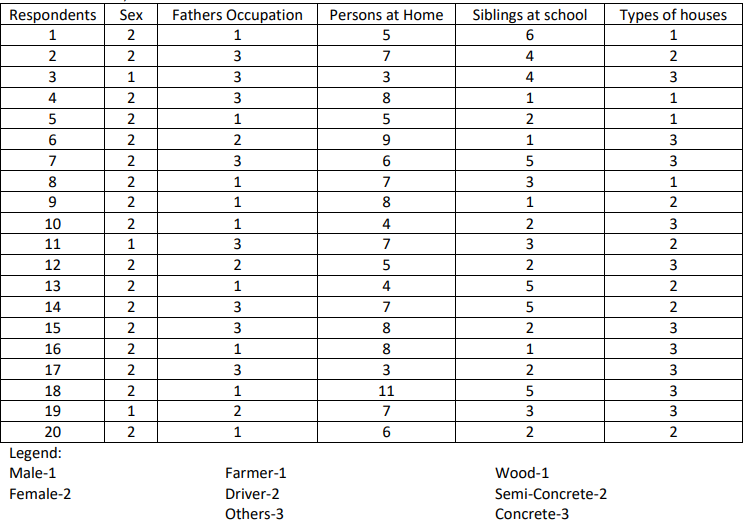
Worksheet-3b in R

**Instructions:**

* Use RStudio or the RStudio Cloud accomplish this worksheet.
* Save the R script as *RWorksheet\_lastname#3b.R*.
* On your own *GitHub repository*, push the R script, the Rmd file, as well as this pdf worksheet to the repo you have created before.
* Do not forget to comment your Git repo on our VLE
* Accomplish this worksheet by answering the questions being asked and writing the code manually.

1. Create a data frame using the table below.



| a. Write the codes.

dframe <- data.frame(

Respondents = 1:20,

Sex = c(2,2,1,2,2,2,2,2,

2,2,1,2,2,2,2,2,

2,2,1,2),

Fathers\_Occupation = c(1,3,3,3,1,2,3,

1,1,1,3,2,1,3,3,1,

3,1,2,1),

Persons\_Home = c(5,7,3,8,5,9,6,7,8

,4,7,5,4,7,8,8,3

,11,7,6),

Siblings\_school = c(6,4,4,1,2,

1,5,3,1,2,3,2,

5,5,2,1,2,5,3,2),

Types\_houses = c(1,2,3,1,1,

3,3,1,2,3,2,3,2,

2,3,3,3,3,3,2))

dframe

* 1. Describe the data. Get the structure or the summary of the data

> summary(dframe)

Respondents Sex Fathers\_Occupation Persons\_Home Siblings\_school

Min. : 1.00 Min. :1.00 Min. :1.00 Min. : 3.0 Min. :1.00

1st Qu.: 5.75 1st Qu.:2.00 1st Qu.:1.00 1st Qu.: 5.0 1st Qu.:2.00

Median :10.50 Median :2.00 Median :2.00 Median : 7.0 Median :2.50

Mean :10.50 Mean :1.85 Mean :1.95 Mean : 6.4 Mean :2.95

3rd Qu.:15.25 3rd Qu.:2.00 3rd Qu.:3.00 3rd Qu.: 8.0 3rd Qu.:4.25

Max. :20.00 Max. :2.00 Max. :3.00 Max. :11.0 Max. :6.00

Types\_houses

Min. :1.0

1st Qu.:2.0

Median :2.5

Mean :2.3

3rd Qu.:3.0

Max. :3.0

* 1. Is the mean number of siblings attending is 5?

mean(dframe$Siblings\_school)

[1] 2.95

* 1. Extract the 1st two rows and then all the columns using the subsetting functions. Write the codes and its output.

ext1 <- subset(dframe[c(1:2), ])

ext1

> ext1

Respondents Sex Fathers\_Occupation Persons\_Home Siblings\_school Types\_houses

1 1 2 1 5 6 1

2 2 2 3 7 4 2

* 1. Extract 3rd and 5th row with 2nd and 4th column. Write the codes and its result.

ext2 <- subset(dframe[c(3,5), c(2,4)])

ext2

> ext2 <- subset(dframe[c(3,5), c(2,4)])

> ext2

Sex Persons\_Home

3 1 3

5 2 5

* 1. Select the variable types of houses then store the vector that results as types\_houses. Write the codes.

