Worksheet-3b in R

Instructions:

- Use RStudio or the posit(RStudio) Cloud accomplish this worksheet.

- Inside the folder worksheet#3, create an .Rmd (R Markdown) for this worksheet and saved it as *RWorksheet_lastname#3b.Rmd*
- **Knit to pdf** to render a pdf file.
- On your own *GitHub repository*, push the .Rmd file, as well as the pdf worksheet knitted 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.

   b. Describe the data. Get the structure or the summary of the data

Respondents	Sex	Fathers Occupation	Persons at Home	Siblings at school	Types of houses
1	2	1	5	6	1
2	2	3	7	4	2
3	1	3	3	4	3
4	2	3	8	1	1
5	2	1	5	2	1
6	2	2	9	1	3
7	2	3	6	5	3
8	2	1	7	3	1
9	2	1	8	1	2
10	2	1	4	2	3
11	1	3	7	3	2
12	2	2	5	2	3
13	2	1	4	5	2
14	2	3	7	5	2
15	2	3	8	2	3
16	2	1	8	1	3
17	2	3	3	2	3
18	2	1	11	5	3
19	1	2	7	3	3
20	2	1	6	2	2

Legend:

Male-1	Farmer-1	Wood-1
Female-2	Driver-2	Semi-Concrete-2
	Others-3	Concrete-3

Figure 1: R Chunk

c. Is the mean number of siblings attending is 5?

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

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

f. Select the variable types of houses then store the vector that results as types_houses. Write the codes.

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

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

2. 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.

3. Create a `.csv` file of this. Save it as `HouseholdData.csv`

Respondents	Sex	Fathers Occupation	Persons at Home	Siblings at School	Types of Houses
1	Male	1	5	2	Wood
2	Female	2	7	3	Congrete
3	Female	3	3	0	Congrete
4	Male	3	8	5	Wood
5	Male	1	6	2	Semi-congrete
6	Female	2	4	3	Semi-congrete
7	Female	2	4	1	Wood
8	Male	3	2	2	Semi-congrete
9	Female	1	11	6	Semi-congrete
10	Male	3	6	2	Congrete

Figure 2: Figure 2: Sentiment Analysis

a. Import the csv file into the R environment. Write the codes.

b. Convert the `Sex` into factor using `factor()` function and change it into integer.[Legend: Male = 1 and Female = 2]. Write the R codes and its output.

c. Convert the `Type of Houses` into factor and change it into integer. [Legend: Wood = 1; Congrete = 2; Semi-Congrete = 3]. Write the R codes and its output.

d. On father's occupation, factor it as Farmer = 1; Driver = 2; and Others = 3. What is the R code and its output?

e. Select only all females respondent that has a father whose occupation is driver. Write the codes and its output.

f. Select the respondents that have greater than or equal to 5 number of siblings attending school. Write the codes and its output.

4. Interpret the graph.

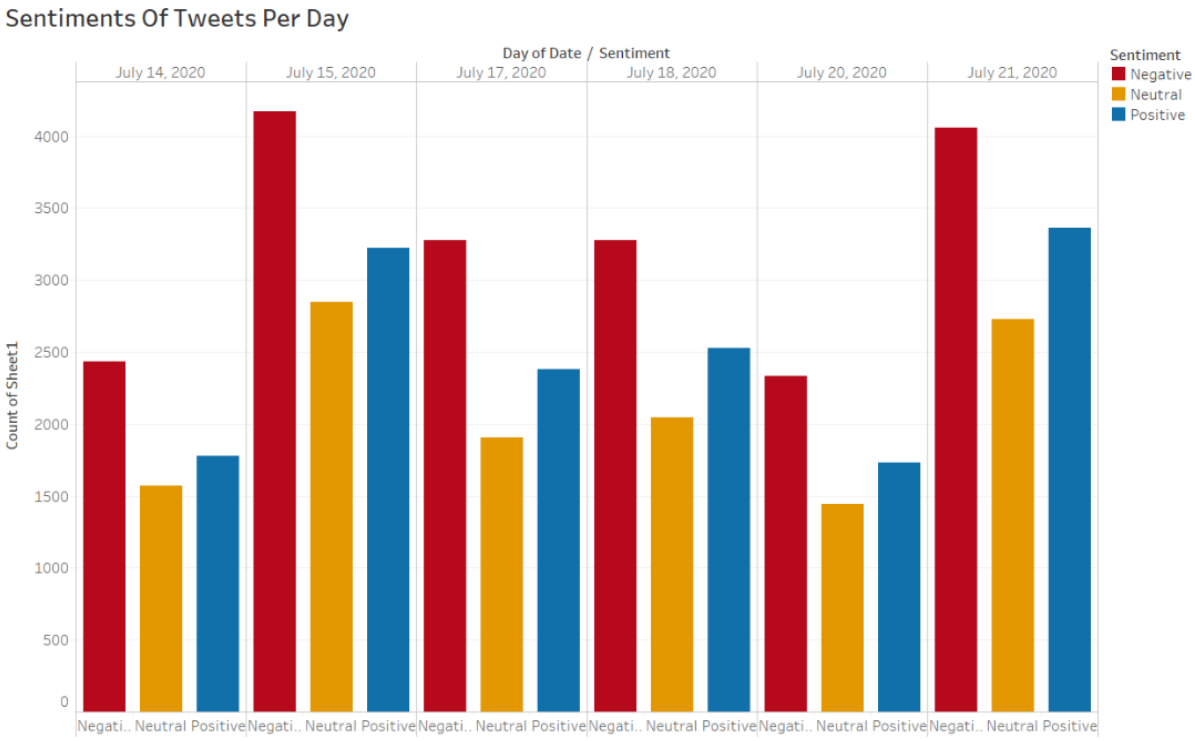

Figure 3: Figure 2: Sentiment Analysis