Assignment_4

Matilde Jacobsen

10/4/2020

Set working directory

As a standard I always start out by setting my working directory to the current folder:

```
#Setting working directory
setwd("/Users/matilde/Desktop/AU/Cultural Data Science/R/CDS2020_1")
```

Exercise 1:

Use R to figure out how many elements in the vector below are greater than 2 . (You need to filter out the NAs first)

```
#Loading in given vector
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)

#filter NA's
rooms_clean <- na.omit(rooms)

#get values in vector greater than 2
sum(rooms_clean > 2)
```

```
## [1] 8
```

Step-by-step: I loaded the vector into an object by using the assignment arrow "<-" Then I used the function na.omit to filter away all "NA's" from the vector and assigned it to a new object I called "rooms_clean" to leave the original vector in case I needed to use it later Finally, I used the function sum to get a count of what is written inside the parentheses, where I used the greater than sign ">" to indicate that I only wanted the counts of numbers greater than 2. Result: It showed me that there are 8 numbers in the vector rooms that are of a value greater than 2

Exercise 2:

What is the result of running median() function on the above 'rooms' vector? (again, best remove the NAs)

```
#running the median() function on the rooms vector
median(rooms, na.rm = T)
```

```
## [1] 1.5
```

Result: Using the generic function median() on the rooms vector gave me the value in the middle of the vector which supposedly is 1.5
Note: Instead of using the "rooms_clean" vector, I here specified in the function with the argument "na.rm = T" that I wanted it to remove NA's

Exercise 3:

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 script you used to create the object, b) the 'interviews' object in the Environment and the c) 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). Place here the URL leading to the screenshot in your repository.

```
#Installing packages
pacman::p_load(tidyverse)

#Downloading the SAFI_clean dataset from website given in the data carpentry instructions page
download.file("https://ndownloader.figshare.com/files/11492171",
              "data/SAFI_clean.csv", mode = "wb")

#Cretaing "interviews" from the SAFI_clean.csv file
interviews <- read.csv("Data/SAFI_clean.csv")
```

Notes: I use the package "pacman" to get the function "p_load", which both "installs" and saves the package in my "library". I can keep adding packages I want to install inside of the parantheses by adding a comma after the package name and type the name of another package I want to install.