RWorksheet_3b.Rmd

Lore Faith Sabio

2023-10-10

```
Respondents <- c(1:20)
Sex <- c(2,2,1,2,2,2,2,2,2,2,1,2,2,2,2,2,2,2,1,2)
FatherOccupation <- c(1,3,3,3,1,2,3,1,1,1,3,2,1,3,3,1,3,1,2,1)
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)
TypeOfHouses <- c (1,2,3,1,1,3,3,1,2,3,2,3,2,2,3,3,3,3,3,2)

householdData <- data.frame(Respondents, Sex,FatherOccupation,PersonsAtHome,SiblingsAtSchool,TypeOfHouse
householdData
```

```
   Respondents Sex FatherOccupation PersonsAtHome SiblingsAtSchool TypeOfHouses
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
```

```
#This data compose of 20 respondents with 7 males and 13 females. Also, the records of their fathers oc
summary(householdData)
```

```
  Respondents          Sex       FatherOccupation 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  TypeOfHouses
 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
```

```
meanSiblings <- mean(householdData$SiblingsAtSchool)
meanSiblings == 5
```

```
[1] FALSE
```

```
subSet1 <- householdData[1:2, ]
subSet1
```

```
  Respondents Sex FatherOccupation PersonsAtHome SiblingsAtSchool TypeOfHouses
1           1   2                1             5                6            1
2           2   2                3             7                4            2
```

```
subSet2 <- householdData[c(3,5), c(2,4)]
subSet2
```

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

```
types_houses <- householdData$TypeOfHouses
types_houses
```

