Worksheet 3b

Ben Robert T. Salve.

2022-11-23

#Ben Robert T. Salve. BSIT 2-A

#1) ##a.

```
data_frame <- 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_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))
data_frame
```

```
##    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
```

##b.

```
summary(data_frame)
```

```
##   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
```

The summary displayed the Min, 1st Qu., Median, Mean, 3rd Qu., and Max.

The data has male and female respondents, fathers occupation, persons at home, and

siblings at school. This is a survey data.

##c. ##Answer: No, siblings at school has 2.95 mean.

##d.

```
subset(data_frame[1:2, ])
```

```
##   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
```

##e.

```
subset(data_frame[3:5, 2:4])
```

```
##   Sex Fathers_occupation Persons_at_home
## 3   1                  3               3
## 4   2                  3               8
## 5   2                  1               5
```

##f.

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
types_houses <- select(data_frame, Types_of_houses)
types_houses
```

```
##    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
```

##g.

```
respondents_and_father <- subset(data_frame[c(1:20),c(1:3)])
respondents_and_father
```

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

```
male_only <- respondents_and_father[data_frame$Sex == '1',]
male_only
```

```
##    Respondents Sex Fathers_occupation
## 3            3   1                  3
## 11          11   1                  3
## 19          19   1                  2
```

The output for this code: There is no male respondent with a father that is a farmer.

##h.

```
data1 <- subset(data_frame[c(1:20),c(1,2,5)])
data1
```

```
##    Respondents Sex Siblings_at_school
## 1            1   2                  6
## 2            2   2                  4
## 3            3   1                  4
## 4            4   2                  1
## 5            5   2                  2
## 6            6   2                  1
## 7            7   2                  5
## 8            8   2                  3
## 9            9   2                  1
```

```
## 10          10   2                 2
## 11          11   1                 3
## 12          12   2                 2
## 13          13   2                 5
## 14          14   2                 5
## 15          15   2                 2
## 16          16   2                 1
## 17          17   2                 2
## 18          18   2                 5
## 19          19   1                 3
## 20          20   2                 2
```

```
female_only <- data1[data_frame$Sex == '2',]
female_only
```

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

```
call <- data_frame[,5] >= 5
call
```

```
##  [1]  TRUE FALSE FALSE FALSE FALSE FALSE  TRUE FALSE FALSE FALSE FALSE FALSE
## [13]  TRUE  TRUE FALSE FALSE FALSE  TRUE FALSE FALSE
```

```
sum(call)
```

```
## [1] 5
```

```
data1[call,]
```

```
##    Respondents Sex Siblings_at_school
## 1            1   2                 6
## 7            7   2                 5
## 13          13   2                 5
## 14          14   2                 5
## 18          18   2                 5
```

The output for this code: There are 5 female respondents

that have greater than or equal to 5 number of siblings attending school.

#2)

```
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. ## The results it displayed whats inside the data frame which are empty, 0 obs. of ## 5 variables. Instead of displaying <0 rows> (or 0-length row.names) it displayed: ## 'data.frame': 0 obs. of 5 variables: ## $ Ints : int ## $ Doubles : num ## $ Characters: chr ## $ Logicals : logi ## $ Factors : Factor w/ 0 levels: ## NULL ## Due to print(str(df)) it displayed the varibles of the dataset in vertical with the ## following functions.

#3)

The title of the bar graph is sentiments of tweets per day. It has a legend at the right

side, red for negative, yellow for neutral, and blue for positive. In day 1

July 14, 2020 the negative sentiments almost reach 2,500. In day 2 July 15, 2020 the

negative sentiments sky rocketed to 4,000 plus. While in day 3 and 4 negative sentiments

went down around 3,000 plus. Then day 5 it went down again to 2,000 plus then went up at

day 6. The graph is mostly negative sentiments from day 1 to 6.