Worksheet-3b in R

Daisy Ann O. Pajarillo

2022-11-14

1. Create a data frame using the table below.

a. Write the codes.

```
Respondents <- (1:20)
Respondents
```

```
##  [1]  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
```

```
sex <- c(2,2,1,2,2,2,2,2,2,2,1,2,2,2,2,2,2,2,1,2)
sex
```

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

```
father_ocpt <- c(1,3,3,3,1,2,3,1,1,1,3,2,1,3,3,1,3,1,2,1)
father_ocpt
```

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

```
person_home <- c(5,7,3,8,5,9,6,7,8,4,7,5,4,7,8,8,3,11,7,6)
person_home
```

```
##  [1]  5  7  3  8  5  9  6  7  8  4  7  5  4  7  8  8  3 11  7  6
```

```
sibling <- c(6,4,4,1,2,1,5,3,1,2,3,2,5,5,2,1,2,5,3,2)
sibling
```

```
##  [1] 6 4 4 1 2 1 5 3 1 2 3 2 5 5 2 1 2 5 3 2
```

```
house <- c(1,2,3,1,1,3,3,1,2,3,2,3,2,2,3,3,3,3,3,2)
house
```

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

```
A <- data.frame(Respondents,sex,father_ocpt,person_home,sibling,house)
A
```

```
##    Respondents sex father_ocpt person_home sibling house
## 1            1   2           1           5       6     1
## 2            2   2           3           7       4     2
## 3            3   1           3           3       4     3
## 4            4   2           3           8       1     1
## 5            5   2           1           5       2     1
## 6            6   2           2           9       1     3
## 7            7   2           3           6       5     3
## 8            8   2           1           7       3     1
## 9            9   2           1           8       1     2
## 10          10   2           1           4       2     3
## 11          11   1           3           7       3     2
```

```
## 12          12   2           2           5       2     3
## 13          13   2           1           4       5     2
## 14          14   2           3           7       5     2
## 15          15   2           3           8       2     3
## 16          16   2           1           8       1     3
## 17          17   2           3           3       2     3
## 18          18   2           1          11       5     3
## 19          19   1           2           7       3     3
## 20          20   2           1           6       2     2
```

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

```
It computes summary statistics of data including the mean,
median, and maximum.
```

```
summary(A)
```

```
##   Respondents         sex        father_ocpt    person_home      sibling
##  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
##      house
##  Min.   :1.0
##  1st Qu.:2.0
##  Median :2.5
##  Mean   :2.3
##  3rd Qu.:3.0
##  Max.   :3.0
```

c. Is the mean number of siblings attending is 5? Answer: No

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

```
r <- subset(A[1:2, 1:6, drop= FALSE])
r
```

```
##   Respondents sex father_ocpt person_home sibling house
## 1           1   2           1           5       6     1
## 2           2   2           3           7       4     2
```

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

```
r2 <- subset(A[c(3,5), c(2,4)])
r2
```

```
##   sex person_home
## 3   1           3
## 5   2           5
```

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

```
r3 <- subset(A[c(1:20), c(2,6)])
r3
```

```
##    sex house
## 1    2     1
```

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

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

```
sel <- subset(A[c(1:20), c(2,3)] )
sel
```

```
##    sex father_ocpt
## 1    2           1
## 2    2           3
## 3    1           3
## 4    2           3
## 5    2           1
## 6    2           2
## 7    2           3
## 8    2           1
## 9    2           1
## 10   2           1
## 11   1           3
## 12   2           2
## 13   2           1
## 14   2           3
## 15   2           3
## 16   2           1
## 17   2           3
## 18   2           1
## 19   1           2
## 20   2           1
```

```
Male <- sel[A$father_ocpt=='1',]
Male
```

```
##    sex father_ocpt
## 1    2           1
## 5    2           1
## 8    2           1
## 9    2           1
```

```
## 10   2           1
## 13   2           1
## 16   2           1
## 18   2           1
## 20   2           1
```

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.

```
sel2 <- subset(A[c(1:20), c(2,5)])
sel2
```

```
##    sex sibling
## 1    2       6
## 2    2       4
## 3    1       4
## 4    2       1
## 5    2       2
## 6    2       1
## 7    2       5
## 8    2       3
## 9    2       1
## 10   2       2
## 11   1       3
## 12   2       2
## 13   2       5
## 14   2       5
## 15   2       2
## 16   2       1
## 17   2       2
## 18   2       5
## 19   1       3
## 20   2       2
```

```
Female <- sel2[A$sibling>='5',]
Female
```

```
##    sex sibling
## 1    2       6
## 7    2       5
## 13   2       5
## 14   2       5
## 18   2       5
```

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:")
```

```
## [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
```

a. Describe the results.

```
It give information of the class of the object.
```

3. Interpret the graph

Answer: The graph represents that people's negative sentiments each day towards Donald Trump from the date July 14, 2021 to July 21, 2021 are higher than positive and neutral