RWorksheet_Sorenio#3b.Rmd

2024-10-03

```
# 1.a

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)

df <- data.frame(Respondents = Respondents, Sex = Sex, Fathers_Occupation = Fathers_Occupation, Persons_

print(df)
```

```
##    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
# The data includes information from 20 people about their families and homes. Most respondents are fem

summary(df)
```

```
##   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
```

```
str(df)
```

```
## 'data.frame':    20 obs. of  6 variables:
##  $ Respondents       : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ Sex               : num  2 2 1 2 2 2 2 2 2 2 ...
##  $ Fathers_Occupation: num  1 3 3 3 1 2 3 1 1 1 ...
##  $ Persons_at_home   : num  5 7 3 8 5 9 6 7 8 4 ...
##  $ Siblings_at_school: num  6 4 4 1 2 1 5 3 1 2 ...
##  $ Types_of_houses   : num  1 2 3 1 1 3 3 1 2 3 ...
```

```
# c
# The mean of the number of siblings is 2.95 and not 5.

meansibs <- mean(df$Siblings_at_school)

print(meansibs)
```

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

```
# d

subsetdf <- df[1:2, ]

print(subsetdf)
```

```
##   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

subsetdf_2 <- df[c(3, 5), c(2, 4)]

print(subsetdf_2)
```

```
##   Sex Persons_at_home
## 3   1               3
## 5   2               5
```

```
# f

typeshouses <- df$Types_of_houses

print(typeshouses)
```

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

```
# g

maleF <- subset(df, Sex == 1 & Fathers_Occupation == 1)

print(maleF)
```

```
## [1] Respondents        Sex                Fathers_Occupation Persons_at_home
## [5] Siblings_at_school Types_of_houses
## <0 rows> (or 0-length row.names)
```

```
# h

femsibs <- subset(df, Sex == 2 & Siblings_at_school >= 5)

print(femsibs)
```

```
##    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
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
# An empty data frame w/ 5 columns for integers, decimal, text, and for true or false value. There are
```

```
# 3
respondents <- c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10)
sex <- c("Male", "Female", "Female", "Male", "Male", "Female", "Female", "Male", "Female", "Male")
fatherOccu <- c(1, 2, 3, 3, 1, 2, 2, 3, 1, 3)
personAH <- c(5, 7, 3, 8, 6, 4, 4, 2, 11, 6)
sibs <- c(2, 3, 0, 5, 2, 3, 1, 2, 6, 2)
typehouses <- c("Wood", "Congrete", "Congrete", "Wood", "Semi-congrete", "Semi-congrete", "Wood", "Semi

HHdata <- data.frame(Respondents = respondents, Sex = sex, Fathers_Occupation = fatherOccu, Persons_at_
HHdata
```

```
##    Respondents    Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 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
##    Types_of_Houses
## 1             Wood
## 2         Congrete
## 3         Congrete
## 4             Wood
## 5    Semi-congrete
## 6    Semi-congrete
## 7             Wood
## 8    Semi-congrete
## 9    Semi-congrete
## 10        Congrete
```

```
write.csv(HHdata, file = "HouseholdData.csv", FALSE)
# a
import <- read.csv("HouseholdData.csv")
import
```

```
##     X Respondents    Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 1   1           1   Male                  1               5                  2
## 2   2           2 Female                  2               7                  3
## 3   3           3 Female                  3               3                  0
## 4   4           4   Male                  3               8                  5
## 5   5           5   Male                  1               6                  2
## 6   6           6 Female                  2               4                  3
## 7   7           7 Female                  2               4                  1
## 8   8           8   Male                  3               2                  2
## 9   9           9 Female                  1              11                  6
## 10 10          10   Male                  3               6                  2
##    Types_of_Houses
## 1             Wood
## 2         Congrete
## 3         Congrete
## 4             Wood
## 5    Semi-congrete
## 6    Semi-congrete
## 7             Wood
## 8    Semi-congrete
## 9    Semi-congrete
## 10        Congrete
```

