RWorksheet_Eusuya#3B

Jonjeyl M. Eusuya

2024-10-10

1

A

```
data <- data.frame(
  Respondents = c(1:20),
  Sex = c(2, 2, 1, 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 2, 2, 2, 2, 2, 1, 2),
  FathersOccupation = c(1, 3, 3, 3, 1, 2, 3, 1, 1, 1, 3, 2, 1, 3, 3, 1, 3, 1, 2, 1) ,
  PersonsAtHome =  c(5, 7, 3, 8, 5, 9, 6, 7, 8, 4, 7, 5, 4, 7, 8, 8, 3, 11, 7, 6),
  SiblingsAtSchool =  c(6, 4, 4, 1, 2, 1, 5, 3, 1, 2, 3, 2, 5, 5, 2, 1, 2, 5, 3, 2),
  TypesOfHouses = c(1, 2, 3, 1, 1, 3, 3, 1, 2, 3, 2, 3, 2, 2, 3, 3, 3, 3, 3, 2)
)
data
```

```
##    Respondents Sex FathersOccupation PersonsAtHome SiblingsAtSchool
## 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
##    TypesOfHouses
## 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

The data has 20 rows and 6 columns

```
summary(data)
```

```
##   Respondents        Sex       FathersOccupation PersonsAtHome
##  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
##  SiblingsAtSchool TypesOfHouses
##  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
```

C

```
SiblingsAtSchool =  c(6, 4, 4, 1, 2, 1, 5, 3, 1, 2, 3, 2, 5, 5, 2, 1, 2, 5, 3, 2)
mean(SiblingsAtSchool)
```

```
## [1] 2.95
# No, the mean is 2.95
```

D

```
sub <- data[1:2, ]
sub
```

```
##   Respondents Sex FathersOccupation PersonsAtHome SiblingsAtSchool
## 1           1   2                 1             5                6
```

```
## 2             2   2             3             7             4
##   TypesOfHouses
## 1            1
## 2            2
```

E

```r
sub2 <- data[c(3, 5), c(2,4)]
sub2
```

```
##   Sex PersonsAtHome
## 3   1             3
## 5   2             5
```

F

```r
types_houses <- data[, "TypesOfHouses"]
types_houses
```

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

G

```r
males <- subset(data, Sex == 1 & FathersOccupation == 1)
males
```

```
## [1] Respondents      Sex              FathersOccupation PersonsAtHome
## [5] SiblingsAtSchool TypesOfHouses
## <0 rows> (or 0-length row.names)
```

H

```r
females <- subset(data, Sex ==2 & SiblingsAtSchool >=5)
females
```

```
##    Respondents Sex FathersOccupation PersonsAtHome SiblingsAtSchool
## 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
##    TypesOfHouses
## 1             1
## 7             3
## 13            2
## 14            2
## 18            3
```

2

A

```
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
# It shows an empty data frame
```

3

A

```
library(readxl)
HouseholdData <- read_excel("HouseholdData.csv")
HouseholdData
```

```
## # A tibble: 10 x 6
##     Respondents Sex     FathersOccupation PersonsAtHome SiblingsAtSchool
##           <dbl> <chr>               <dbl>         <dbl>            <dbl>
## 1            1 Male                    1             5                2
## 2            2 Female                  2             7                3
## 3            3 Female                  3             3                0
## 4            4 Male                    3             8                5
## 5            5 Male                    1             6                2
## 6            6 Female                  2             4                3
## 7            7 Female                  2             4                1
## 8            8 Male                    3             2                2
## 9            9 Female                  1            11                6
## 10          10 Male                    3             6                2
## # i 1 more variable: TypesOfHouses <chr>
```

B

```
HouseholdData$Sex <- factor(HouseholdData$Sex, levels = c("Male", "Female"), labels = c(1,2))
HouseholdData
```

```
## # A tibble: 10 x 6
##    Respondents Sex   FathersOccupation PersonsAtHome SiblingsAtSchool
##          <dbl> <fct>             <dbl>         <dbl>            <dbl>
## 1            1 1                     1             5                2
## 2            2 2                     2             7                3
## 3            3 2                     3             3                0
## 4            4 1                     3             8                5
## 5            5 1                     1             6                2
## 6            6 2                     2             4                3
## 7            7 2                     2             4                1
## 8            8 1                     3             2                2
## 9            9 2                     1            11                6
## 10          10 1                     3             6                2
## # i 1 more variable: TypesOfHouses <chr>
```

C

```
HouseholdData$TypesOfHouses <- factor(HouseholdData$TypesOfHouses, levels = c("Wood", "Congrete", "Semi-
HouseholdData
```

```
## # A tibble: 10 x 6
##    Respondents Sex   FathersOccupation PersonsAtHome SiblingsAtSchool
##          <dbl> <fct>             <dbl>         <dbl>            <dbl>
## 1            1 1                     1             5                2
## 2            2 2                     2             7                3
## 3            3 2                     3             3                0
## 4            4 1                     3             8                5
## 5            5 1                     1             6                2
## 6            6 2                     2             4                3
## 7            7 2                     2             4                1
## 8            8 1                     3             2                2
## 9            9 2                     1            11                6
## 10          10 1                     3             6                2
## # i 1 more variable: TypesOfHouses <fct>
```

D

```
HouseholdData$FathersOccupation <- factor(HouseholdData$FathersOccupation, levels = c(1, 2, 3), labels =
HouseholdData
```

```
## # A tibble: 10 x 6
##    Respondents Sex   FathersOccupation PersonsAtHome SiblingsAtSchool
##          <dbl> <fct> <fct>                     <dbl>            <dbl>
## 1            1 1     Farmer                        5                2
## 2            2 2     Driver                        7                3
## 3            3 2     Others                        3                0
## 4            4 1     Others                        8                5
## 5            5 1     Farmer                        6                2
## 6            6 2     Driver                        4                3
## 7            7 2     Driver                        4                1
```

```
##  8          8 1    Others                          2              2
##  9          9 2    Farmer                         11              6
## 10         10 1    Others                          6              2
## # i 1 more variable: TypesOfHouses <fct>
```

E

```
females2 <- subset(HouseholdData, Sex==2 & FathersOccupation =="Driver")
females2
```

```
## # A tibble: 3 x 6
##    Respondents Sex   FathersOccupation PersonsAtHome SiblingsAtSchool
##          <dbl> <fct> <fct>                     <dbl>            <dbl>
## 1            2 2     Driver                        7                3
## 2            6 2     Driver                        4                3
## 3            7 2     Driver                        4                1
## # i 1 more variable: TypesOfHouses <fct>
```

F

```
sub3 <- subset(HouseholdData, Respondents & SiblingsAtSchool >=  5)
sub3
```

```
## # A tibble: 2 x 6
##    Respondents Sex   FathersOccupation PersonsAtHome SiblingsAtSchool
##          <dbl> <fct> <fct>                     <dbl>            <dbl>
## 1            4 1     Others                        8                5
## 2            9 2     Farmer                       11                6
## # i 1 more variable: TypesOfHouses <fct>
```

4

The graph shows the positive, negative and neutral sentiments of tweets per day from July 14 to 21, 2020.