Worksheet #3b

Grace Anne E. Capanang

2022-11-20

1a. Create a data frame using the table below. Write the codes.

```
Respondents <- seq(1:20)
Sex <- rep(c(2, 1, 2, 1, 2, 1, 2), c(2, 1, 7, 1, 7, 1, 1))
Fathers_Occupation <- c(1, 3, 3, 3, 1, 2, 3, 1, 1, 1, 3, 2, 1, 3, 3, 1, 3, 1, 2, 1)
Persons_at_Home <- c(5, 7, 3, 8, 5, 9, 6, 7, 8, 4, 7, 5, 4, 7, 8, 8, 3, 11, 7, 6 )
Siblings_at_School <- c(6, 4, 4, 1, 2, 1, 5, 3, 1, 2, 3, 2, 5, 5, 2, 1, 2, 5, 3, 2)
Types_of_Houses <- c(1, 2, 3, 1, 1, 3, 3, 1, 2, 3, 2, 3, 2, 2, 3, 3, 3, 3, 3, 2)


Profile <- data.frame (
  Respondents,
  Sex,
  Fathers_Occupation,
  Persons_at_Home,
  Siblings_at_School,
  Types_of_Houses
)
Profile
```

```
##    Respondents Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 1            1   2                  1               5                  6
## 2            2   2                  3               7                  4
## 3            3   1                  3               3                  4
## 4            4   2                  3               8                  1
## 5            5   2                  1               5                  2
## 6            6   2                  2               9                  1
## 7            7   2                  3               6                  5
## 8            8   2                  1               7                  3
## 9            9   2                  1               8                  1
## 10          10   2                  1               4                  2
## 11          11   1                  3               7                  3
## 12          12   2                  2               5                  2
## 13          13   2                  1               4                  5
## 14          14   2                  3               7                  5
## 15          15   2                  3               8                  2
## 16          16   2                  1               8                  1
## 17          17   2                  3               3                  2
## 18          18   2                  1              11                  5
## 19          19   1                  2               7                  3
## 20          20   2                  1               6                  2
```

```
##    Types_of_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
```

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

#Answer: The data consist of number of respondents and their profile, common important matters regarding to a person's personal profile such as their sex, father's occupation, number of persons at home, siblings they have attending the school, and the type of house they live in. There is a legend which serve as a guide to the variables' data: (1) for the variable 'Sex' number 1 implied to Male and number 2 for Female, (2) for 'Father's Occupation' the number 1 is for Farmer, 2 for Driver and 3 for other professions aside from the two first mentioned, and (3) for the 'Types of Houses' 1 is for house made of Wood, 2 for Semi-concrete, and 3 for Concrete ones.

```
summary(Profile)
```

```
##   Respondents          Sex        Fathers_Occupation Persons_at_Home
##   Min.   : 1.00   Min.   :1.00   Min.   :1.00        Min.   : 3.0
##   1st Qu.: 5.75   1st Qu.:2.00   1st Qu.:1.00        1st Qu.: 5.0
##   Median :10.50   Median :2.00   Median :2.00        Median : 7.0
##   Mean   :10.50   Mean   :1.85   Mean   :1.95        Mean   : 6.4
##   3rd Qu.:15.25   3rd Qu.:2.00   3rd Qu.:3.00        3rd Qu.: 8.0
##   Max.   :20.00   Max.   :2.00   Max.   :3.00        Max.   :11.0
##   Siblings_at_School Types_of_Houses
##   Min.   :1.00        Min.   :1.0
##   1st Qu.:2.00        1st Qu.:2.0
##   Median :2.50        Median :2.5
##   Mean   :2.95        Mean   :2.3
##   3rd Qu.:4.25        3rd Qu.:3.0
##   Max.   :6.00        Max.   :3.0
```

1c Is the mean number of siblings attending is 5?

```
siblingsAtSchoolMean <- mean(Siblings_at_School)
siblingsAtSchoolMean
```

```
## [1] 2.95
```

#Answer: No, the mean number of siblings attending school is 2.95

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

```
Row1stAllCol <- subset(Profile[1:2, 1:6])
Row1stAllCol
```

```
##   Respondents Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 1           1   2                  1               5                  6
## 2           2   2                  3               7                  4
##   Types_of_Houses
## 1               1
## 2               2
```

