

RWorksheet_lauron#3b.Rmd

Mary Ghale C. Lauron

2025-10-13

```
#Create a data frame

personalInfo <- data.frame(Respondents = c(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, 1, 2, 2, 2, 2, 2, 2, 2, 1, 2),
FathersWork = c(1, 3, 3, 3, 1, 2, 3, 1, 1, 1, 3, 2, 1, 3, 3, 1, 3, 1, 2, 1),
Persons_Home = c(5, 7, 3, 8, 5, 9, 6, 7, 8, 4, 7, 5, 4, 7, 8, 8, 3, 11, 7, 6),
studSibling = c(6, 4, 4, 1, 2, 1, 5, 3, 1, 2, 3, 2, 5, 5, 2, 1, 2, 5, 3, 2),
Housetype = c(1, 2, 3, 1, 1, 3, 3, 1, 2, 3, 2, 3, 2, 2, 3, 3, 3, 3, 3, 2)
)
personalInfo

##   Respondents Sex FathersWork Persons_Home studSibling Housetype
## 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
str(personalInfo)

## 'data.frame': 20 obs. of 6 variables:
## $ Respondents : num  1 2 3 4 5 6 7 8 9 10 ...
## $ Sex         : num  2 2 1 2 2 2 2 2 2 2 ...
## $ FathersWork : num  1 3 3 3 1 2 3 1 1 1 ...
## $ Persons_Home: num  5 7 3 8 5 9 6 7 8 4 ...
## $ studSibling : num  6 4 4 1 2 1 5 3 1 2 ...
## $ Housetype   : num  1 2 3 1 1 3 3 1 2 3 ...
```

```

summary(personalInfo)

##   Respondents      Sex   FathersWork Persons_Home studSibling
## 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
##   Housetype
## Min.   :1.0
## 1st Qu.:2.0
## Median :2.5
## Mean   :2.3
## 3rd Qu.:3.0
## Max.   :3.0

#c
mean_sib <- mean(personalInfo$studSibling)
mean_sib

## [1] 2.95
#- no, it's lower than 5

#d
fst_2rows <- personalInfo[1:2, ]
fst_2rows

##   Respondents Sex FathersWork Persons_Home studSibling Housetype
## 1           1   2           1           5           6           1
## 2           2   2           3           7           4           2

#e
extract<- personalInfo[c(3,5), c(2,4)]
extract

##   Sex Persons_Home
## 3   1            3
## 5   2            5

#f
Housetype<- personalInfo$Housetype
Housetype

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

#g
maleresp_farmer <- subset(personalInfo, Sex == 1 & FathersWork == 1)
maleresp_farmer

## [1] Respondents  Sex          FathersWork  Persons_Home studSibling
## [6] Housetype
## <0 rows> (or 0-length row.names)

#h
femaleresp_siblings <- subset(personalInfo, Sex == 2 & studSibling>= 5)
femaleresp_siblings

```

```

##      Respondents Sex FathersWork Persons_Home studSibling Housetype
## 1            1     2           1         5          6        1
## 7            7     2           3         6          5        3
## 13           13    2           1         4          5        2
## 14           14    2           3         7          5        2
## 18           18    2           1        11          5        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

#3
household <- data.frame(
  Respondents = 1:10,
  Sex = c("Male", "Female", "Female", "Male", "Male", "Female", "Female", "Male",
  "Female", "Male"),
  FathersWork = c(1, 2, 3, 3, 1, 2, 2, 3, 1, 3),
  Persons_Home = c(5, 7, 3, 8, 5, 4, 4, 2, 11, 6),
  studSibling = c(2, 3, 0, 5, 2, 4, 4, 2, 6, 6),
  Housetype = c("Wood", "Concrete", "Concrete", "Wood", "Semi-concrete",
  "Semi-concrete", "Wood", "Semi-concrete", "Semi-concrete", "Concrete")
)
write.csv(household, "HouseholdData.csv", row.names = FALSE)

#3a Import CSV file
household_data <- read.csv("HouseholdData.csv", stringsAsFactors = FALSE)
household_data

##      Respondents   Sex FathersWork Persons_Home studSibling Housetype
## 1            1 Male           1         5          2       Wood
## 2            2 Female         2         7          3     Concrete
## 3            3 Female         3         3          0     Concrete
## 4            4 Male           3         8          5       Wood
## 5            5 Male           1         5          2 Semi-concrete
## 6            6 Female         2         4          4 Semi-concrete
## 7            7 Female         2         4          4       Wood

```

```

## 8      8 Male      3      2      2 Semi-concrete
## 9      9 Female    1     11      6 Semi-concrete
## 10     10 Male     3      6      6 Concrete

#3b
# Convert Sex into factor
household_data$Sex <- factor(household_data$Sex, levels = c("Male", "Female"),
                               labels = c(1, 2))
household_data$Sex

## [1] 1 2 2 1 1 2 2 1 2 1
## Levels: 1 2

#3c
household_data$Housetype <- factor(
  household_data$Housetype,
  levels = c("Wood", "Concrete", "Semi-concrete"),
  labels = c(1, 2, 3)
)
household_data$Housetype

## [1] 1 2 2 1 3 3 1 3 3 2
## Levels: 1 2 3

#3d
household_data$FathersWork <- factor(
  household_data$FathersWork,
  levels = c(1, 2, 3),
  labels = c("Farmer", "Driver", "Others")
)

household_data$FathersWork

## [1] Farmer Driver Others Others Farmer Driver Driver Others Farmer Others
## Levels: Farmer Driver Others

#3e
female_driver <- subset(household_data, Sex == "2" & FathersWork == "Driver")
female_driver

## Respondents Sex FathersWork Persons_Home studSibling Housetype
## 2      2 2 Driver      7      3      2
## 6      6 2 Driver      4      4      3
## 7      7 2 Driver      4      4      1

#3f
respondent_siblings <- subset(household_data, studSibling >= 5)
respondent_siblings

## Respondents Sex FathersWork Persons_Home studSibling Housetype
## 4      4 1 Others      8      5      1
## 9      9 2 Farmer     11      6      3
## 10     10 1 Others      6      6      2

#4 Interpret the graph
#The graph illustrates the count of tweets from July 14 to July 21, 2020,
#categorized by sentiment type (positive, negative, and neutral).
#Based on the visualization, it is evident that negative sentiments dominate across
#all dates. The count of negative tweets started high on July 14 and peaked sharply

```

*#on July 15, reaching approximately 4,000 tweets. Meanwhile, positive sentiment
#consistently rank second as the neutral to last.*