

# RWorksheet\_Sanceda#3b

2024-10-05

#1.  
#a.

```
Respondents <- c(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(2,2,1,2,2,2,2,2,2,2,1,2,2,2,2,2,2,1,2)
Sex
```

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

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

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

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

```
## [1] 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)
Siblingsatschool
```

```
## [1] 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)
Typesofhouses
```

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

```
Data <- data.frame(Respondents=Respondents, Sex=Sex, FatherOccupation=FatherOccupation, Personathome=Personathome, Siblingsatschool=Siblingsatschool, Typesofhouses=Typesofhouses)
Data
```

```
## Respondents Sex FatherOccupation Personathome Siblingsatschool Typesofhouses
## 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. The data is more clearer and organized.*

```
summary(Data)
```

```
## Respondents      Sex      FatherOccupation  Personathome
## 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
```

```
str(Data)
```

```
## '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 ...
## $ FatherOccupation: num 1 3 3 3 1 2 3 1 1 1 ...
## $ Personathome : num 5 7 3 8 5 9 6 7 8 4 ...
## $ Siblingsatschool: num 6 4 4 1 2 1 5 3 1 2 ...
## $ Typesofhouses : num 1 2 3 1 1 3 3 1 2 3 ...
```

*#c. The mean number is 2.95, not 5.*

```
mean(Data[, 5])
```

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

#d.

```
rows <- Data[1:2, ]
rows
```

```
## Respondents Sex FatherOccupation Personathome Siblingsatschool Typesofhouses
## 1          1  2              1          5          6          1
## 2          2  2              3          7          4          2
```

#e.

```
xtract <- Data[c(3:5), c(2,4)]
xtract
```

```
## Sex Personathome
## 3  1            3
## 4  2            8
## 5  2            5
```

#f.

```
tHouses <- Data[,6]
tHouses
```

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

#g.

```
male <- subset(Data, Data[,2] == 1 & Data[,3] == 1 )
male
```

```
## [1] Respondents      Sex      FatherOccupation Personathome
## [5] Siblingsatschool Typesofhouses
## <0 rows> (or 0-length row.names)
```

#h.

```
female <- subset(Data, Data[,2] == 2 & Data[,5] >= 5)
female
```

```
## Respondents Sex FatherOccupation Personathome Siblingsatschool Typesofhouses
## 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.

#a.

```
df = data.frame(Ints=integer(), Doubles=double(), Characters=character(), Logicals=logical(), Factors=factor(),
                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.
```

```
#a.
```

```
Respondents <- c(1,2,3,4,5,6,7,8,9,10)
Sex <- c("Male", "Female", "Female", "Male", "Male", "Female", "Female", "Male", "Female", "Male")
Father <- c(1,2,3,3,1,2,2,3,1,3)
Persons <- c(5,7,3,8,6,4,4,2,11,6)
Siblings <- c(2,3,0,5,2,3,1,2,6,2)
Houses <- c("Wood", "Concrete", "Concrete", "Wood", "Semi-concrete", "Semi-concrete", "Wood", "Semi-concrete", "Wood", "Semi-concrete")

DataH <- data.frame(Respondents = Respondents, Sex = Sex, Fathers_Occupation = Father, Persons = Persons, Siblings = Siblings, Types_of_Houses = Houses)
DataH
```

```
##      Respondents      Sex Fathers_Occupation Persons Siblings Types_of_Houses
## 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         6         2    Semi-concrete
## 6             6  Female                2         4         3    Semi-concrete
## 7             7  Female                2         4         1         Wood
## 8             8    Male                3         2         2    Semi-concrete
## 9             9  Female                1        11         6    Semo-concrete
## 10            10    Male                3         6         2         Concrete
```

```
write.csv(DataH, file = "HouseholdData.csv", FALSE)
```

```
#a.
```

```
Imprtd <- read.csv("HouseholdData.csv")
Imprtd
```

```
##      X Respondents      Sex Fathers_Occupation Persons Siblings Types_of_Houses
## 1     1             1    Male                1         5         2         Wood
## 2     2             2  Female                2         7         3         Concrete
## 3     3             3  Female                3         3         0         Concrete
## 4     4             4    Male                3         8         5         Wood
## 5     5             5    Male                1         6         2    Semi-concrete
## 6     6             6  Female                2         4         3    Semi-concrete
## 7     7             7  Female                2         4         1         Wood
## 8     8             8    Male                3         2         2    Semi-concrete
## 9     9             9  Female                1        11         6    Semo-concrete
## 10    10            10    Male                3         6         2         Concrete
```

