

RWorksheet_Francisco#3b

Enrique Francisco

2024-10-13

#1.a

```
make_table <- 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), Fathers_occupation = c(1,2,3,3,1,2,3,1,1,1,3,2,1,3,3,1,3,2,2), PersonsAtHome = c(5,7,3,8,5,9,6,7,8,4,7,5,4,7,8,3,11,7,6), SiblingsatSchool = c(6,4,4,1,2,1,5,3,1,2,3,2,5,5,2,1,5,3,2))
```

```
make_table
```

##	Respondent	Sex	Fathers_Occupation	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 Farmer=1 Wood=1 # Female=2 Driver=2 Semi-Concrete=2 # Others=3 Concrete=3

#1.b

```
summary(make_table)
```

```
##      Respondent      Sex      Fathers_Occupation 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
```

#1.c: No, the mean is 2.95 based on summary

#1.d

```
access <- make_table[1:2,]
access
```

```
##      Respondent Sex Fathers_Occupation PersonsAtHome SiblingsatSchool
## 1             1  2                1                5                6
## 2             2  2                3                7                4
##      TypesOfHouses
## 1                 1
## 2                 2
```

#1.e

```
access2 <- make_table[c(3,5),c(2,4)]
access2
```

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

#1.f

```
types_houses <- make_table$TypesOfHouses
types_houses
```

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

#1.g

```
males_respondent <- subset(make_table, Sex == 1 & Fathers_Occupation == 1)
males_respondent
```

```
##      [1] Respondent      Sex      Fathers_Occupation PersonsAtHome
##      [5] SiblingsatSchool  TypesOfHouses
##      <0 rows> (or 0-length row.names)
```

```
#1.h
female_respondent <- subset(make_table, Sex == 2 & SiblingsatSchool >= 5)
female_respondent
```

```
##      Respondent Sex Fathers_Occupation 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
```

#2.a: It prints the datatype of the object and there's 0 levels

```
#3.a
HousesData <- read.csv("/cloud/project/Worksheet#3/HouseholdData.csv")
HousesData
```

```
##      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
```

#3.b

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

```
##      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
```

#3.c

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

```
##      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
```

#3.d

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

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

#3.e

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

```
##      Respondents Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 2      2      2      Driver      7      3
## 6      6      2      Driver      4      3
## 7      7      2      Driver      4      1
##      Types.of.Houses
## 2      2
## 6      3
## 7      1
```

#3.f

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

```
##      Respondents Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 4      4      1      Others      8      5
## 9      9      2      Farmer     11      6
##      Types.of.Houses
```

## 4	1
## 9	3

#4

#. The graph represents the data of Sentiments of Tweets Per Day in the month of July. It show the diff