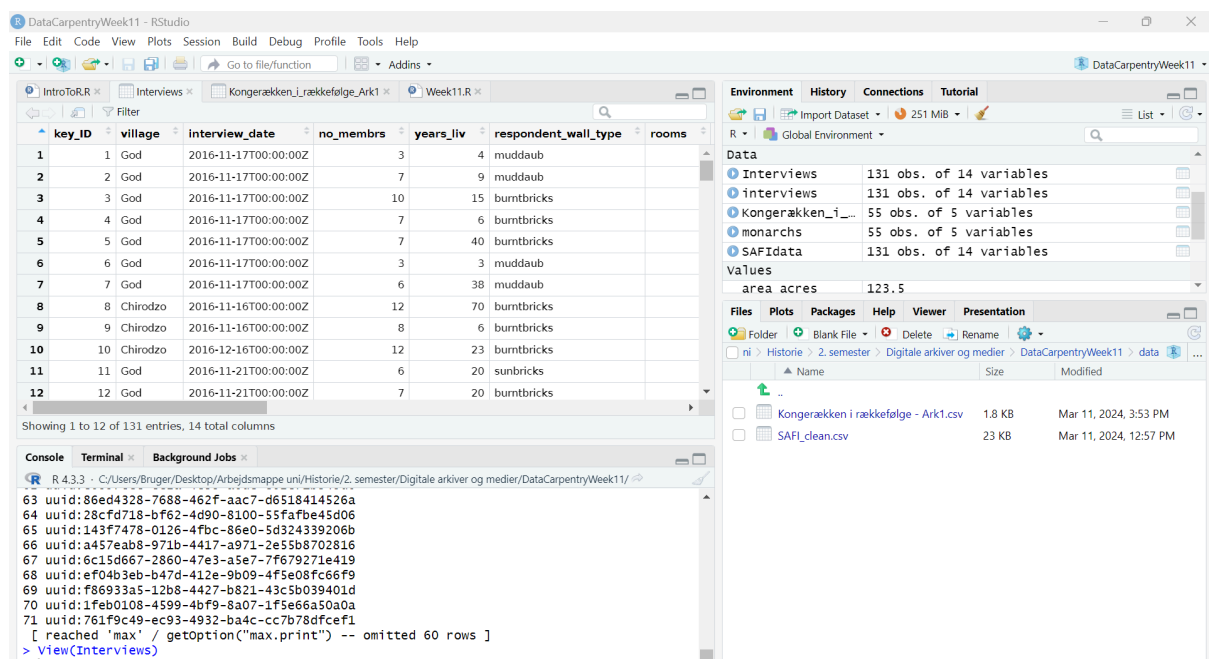
dataset into your R project as 'interviews' digital object (see instructions in <https://datacarpentry.org/r-socialsci/setup.html> and 'Starting with Data' section). Take a screenshot of your RStudio interface showing

a) the line of code you used to create the object,

```
#Spørgsmål 4
install.packages("tidyverse")

Interviews <- read.csv("data/SAFI_clean.csv", na = "NULL")
getwd()
Interviews
```

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 organization ([github.com/Digital-Methods-HASS](https://github.com/Digital-Methods-HASS)) or equivalent. Place here the URL leading to the screenshot in your repository.

5) **Challenge:** Tidy up your Danish monarchs dataset (you created last week) sufficiently so that you can load it into R as a tibble using the `read_csv()` and calculate the `mean()` and `median()` duration of their rule over time. Remember you can reload the dataset infinitely as you discover issues :)

```
monarchs <- read.csv("data/Kongerækken i rækkefølge - Ark1.csv")
monarchs
```

IntroToR.R x monarchs x Interviews x Week11.R x					
Filter					
	name	birth.year	death.year	start.of.reign	end.of.reign
1	gorm_den_gamle	908	958	936	958
2	harald_1_blatand	NA	987	NA	NA
3	svend_1_tveskæg	NA	1014	980	1014
4	harald_2	NA	NA	1014	1018
5	knud_1_den_store	995	1035	1018	1035
6	hardeknud	1020	1042	1035	1042
7	magnus_den_gode	1024	1047	1042	1047
8	svend_2_estridsen	NA	1076	1047	1074
9	harald_3_hen	NA	1080	1074	1080
10	knud_2_den_hellige	NA	1086	1080	1086
11	oluf_1_hunger	NA	1095	1086	1095
12	erik_1_ejegod	1056	1103	1095	1103
13	niels	NA	1134	1104	1134
14	erik_2_emune	NA	1137	1134	1137
15	erik_3_lam	NA	1146	1137	1146
16	svend_3_grathe	NA	1157	1146	1157
17	knud_3	NA	1157	1146	1157
18	valdemar_1_den_store	1131	1182	1157	1182
19	knud_4	1163	1202	1182	1202
20	valdemar_2_sejr	1170	1241	1202	1241
21	erik_4_plovpenning	1216	1250	1241	1250

Showing 1 to 21 of 55 entries, 5 total columns

**How to submit this homework?** Type up the solutions to the tasks (= code that got you the results + the results) , include the URL to the screenshot in a document/pdf and upload to Brightspace.