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

```
selected_data <- subset(householdData, Sex == 1 & FatherOccupation >= 1)
selected_data
```

```
   Respondents Sex FatherOccupation PersonsAtHome SiblingsAtSchool TypeOfHouses
3            3   1                3             3                4            3
11          11   1                3             7                3            2
19          19   1                2             7                3            3
```

```
selected_data1 <- subset(householdData, Sex == 2 & SiblingsAtSchool >= 5)
selected_data1
```

```
   Respondents Sex FatherOccupation PersonsAtHome SiblingsAtSchool TypeOfHouses
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
```

```
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
```

```
#The result shows that the df is the data frame with  observation of 5 variables which are Ints,Doubles
```

```
respondents<-  c(1:10)
sex <- c("Male","Female","Female","Male","Male","Female","Female","Male","Female","Male")
Fathers_Occupation <- c(1,2,3,3,1,2,2,3,1,3)
Persons_At_Home <- c(5,7,3,8,6,4,4,2,11,6)
Siblings_At_School <- c(2,3,0,5,2,3,1,2,6,2)
Type_Of_Houses <- c ("Wood","Congrete","Congrete","Wood","Semi-Congrete","Semi-Congrete","Wood","Semi-C

HouseholdData <- data.frame(respondents,sex,Fathers_Occupation,Persons_At_Home,Siblings_At_School,Type_


csv_file <- "HouseholdData.csv"
write.csv(HouseholdData, file = csv_file)
HouseholdData <- read.csv("HouseholdData.csv")
```

```
HouseholdData$sex <- factor(HouseholdData$sex)
HouseholdData$sex <- as.integer(factor(HouseholdData$sex,
levels = c("Male", "Female"),
labels = c(1, 2)))
HouseholdData
```

```
    X respondents sex Fathers_Occupation Persons_At_Home Siblings_At_School
1   1           1   1                  1               5                  2
2   2           2   2                  2               7                  3
3   3           3   2                  3               3                  0
4   4           4   1                  3               8                  5
5   5           5   1                  1               6                  2
6   6           6   2                  2               4                  3
7   7           7   2                  2               4                  1
8   8           8   1                  3               2                  2
9   9           9   2                  1              11                  6
10 10          10   1                  3               6                  2
   Type_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
```

```
HouseholdData$Type_Of_Houses <- factor(HouseholdData$Type_Of_Houses)
HouseholdData$Type_Of_Houses <- as.integer(factor(HouseholdData$Type_Of_Houses,
levels = c("Wood", "Congrete", "Semi-Congrete"),
labels = c(1, 2, 3)))
```

```r
print(HouseholdData)
```

```
    X respondents sex Fathers_Occupation Persons_At_Home Siblings_At_School
1   1           1   1                  1               5                  2
2   2           2   2                  2               7                  3
3   3           3   2                  3               3                  0
4   4           4   1                  3               8                  5
5   5           5   1                  1               6                  2
6   6           6   2                  2               4                  3
7   7           7   2                  2               4                  1
8   8           8   1                  3               2                  2
9   9           9   2                  1              11                  6
10 10          10   1                  3               6                  2
   Type_Of_Houses
1               1
2               2
3               2
4               1
5               3
6               3
7               1
8               3
9               3
10              2
```

```r
HouseholdData$Fathers_Occupation <- factor(HouseholdData$Fathers_Occupation)
HouseholdData$Fathers_Occupation <- as.character(factor(HouseholdData$Fathers_Occupation,
levels = c(1, 2, 3),
labels = c("Farmer", "Driver", "Others")))
print(HouseholdData)
```

```
    X respondents sex Fathers_Occupation Persons_At_Home Siblings_At_School
1   1           1   1             Farmer               5                  2
2   2           2   2             Driver               7                  3
3   3           3   2             Others               3                  0
4   4           4   1             Others               8                  5
5   5           5   1             Farmer               6                  2
6   6           6   2             Driver               4                  3
7   7           7   2             Driver               4                  1
8   8           8   1             Others               2                  2
9   9           9   2             Farmer              11                  6
10 10          10   1             Others               6                  2
   Type_Of_Houses
1               1
2               2
3               2
4               1
5               3
6               3
7               1
8               3
9               3
10              2
```

```r
selected_data2 <- subset(HouseholdData, sex == 2 & Fathers_Occupation == "Driver")
selected_data2
```

```
  X respondents sex Fathers_Occupation Persons_At_Home Siblings_At_School
2 2           2   2             Driver               7                  3
6 6           6   2             Driver               4                  3
7 7           7   2             Driver               4                  1
  Type_Of_Houses
2              2
6              3
7              1
```

```r
selected_data3 <- subset(HouseholdData, sex == 2 & Siblings_At_School >= 5)
selected_data3
```

```
  X respondents sex Fathers_Occupation Persons_At_Home Siblings_At_School
9 9           9   2             Farmer              11                  6
  Type_Of_Houses
9              3
```

```r
#The bar graph titled "Sentiment of Tweets per Day." This graph appears to represent sentiment analysis
#over a specific time frame, which is Twitter data from the year 2020. The sentiment analysis
#categorizes tweets into three sentiment groups: Negative, Neutral, and Positive.

#The analysis of the graph focuses on several specific dates in July 2020. On July 14, 2020,
#the graph shows that there were nearly 2,500 tweets categorized as negative sentiment,
#around 1,500 tweets categorized as neutral,
#and approximately 1,750 tweets categorized as positive.

#Moving on to July 15, 2020, the negative sentiment category saw a significant increase with over 4,000
#tweets categorized as negative. The neutral sentiment category
#had about 2,750 tweets, and the positive sentiment category had almost 3,200 tweets.

#On July 17, 2020, the negative sentiment count was approximately 3,250 tweets, while the neutral senti
#category had about 1,800 tweets, and the positive sentiment category had almost 2,500 tweets.

#The analysis continues with July 18, 2020, where the negative sentiment count remained at around 3,250
#tweets. The neutral sentiment category had about 2,000 tweets, and the
#positive sentiment category had approximately 2,500 tweets.

#Moving to July 20, 2020, the negative sentiment count was nearly 2,500 tweets. The neutral sentiment
#category had around 1,500 tweets, and almost 1,750 tweets were categorized as positive.

#Finally, on July 21, 2020, the negative sentiment count was around 4,000 tweets, while the neutral sen
#category had about 2,600 tweets, and more than 3,000 tweets were categorized as positive.

#In conclusion, the analysis highlights the variations in sentiment categories over these specific
#dates in July 2020. Notably, July 15 had the highest count of negative sentiment tweets, while July 20
#had the lowest count, both in the negative and neutral categories. In contrast, July 21 had the
#highest count in the positive #sentiment category, and July 20 had the
#lowest count. Overall, July 20, 2020, appeared to have the lowest number of total tweets
#categorized across all sentiment categories.
```