Rworksheet_de la Cruz-Hanz #3b

2023-10-04

```
personalInfo <- data.frame (
  Respondents = 1:20,
  Sex = c(2,2,1,rep(2,7),1,rep(2,7),1,2),
  Fathers_Occupation = c(1,rep(3,3),1:3,rep(1,3),(3:1),3,3,1,3,1,2,1),
  Persons_at_Home = c(5,7,3,8,5,9,6,7,8,4,7,5,4,7,8,8,3,11,7,6),
  Siblings_at_School = c(6,4,4,1,2,1,5,3,1,2,3,2,5,5,2,1,2,5,3,2),
  Types_of_Houses = c(1:3,1,1,3,3,1:3,2,3,2,2,rep(3,5),2)
  )

  personalInfo
```

```
##    Respondents Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 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
##    Types_of_Houses
## 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
```

```
summary(personalInfo)
```

```
##   Respondents          Sex       Fathers_Occupation Persons_at_Home
##  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
##  Siblings_at_School Types_of_Houses
##  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(personalInfo)
```

```
## '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 ...
##  $ Fathers_Occupation: num  1 3 3 3 1 2 3 1 1 1 ...
##  $ Persons_at_Home   : num  5 7 3 8 5 9 6 7 8 4 ...
##  $ Siblings_at_School: num  6 4 4 1 2 1 5 3 1 2 ...
##  $ Types_of_Houses   : num  1 2 3 1 1 3 3 1 2 3 ...
```

```
#It shows the mean, the median the 1st and 3rd Quarter, the minimum and maximum number.
#It also shows the data or observations inside the columns, the column names and how many observations,
# and variables inside the data frame
```

```
meanSiblings <- mean(personalInfo$Siblings_at_School)
meanSiblings
```

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

```
#the means of the siblings attending school is not 5 but 2.95 or 3
```

```
firstTwoRowsCols <- personalInfo[1:2,1:2]
firstTwoRowsCols
```

```
##   Respondents Sex
## 1           1   2
## 2           2   2
```

```
rowsColsBetween <- personalInfo[c(3,5),c(2,4)]
rowsColsBetween
```

```
##   Sex Persons_at_Home
## 3   1               3
## 5   2               5
```

```
types_house <- personalInfo$Types_of_Houses
types_house
```

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

```
maleAndFarmers <- personalInfo[personalInfo$Sex == 1 & personalInfo$Fathers_Occupation == 1, ]
maleAndFarmers
```

```
## [1] Respondents        Sex                Fathers_Occupation Persons_at_Home
## [5] Siblings_at_School Types_of_Houses
## <0 rows> (or 0-length row.names)
```

```
femaleAndSchool <- personalInfo[personalInfo$Sex == 2 & personalInfo$Siblings_at_School >= 5,]
femaleAndSchool
```

```
##    Respondents Sex Fathers_Occupation Persons_at_Home Siblings_at_School
## 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
##    Types_of_Houses
## 1                1
## 7                3
## 13               2
## 14               2
## 18               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:"
```

```r
print(str(df))
```

```
## 'data.frame':    0 obs. of  5 variables:
##  $ Ints      : int
##  $ Doubles   : num
##  $ Characters: chr
##  $ Logicals  : logi
##  $ Factors   : Factor w/ 0 levels:
## NULL
```

```r
#the results show many variables or columns are in the data frame, shows the observations,
#shows the column names and shows the equivalent function for each vector.
```

```r
houseHoldData <- data.frame (
  Respondents = 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),
  Types_of_Houses = c("Wood","Congrete","Congrete","Wood","Semi-Congrete","Semi-Congrete","Wood",
                      "Semi-Congrete","Semi-Congrete","Congrete")
)
houseHoldData
```

```
##    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
```

```r
householdCSV <- "household_data.csv"
write.csv(houseHoldData, file = householdCSV, row.names = FALSE)
householdCSV
```

```
## [1] "household_data.csv"
```

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

houseHoldData
```

```
##    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
```

```
houseHoldData$Types_of_Houses <- factor(houseHoldData$Types_of_Houses,
                                        levels = c("Wood", "Congrete", "Semi-Congrete"))


levels(houseHoldData$Types_of_Houses) <- c(1, 2, 3)
houseHoldData
```

```
##    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
```

```
fathersOccupation <- houseHoldData$Fathers_Occupation <- factor(houseHoldData$Fathers_Occupation,
                                        levels = c(1, 2, 3),
                                        labels = c("Farmer", "Driver", "Others"))
houseHoldData
```

```
##    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
```

```
femaleAndDrivers <- houseHoldData[houseHoldData$Sex == 2 &
                                houseHoldData$Fathers_Occupation == "Driver", ]
femaleAndDrivers
```

```
##   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
```

```
fiveOrMoreSiblings <- houseHoldData[houseHoldData$Siblings_at_School >= 5, ]
fiveOrMoreSiblings
```

```
##   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
```

```
#NUMBER 4 Interpret the graph
#The graph consistently Shows that the negative sentiments are the greatest in number,
# the neutral being the least and the positive sentiments in the middle. The greatest
# value of negative sentiments was recorded on July 15, 2020, the Positive sentiments
# maxing out in July 21, 2020, and the neutral on July 15, 2020 also. All of the
# sentiments' least number were all recorded on July 20, 2020
```