#b.

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

##	X	Respondents	Sex	Fathers_Occupation	Persons	Siblings	Types_of_Houses	
## 1	1	1	1		1	5	2	Wood
## 2	2	2	2		2	7	3	Concrete
## 3	3	3	2		3	3	0	Concrete
## 4	4	4	1		3	8	5	Wood
## 5	5	5	1		1	6	2	Semi-concrete
## 6	6	6	2		2	4	3	Semi-concrete
## 7	7	7	2		2	4	1	Wood
## 8	8	8	1		3	2	2	Semi-concrete
## 9	9	9	2		1	11	6	Semo-concrete
## 10	10	10	1		3	6	2	Concrete

#c.

```
Imprtd$Types_of_Houses <- factor(Imprtd$Types_of_Houses, levels = c("Wood", "Congrete", "Semi-cong"), labels = c(1,2,3))
Imprtd
```

##	X	Respondents	Sex	Fathers_Occupation	Persons	Siblings	Types_of_Houses	
## 1	1	1	1		1	5	2	1
## 2	2	2	2		2	7	3	<NA>
## 3	3	3	2		3	3	0	<NA>
## 4	4	4	1		3	8	5	1
## 5	5	5	1		1	6	2	<NA>
## 6	6	6	2		2	4	3	<NA>
## 7	7	7	2		2	4	1	1
## 8	8	8	1		3	2	2	<NA>
## 9	9	9	2		1	11	6	<NA>
## 10	10	10	1		3	6	2	<NA>

#d.

```
Imprtd$Father <- factor(Imprtd$Father, levels = c("Farmer", "Driver", "Others"), labels = c(1, 2, 3))
Imprtd
```

##	X	Respondents	Sex	Fathers_Occupation	Persons	Siblings	Types_of_Houses	
## 1	1	1	1		1	5	2	1
## 2	2	2	2		2	7	3	<NA>
## 3	3	3	2		3	3	0	<NA>
## 4	4	4	1		3	8	5	1
## 5	5	5	1		1	6	2	<NA>
## 6	6	6	2		2	4	3	<NA>
## 7	7	7	2		2	4	1	1
## 8	8	8	1		3	2	2	<NA>
## 9	9	9	2		1	11	6	<NA>
## 10	10	10	1		3	6	2	<NA>

## Father

## 1	<NA>
------	------

```
## 2    <NA>
## 3    <NA>
## 4    <NA>
## 5    <NA>
## 6    <NA>
## 7    <NA>
## 8    <NA>
## 9    <NA>
## 10   <NA>
```

*#e.*

```
FtDrivers <- subset(Imprtd, Sex == 2 & Father == 2)
FtDrivers
```

```
## [1] X Respondents Sex Fathers_Occupation
## [5] Persons Siblings Types_of_Houses Father
## <0 rows> (or 0-length row.names)
```

*#f.*

```
Sblngs <- subset(Imprtd, Siblings >= 5)
Sblngs
```

```
## X Respondents Sex Fathers_Occupation Persons Siblings Types_of_Houses Father
## 4 4 4 1 3 8 5 1 <NA>
## 9 9 9 2 1 11 6 <NA> <NA>
```

*#4.*

*#The graph shows the sentiment of tweets day by day. Negative tweets being  
#the highest followed by Positive and lastly, Neutral. This could change,  
#depending on the issues and trends in society and also on the internet.*