

RWorksheet_Marquez#3b

Kurt Sayro D. Marquez

2024-10-14

#1. #a.

```
table1 <- data.frame(Respondent = c(1:20),  
  Sex = c(2,2,1,2,2,2,2,2,2,2,1,2,2,2,2,2,2,1,2),  
  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))
```

table1

##	Respondent	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
```

#legend: Male=1 Female=2 Farmer=1 Driver=2 Others=3 Wood=1 Semi-Concrete=2 Concrete=3 #b.

```
summary(table1)
```

```
##      Respondent      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. No, the mean is 2.95

#d.

```
table1[1:2,]
```

```
##      Respondent Sex FathersOccupation PersonsAtHome SiblingsAtSchool TypesOfHouses
## 1             1  2             1             5             6             1
## 2             2  2             3             7             4             2
```

#e.

```
table1[c(3,5), c(2,4)]
```

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

#f.

```
types_houses <- table1$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.

```
MaleFarmers <- subset(table1, Sex == 1 & FathersOccupation == 1)
MaleFarmers
```

```
## [1] Respondent      Sex      FathersOccupation PersonsAtHome
## [5] SiblingsAtSchool TypesOfHouses
```

```
## <0 rows> (or 0-length row.names)
```

```
#h.
```

```
Female_Siblings <- subset(table1, Sex == 2 & SiblingsAtSchool >= 5)
Female_Siblings
```

```
##      Respondent 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.
```

```
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. It displays the data type of the object #3. #a.
```

```
HouseData <- read.csv("/cloud/project/RWorkSheet_Marquez#3b/HouseholdData.csv")
HouseData
```

```
##      Respondents      Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 1             1   Male             1             1             5
## 2             2 Female             2             2             7
## 3             3 Female             3             3             0
## 4             4   Male             3             3             8
## 5             5   Male             1             1             6
## 6             6 Female             2             2             4
## 7             7 Female             2             4             1
## 8             8   Male             3             3             2
## 9             9 Female             1             1            11
## 10            10   Male             3             3             6
```

```
## Types.of.Houses
## 1      Wood
## 2      Concrete
## 3      Concrete
## 4      Wood
## 5      Semi-concrete
## 6      Semi-concrete
## 7      Wood
## 8      Semi-concrete
## 9      Semi-concrete
## 10     Concrete
```

#b.

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

```
## Respondents Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 1      1      1      1      1      5
## 2      2      2      2      2      7
## 3      3      2      3      3      0
## 4      4      1      3      3      8
## 5      5      1      1      1      6
## 6      6      2      2      2      4
## 7      7      2      2      4      1
## 8      8      1      3      3      2
## 9      9      2      1      1     11
## 10     10     1      3      3      6
```

```
## Types.of.Houses
## 1      Wood
## 2      Concrete
## 3      Concrete
## 4      Wood
## 5      Semi-concrete
## 6      Semi-concrete
## 7      Wood
## 8      Semi-concrete
## 9      Semi-concrete
## 10     Concrete
```

#c.

```
HouseData$Types.of.Houses <- factor(HouseData$Types.of.Houses, levels = c("Wood", "Congrete", "Semi-concrete"), labels = c(1,2,3))
HouseData
```

```
## Respondents Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 1      1      1      1      1      5
## 2      2      2      2      2      7
## 3      3      2      3      3      0
## 4      4      1      3      3      8
## 5      5      1      1      1      6
## 6      6      2      2      2      4
## 7      7      2      2      4      1
## 8      8      1      3      3      2
## 9      9      2      1      1     11
## 10     10     1      3      3      6
```

```
## Types.of.Houses
```

```
## 1      1
## 2      <NA>
## 3      <NA>
## 4      1
## 5      <NA>
## 6      <NA>
## 7      1
## 8      <NA>
## 9      <NA>
## 10     <NA>
```

#d.

```
HouseData$Fathers.Occupation <- factor(HouseData$Fathers.Occupation, levels = c(1,2,3), label = c("Farmer", "Driver", "Others"))
HouseData
```

```
##      Respondents Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 1      1      1      Farmer      1      5
## 2      2      2      Driver      2      7
## 3      3      2      Others      3      0
## 4      4      1      Others      3      8
## 5      5      1      Farmer      1      6
## 6      6      2      Driver      2      4
## 7      7      2      Driver      4      1
## 8      8      1      Others      3      2
## 9      9      2      Farmer      1     11
## 10     10     1      Others      3      6
```

```
##      Types.of.Houses
## 1      1
## 2      <NA>
## 3      <NA>
## 4      1
## 5      <NA>
## 6      <NA>
## 7      1
## 8      <NA>
## 9      <NA>
## 10     <NA>
```

#e.

```
Female_DriverDad <- subset(HouseData, Sex == 2 & Fathers.Occupation == "Driver")
Female_DriverDad
```

```
##      Respondents Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 2      2      2      Driver      2      7
## 6      6      2      Driver      2      4
## 7      7      2      Driver      4      1
##      Types.of.Houses
## 2      <NA>
## 6      <NA>
## 7      1
```

#f.

```
manysibs <- subset(HouseData, Respondents & Siblings.at.School >= 5)
manysibs
```

```
##      Respondents Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 1          1    1          Farmer              1              5
## 2          2    2          Driver              2              7
## 4          4    1          Others              3              8
## 5          5    1          Farmer              1              6
## 9          9    2          Farmer              1             11
## 10         10    1          Others              3              6
##      Types.of.Houses
## 1          1
## 2         <NA>
## 4          1
## 5         <NA>
## 9         <NA>
## 10        <NA>
```

#4. The graph represents the data of Sentiments of Tweets Per Day in the month of July. It shows the positive, neutral, and negative tweets per day.

```
data <- data.frame(
  Date = rep(c("July 14, 2020", "July 15, 2020", "July 17, 2020", "July 18, 2020", "July 20, 2020", "July 21, 2020"), times = 6),
  Sentiment = rep(c("Negative", "Neutral", "Positive"), times = 6),
  Count = c(2500, 1500, 2000, 4200, 3000, 3100,
            2000, 1200, 1800, 3000, 2500, 2700,
            2800, 2100, 2600, 1800, 1200, 2200))
data$Date <- factor(data$Date, levels = c("July 14, 2020", "July 15, 2020", "July 17, 2020",
                                           "July 18, 2020", "July 20, 2020", "July 21, 2020"))
data$Date
```

```
## [1] July 14, 2020 July 14, 2020 July 14, 2020 July 15, 2020 July 15, 2020
## [6] July 15, 2020 July 17, 2020 July 17, 2020 July 17, 2020 July 18, 2020
## [11] July 18, 2020 July 18, 2020 July 20, 2020 July 20, 2020 July 20, 2020
## [16] July 21, 2020 July 21, 2020 July 21, 2020
## 6 Levels: July 14, 2020 July 15, 2020 July 17, 2020 ... July 21, 2020
```