For this assignment, you need to present the results, the code you used to answer a few questions, and then take a screenshot of your working environment.

Submit a textfile with typed up solutions here OR upload the document with solutions and the screenshot to your repository on Github and provide here only your Github URL. Make sure your homework files are clearly marked and readily findable there.

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)

The function: rooms[rooms > 2] will figure out how many elements from the vector above is greater than 2

The following elements are greater than two: 5, 3, NA, 3, 3, NA, 8, 3, 4, NA, 3, 7, NA

To find the elements above two, without the NA's I used the following functions and named these elements rooms_no_na

First, I removed the NA's:
rooms_no_na <- rooms[!is.na(rooms)]

Then the vector looked like this:
rooms_no_na: 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 this function to find the elements greater than two:
rooms_no_na[rooms_no_na > 2]

The following elements are greater than two: 5 3 3 3 8 3 4 3 7

9 elements are greater than two

2) Which function tells you the type of data the 'rooms' vector above contains?
The function called class(rooms_no_na), which inspects the vector, will tell you which type of data 'rooms' are. **The answer is: numeric**
Even when I write: class(rooms), which contains the NA's, the answer is still numeric since NA is registered as missing data and the NA's are coloured blue like the numbers.

The function mode(rooms) and mode(rooms_no_na) gives me the same answer

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

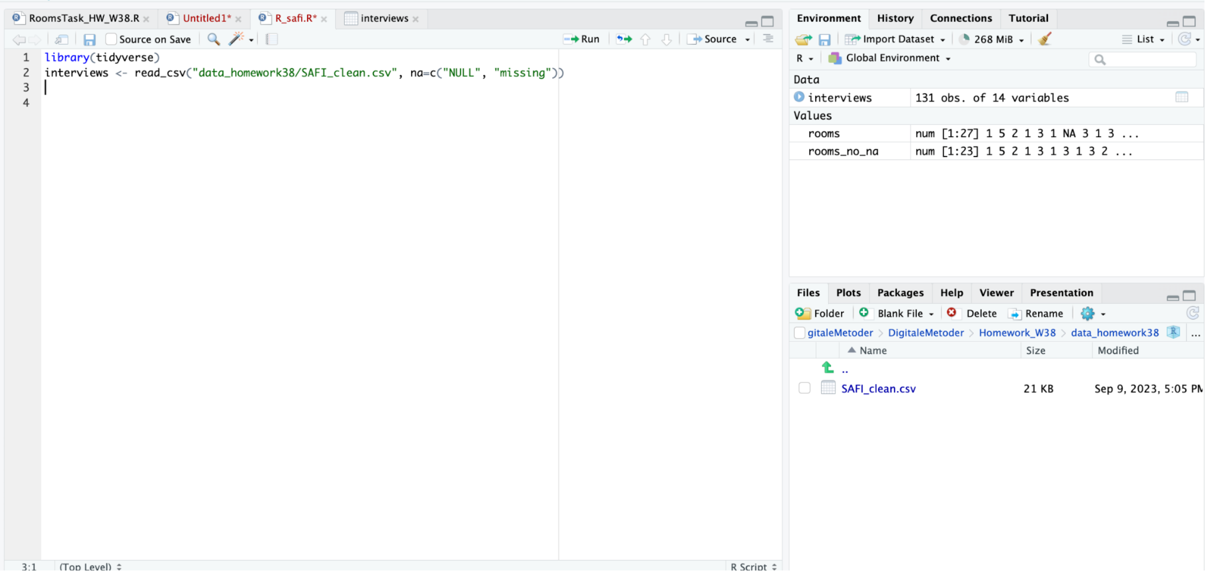
Because the vector contains NA and the function has not been told what to do about this missing data the result is NA.
To look at the median of all the other data in the vector other than the NA's you can write: median(rooms, na.rm = TRUE)
The result of this is: 2

But since I have made a vector called rooms_no_na which has removed the NA's this makes it even easier to find the median: median(rooms_no_na)
The result of this is: 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,
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.