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

print(HHdata)
```

```
##    Respondents Sex Fathers_Occupation Persons_at_Home Siblings_at_School
```

```
## 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
##    Types_of_Houses
## 1             Wood
## 2         Congrete
## 3         Congrete
## 4             Wood
## 5    Semi-congrete
## 6    Semi-congrete
## 7             Wood
## 8    Semi-congrete
## 9    Semi-congrete
## 10        Congrete
```

```r
# c
HHdata$Types_of_Houses <- factor(HHdata$Types_of_Houses, levels = c("Wood", "Congrete", "Semi-congrete"))

print(HHdata)
```

```
##    Respondents Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 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
##    Types_of_Houses
## 1                1
## 2                2
## 3                2
## 4                1
## 5                3
## 6                3
## 7                1
## 8                3
## 9                3
## 10               2
```

```r
# hhdata from first part
HHdata <- read.csv("HouseholdData.csv")

print(HHdata)
```

```
##     X Respondents    Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 1   1           1   Male                  1               5                  2
## 2   2           2 Female                  2               7                  3
## 3   3           3 Female                  3               3                  0
## 4   4           4   Male                  3               8                  5
## 5   5           5   Male                  1               6                  2
## 6   6           6 Female                  2               4                  3
## 7   7           7 Female                  2               4                  1
## 8   8           8   Male                  3               2                  2
## 9   9           9 Female                  1              11                  6
## 10 10          10   Male                  3               6                  2
##    Types_of_Houses
## 1             Wood
## 2         Congrete
## 3         Congrete
## 4             Wood
## 5    Semi-congrete
## 6    Semi-congrete
## 7             Wood
## 8    Semi-congrete
## 9    Semi-congrete
## 10        Congrete
```

```r
# d
HHdata$Fathers_Occupation <- factor(HHdata$Fathers_Occupation,
                                    levels = c(1, 2, 3),
                                    labels = c("Farmer", "Driver", "Others"))

print(HHdata)
```

```
##     X Respondents    Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 1   1           1   Male             Farmer               5                  2
## 2   2           2 Female             Driver               7                  3
## 3   3           3 Female             Others               3                  0
## 4   4           4   Male             Others               8                  5
## 5   5           5   Male             Farmer               6                  2
## 6   6           6 Female             Driver               4                  3
## 7   7           7 Female             Driver               4                  1
## 8   8           8   Male             Others               2                  2
## 9   9           9 Female             Farmer              11                  6
## 10 10          10   Male             Others               6                  2
##    Types_of_Houses
## 1             Wood
## 2         Congrete
## 3         Congrete
## 4             Wood
## 5    Semi-congrete
## 6    Semi-congrete
## 7             Wood
## 8    Semi-congrete
## 9    Semi-congrete
## 10        Congrete
```

```r
# e
female_drivers <- subset(HHdata, Sex == "Female" & Fathers_Occupation == "Driver")

print(female_drivers)
```

```
##   X Respondents    Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 2 2           2 Female             Driver               7                  3
## 6 6           6 Female             Driver               4                  3
## 7 7           7 Female             Driver               4                  1
##   Types_of_Houses
## 2        Congrete
## 6   Semi-congrete
## 7            Wood
```

```r
# f
sibling_filter <- subset(HHdata, Siblings_at_School >= 5)


print(sibling_filter)
```

```
##   X Respondents    Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 4 4           4   Male             Others               8                  5
## 9 9           9 Female             Farmer              11                  6
##   Types_of_Houses
## 4            Wood
## 9   Semi-congrete
```

4. The graph is all about the sentiments of tweets: Negative in red color, Neutral in orange color, Positive in blue color. Some days (July 15 and 21) had many negative tweets , while other days had more positive or neutral feelings. This suggests that people's opinions were mixed and changed based on events happening during that week and it mirrors their tweets.