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

```
RowThrFifthColSecoFourth <- subset (Profile[c(3,5), c(2,4)])
RowThrFifthColSecoFourth
```

```
##   Sex Persons_at_Home
## 3   1               3
## 5   2               5
```

1f. Select the variable types of houses then store the vector that results as types_houses. Write the codes.

```
types_houses <- Profile$Types_of_Houses
types_houses
```

```
##  [1] 1 2 3 1 1 3 3 1 2 3 2 3 2 2 3 3 3 3 3 2
```

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

```
farmer <- subset(Profile, Fathers_Occupation == '1')
farmer
```

```
##    Respondents Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 1            1   2                  1               5                  6
## 5            5   2                  1               5                  2
## 8            8   2                  1               7                  3
## 9            9   2                  1               8                  1
## 10          10   2                  1               4                  2
## 13          13   2                  1               4                  5
## 16          16   2                  1               8                  1
## 18          18   2                  1              11                  5
## 20          20   2                  1               6                  2
##    Types_of_Houses
## 1                1
## 5                1
```

```
## 8                    1
## 9                    2
## 10                   3
## 13                   2
## 16                   3
## 18                   3
## 20                   2
```

```
MalesRespondentFarmerFathers <- subset(farmer, Sex == '1')
MalesRespondentFarmerFathers
```

```
## [1] Respondents       Sex                 Fathers_Occupation Persons_at_Home
## [5] Siblings_at_School Types_of_Houses
## <0 rows> (or 0-length row.names)
```

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

```
Females_respondent <- subset(Profile, Sex == '2')
Females_respondent
```

```
##    Respondents Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 1            1   2                  1               5                  6
## 2            2   2                  3               7                  4
## 4            4   2                  3               8                  1
## 5            5   2                  1               5                  2
## 6            6   2                  2               9                  1
## 7            7   2                  3               6                  5
## 8            8   2                  1               7                  3
## 9            9   2                  1               8                  1
## 10          10   2                  1               4                  2
## 12          12   2                  2               5                  2
## 13          13   2                  1               4                  5
## 14          14   2                  3               7                  5
## 15          15   2                  3               8                  2
## 16          16   2                  1               8                  1
## 17          17   2                  3               3                  2
## 18          18   2                  1              11                  5
## 20          20   2                  1               6                  2
##    Types_of_Houses
## 1                1
## 2                2
## 4                1
## 5                1
## 6                3
## 7                3
## 8                1
## 9                2
## 10               3
## 12               3
## 13               2
## 14               2
```

```
## 15               3
## 16               3
## 17               3
## 18               3
## 20               2
```

```r
FemaleSiblingsMoreThanFive <- subset(Females_respondent, Siblings_at_School >= '5')
FemaleSiblingsMoreThanFive
```

```
##     Respondents Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 1             1   2                  1               5                  6
## 7             7   2                  3               6                  5
## 13           13   2                  1               4                  5
## 14           14   2                  3               7                  5
## 18           18   2                  1              11                  5
##     Types_of_Houses
## 1                 1
## 7                 3
## 13                2
## 14                2
## 18                3
```

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

```r
df <- data.frame (
  Ints = integer(),
  Doubles=double(),
  Characters = character(),
  Logicals=logical(),
  stringsAsFactors = FALSE
)

print ("Structures of the empty dataframe: ")
```

```
## [1] "Structures of the empty dataframe: "
```

```r
print(str(df))
```

```
## 'data.frame':    0 obs. of  4 variables:
##  $ Ints      : int
##  $ Doubles   : num
##  $ Characters: chr
##  $ Logicals  : logi
## NULL
```

2a. Describe the results.

#Answer: The output in this item is: there is no objects inside the data frame. There are four variables namely 'Ints', 'Doubles', 'Characters' and 'Logicals' but there are no data inside them. As it is, an empty dataframe.

#3. Interpret the graph. Answer:

My interpretation of the graph about the tweets' sentiments per day of Donald Trump is that each day the count of negative sentiments dominates; means there is much more negative sentiments than positive and neutrals. And then it was followed by positive, and then neutral in ranking. In conclusion a lot of twitter user does not like Donald Trump as a politician, hence the negative opinions of him.