Here is the URL to my GitHub and the screenshots (which I have also added to this document in case they are not available at GitHub)
https://github.com/Digital-Methods-HASS/AU721527_Sydow_Clara.git

The screenshots:

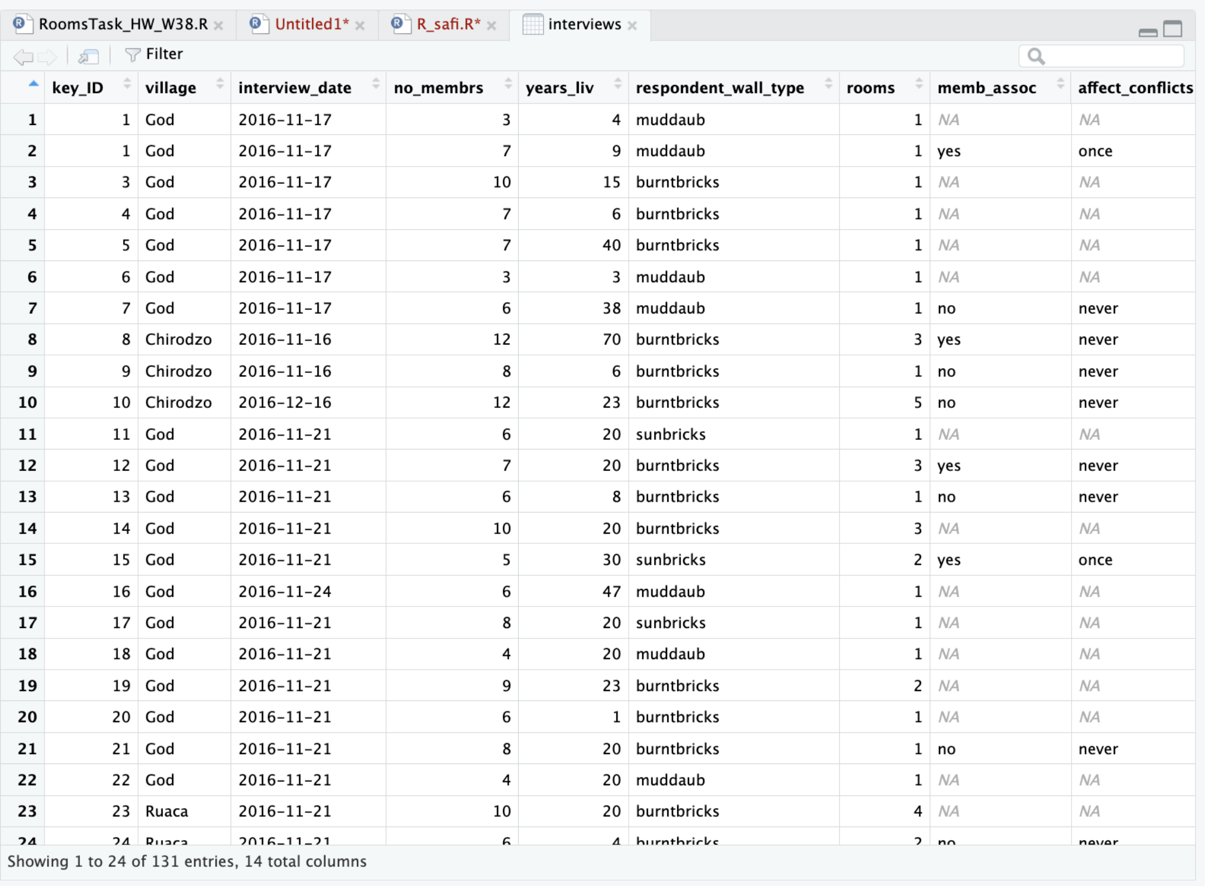


```r
library(tidyverse)
interviews <- read_csv("data_homework38/SAFI_clean.csv", na=c("NULL", "missing"))
```

```
Data
interviews     131 obs. of 14 variables
Values
  rooms        num [1:27] 1 5 2 1 3 1 NA 3 1 3 ...
  rooms_no_na  num [1:23] 1 5 2 1 3 1 3 1 3 2 ...
```