ext3 <- subset(dframe, select = Types\_houses)

ext3

> ext3 <- subset(dframe, select = Types\_houses)

> ext3

Types\_houses

1 1

2 2

3 3

4 1

5 1

6 3

7 3

8 1

9 2

10 3

11 2

12 3

13 2

14 2

15 3

16 3

17 3

18 3

19 3

20 2

* 1. Select only all Males respondent that their father occupation was farmer. Write the codes and its output.

ext4 <- subset(dframe,

+ Sex == 1 & Fathers\_Occupation < 1,

+ select = c(Sex, Fathers\_Occupation),

+ drop = FALSE)

> ext4

[1] Sex Fathers\_Occupation

<0 rows> (or 0-length row.names)

* 1. Select only all females respondent that have greater than or equal to 5 number of siblings attending school. Write the codes and its outputs.

ext5 <- subset(dframe,

Sex == 2 & Siblings\_school >= 5,

select = c(Sex, Siblings\_school),

drop = FALSE)

> ext5

Sex Siblings\_school

1 2 6

7 2 5

13 2 5

14 2 5

18 2 5

1. Write a R program to create an empty data frame. Using the following codes:

df = data.frame(Ints=integer(),

Doubles=double(), Characters=character(), Logicals=logical(),

Factors=factor(), stringsAsFactors=FALSE)

print("Structure of the empty dataframe:") print(str(df))

a. Describe the results.

df = data.frame(Ints=integer(),

Doubles=double(), Characters=character(),

Logicals=logical(),

Factors=factor(),

stringsAsFactors=FALSE)

>

> print("Structure of the empty dataframe:")

[1] "Structure of the empty dataframe:"

> print(str(df))

'data.frame': 0 obs. of 5 variables:

$ Ints : int

$ Doubles : num

$ Characters: chr

$ Logicals : logi

$ Factors : Factor w/ 0 levels:

NULL

1. Interpret the graph.

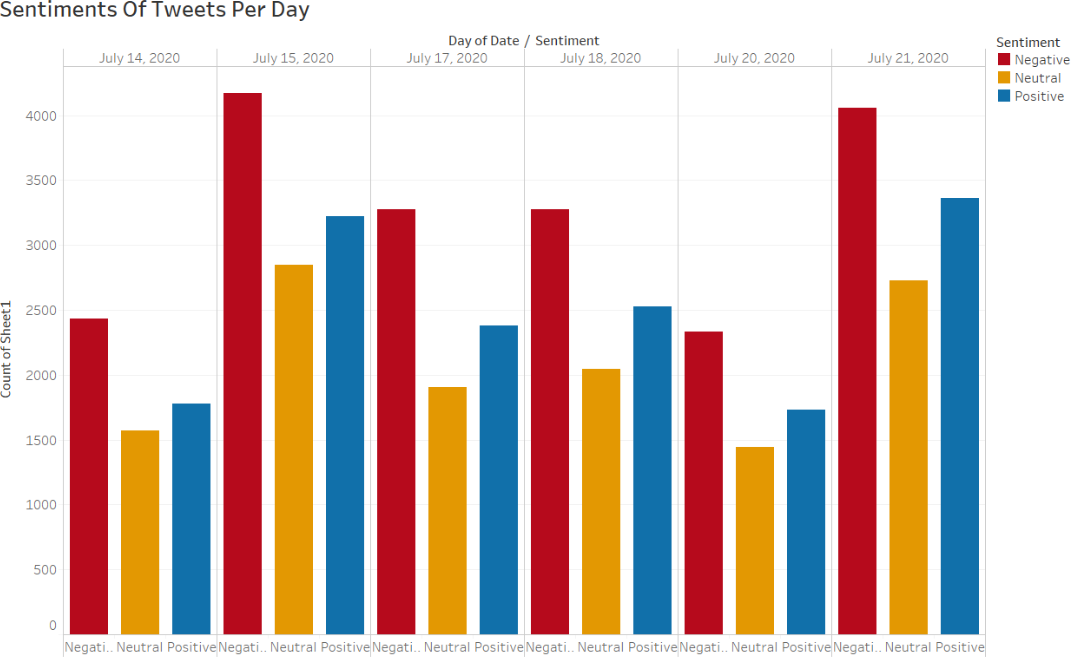


Figure 1: Sentiments of Tweets per day - Donald Trump