W46: Start with R

1) Use R to figure out how many elements in the vector below are greater than 2.
rooms <- c(1, 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)

Explanation:
First I apply rooms with the numbers (and NAs) listed in the task. If I have to find all the numbers that are greater than 2 I have to sort out the NAs. This is done with the code: rooms[!is.na(rooms)] Now I make a new rooms called "rooms1" – This version of rooms only contains numbers, and as such has sorted out the "NA".

I can now apply the subsetting with the sqare brackets and use the following code to find all the numbers that are greater than 2: rooms1[rooms1>2]. I then get the answer: 3 3 3 8 3 4 3 7 – A total of 8 digits are greater than 2

Script:

```
rooms <-c(1, 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)

rooms1<- rooms[!is.na(rooms)]

rooms1[rooms1>2]
```

Console:

```
> rooms1[rooms1>2]
[1] 3 3 3 8 3 4 3 7
```

2) What type of data is in the 'rooms' vector?
I use the command class(rooms) and the answer is numeric

3) What is the result of running the median() function on the above 'rooms' vector?
a)The result is 1.5 using the median(rooms, na.rm=TRUE) function

```
> median(rooms, na.rm=TRUE)
[1] 1.5
```

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

```
105
106 #Working with spreadsheets:
107
108 ?read_csv()
109 interviews <- read_csv("./data/SAFI_clean.csv", na="NULL")
110 view(interviews)
111 dim(interviews) #Dimensions, how long and wide is your dataset
112 ncol(interviews) #number of coloums
113 head(interviews) #Finds the first 5
114 tail(interviews) #Finds the last 5
115 names(interviews)
116 glimpse(interviews)
117
```

135:4 (Top Level) R Script

Console Terminal Jobs

R 4.1.1 · ~/Digital methods/

```
── Conflicts ──────────────────────────────────────── tidyverse_conflicts() ──
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
> ?read_csv()
> read_csv("./data/SAFI_clean.csv")
Rows: 131 Columns: 14
── Column specification ────────────────────────────────────────────────────────
Delimiter: ","
chr  (7): village, respondent_wall_type, memb_assoc, affect_conflicts, items_owned, months_lack_food, instanceID
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.
# A tibble: 131 × 14
```

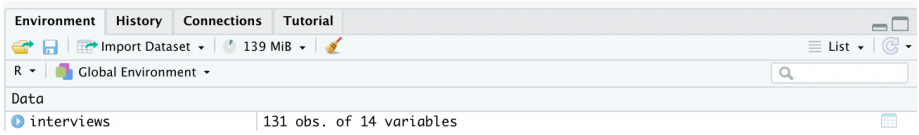
	key_ID	village	interview_date	no_membrs	years_liv	respondent_wall_type	rooms	memb_assoc	affect_conflicts	liv_count
	<dbl>	<chr>	<dttm>	<dbl>	<dbl>	<chr>	<dbl>	<chr>	<chr>	<dbl>
1	1	God	2016-11-17 00:00:00	3	4	muddaub	1	NULL	NULL	1
2	1	God	2016-11-17 00:00:00	7	9	muddaub	1	yes	once	3
3	3	God	2016-11-17 00:00:00	10	15	burntbricks	1	NULL	NULL	1
4	4	God	2016-11-17 00:00:00	7	6	burntbricks	1	NULL	NULL	2
5	5	God	2016-11-17 00:00:00	7	40	burntbricks	1	NULL	NULL	4
6	6	God	2016-11-17 00:00:00	3	3	muddaub	1	NULL	NULL	1
7	7	God	2016-11-17 00:00:00	6	38	muddaub	1	no	never	1
8	8	Chirodzo	2016-11-16 00:00:00	12	70	burntbricks	3	yes	never	2

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

a. interviews <- read_csv("./data/SAFI_clean.csv", na="NULL")

b) the 'interviews' object in the Environment

a.



c) the file structure of your R project in the bottom right "Files" pane.

a.