gitaleMetoder > DigitaleMetoder > Homework_W38 > data_homework38

Name	Size	Modified
..		
SAFI_clean.csv	21 KB	Sep 9, 2023, 5:05 PM

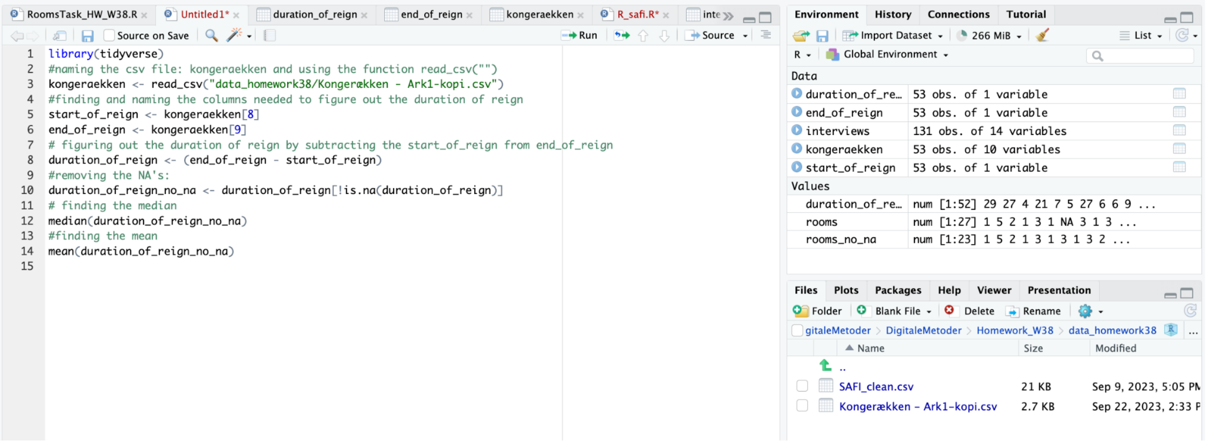


	key_ID	village	interview_date	no_membrs	years_liv	respondent_wall_type	rooms	memb_assoc	affect_conflicts
1	1	God	2016-11-17	3	4	muddaub	1	NA	NA
2	1	God	2016-11-17	7	9	muddaub	1	yes	once
3	3	God	2016-11-17	10	15	burntbricks	1	NA	NA
4	4	God	2016-11-17	7	6	burntbricks	1	NA	NA
5	5	God	2016-11-17	7	40	burntbricks	1	NA	NA
6	6	God	2016-11-17	3	3	muddaub	1	NA	NA
7	7	God	2016-11-17	6	38	muddaub	1	no	never
8	8	Chirodzo	2016-11-16	12	70	burntbricks	3	yes	never
9	9	Chirodzo	2016-11-16	8	6	burntbricks	1	no	never
10	10	Chirodzo	2016-12-16	12	23	burntbricks	5	no	never
11	11	God	2016-11-21	6	20	sunbricks	1	NA	NA
12	12	God	2016-11-21	7	20	burntbricks	3	yes	never
13	13	God	2016-11-21	6	8	burntbricks	1	no	never
14	14	God	2016-11-21	10	20	burntbricks	3	NA	NA
15	15	God	2016-11-21	5	30	sunbricks	2	yes	once
16	16	God	2016-11-24	6	47	muddaub	1	NA	NA
17	17	God	2016-11-21	8	20	sunbricks	1	NA	NA
18	18	God	2016-11-21	4	20	muddaub	1	NA	NA
19	19	God	2016-11-21	9	23	burntbricks	2	NA	NA
20	20	God	2016-11-21	6	1	burntbricks	1	NA	NA
21	21	God	2016-11-21	8	20	burntbricks	1	no	never
22	22	God	2016-11-21	4	20	muddaub	1	NA	NA
23	23	Ruaca	2016-11-21	10	20	burntbricks	4	NA	NA
24	24	Ruaca	2016-11-21	6	4	burntbricks	2	no	never

Showing 1 to 24 of 131 entries, 14 total columns

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 rule over time.

(description in the screenshot of the process)



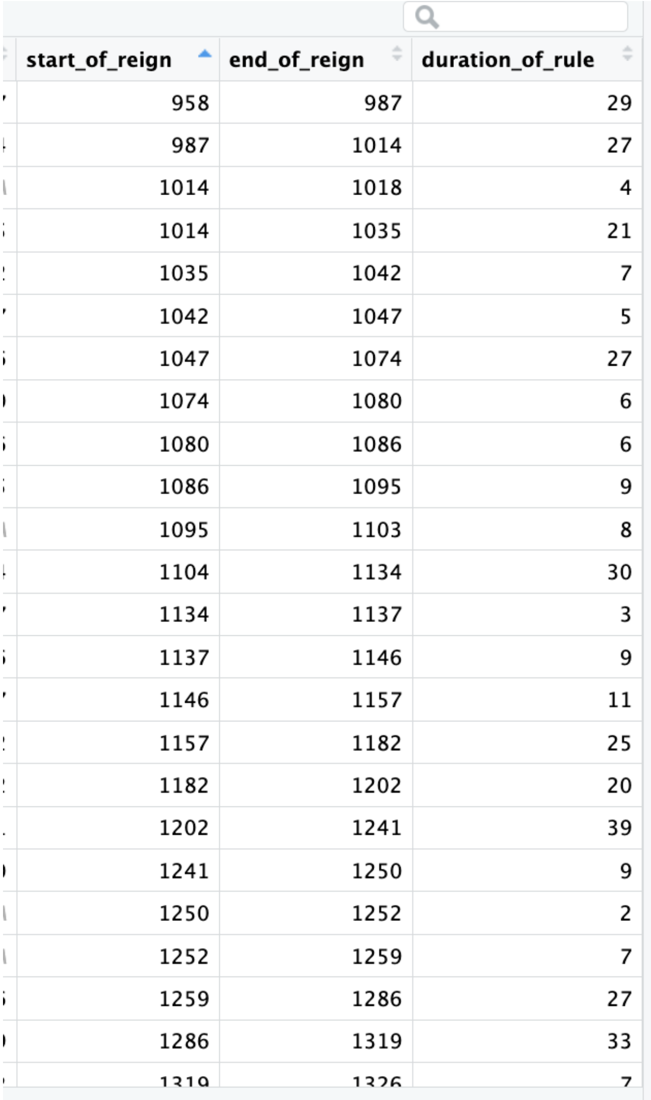
```r
library(tidyverse)
#naming the csv file: kongeraekken and using the function read_csv("")
kongeraekken <- read_csv("data_homework38/Kongerækken - Ark1-kopi.csv")
#finding and naming the columns needed to figure out the duration of reign
start_of_reign <- kongeraekken[8]
end_of_reign <- kongeraekken[9]
# figuring out the duration of reign by subtracting the start_of_reign from end_of_reign
duration_of_reign <- (end_of_reign - start_of_reign)
#removing the NA's:
duration_of_reign_no_na <- duration_of_reign[!is.na(duration_of_reign)]
# finding the median
median(duration_of_reign_no_na)
#finding the mean
mean(duration_of_reign_no_na)

```

```
Data
duration_of_re…   53 obs. of 1 variable
end_of_reign      53 obs. of 1 variable
interviews        131 obs. of 14 variables
kongeraekken      53 obs. of 10 variables
start_of_reign    53 obs. of 1 variable
Values
duration_of_re…   num [1:52] 29 27 4 21 7 5 27 6 6 9 ...
rooms             num [1:27] 1 5 2 1 3 1 NA 3 1 3 ...
rooms_no_na       num [1:23] 1 5 2 1 3 1 3 1 3 2 ...
```

gitaleMetoder > DigitaleMetoder > Homework_W38 > data_homework38

Name	Size	Modified
..		
SAFI_clean.csv	21 KB	Sep 9, 2023, 5:05 PM
Kongerækken - Ark1-kopi.csv	2.7 KB	Sep 22, 2023, 2:33 P

a section of the three columns used:



start_of_reign	end_of_reign	duration_of_rule
958	987	29
987	1014	27
1014	1018	4
1014	1035	21
1035	1042	7
1042	1047	5
1047	1074	27
1074	1080	6
1080	1086	6
1086	1095	9
1095	1103	8
1104	1134	30
1134	1137	3
1137	1146	9
1146	1157	11
1157	1182	25
1182	1202	20
1202	1241	39
1241	1250	9
1250	1252	2
1252	1259	7
1259	1286	27
1286	1319	33
1319	1326	7

The results:

median(duration_of_reign_no_na): 20.5 years

mean(duration_of_reign_no_na): 20.48077 years