RWorksheet-3_Amuan#3B

Marivic Amuan

2023-10-16

```r
#1 A. Create a data frame using the table below.

respondent <- c(1:20)
sex <- c(2,2,1,2,2,2,2,2,2,2,2,1,2,2,2,2,2,2,2,1,2)
fathersOccuption <- 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)
typeOfHouse <- 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 = respondent,
  Sex = sex,
  FathersOccupation = fathersOccuption,
  PersonAtHome = personsAtHome,
  SiblingsAtSchool = siblingsAtSchool,
  HouseType = typeOfHouse
)

householdData
```

```
##    Respondents Sex FathersOccupation PersonAtHome SiblingsAtSchool HouseType
## 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
```

```r
#1 B. Describe the data. Get the structure or the summary of the data.

    str(householdData)
```

```
## '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 ...
##  $ FathersOccupation: 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 ...
##  $ HouseType       : num  1 2 3 1 1 3 3 1 2 3 ...
```

```r
    summary(householdData)
```

```
##   Respondents         Sex         FathersOccupation  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   HouseType
##  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
```

```r
#The data frame consists of 20 observations and 6 variables.
#EXPLAINATIONS:

  #Respondents: A variable is type of integer and there are 20 respondents in dataset.
  #Sex: A variable is type of numeric with values 1 and 2 represent the gender of male and female.
  #FathersOccupation: A variable is type of numeric with values 1, 2, and 3 , maybe it represent the le
  #PersonAtHome: A variable is type of numeric, and it represent the number of people at home for each
  #SiblingsAtSchool: A variable is type of numeric and represent the number of siblings that the respon
  #HouseType: A variable is numeric with values 1, 2, and 3, it represent kind of houses.


#1 C. Is the mean number of siblings attending is 5?

    siblingsMean <- mean(householdData$SiblingsAtSchool)

    siblingsMean
```

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

```r
#EXPLANATION: No, The mean number of siblings attending is 2.95.

#1 D. Extract the 1st two rows and then all the columns using the subsetting functions.
#Write the codes and its output.

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

```
##   Respondents Sex FathersOccupation PersonAtHome SiblingsAtSchool HouseType
## 1           1   2                1            5                6         1
## 2           2   2                3            7                4         2
```

#1 E. Extract 3rd and 5th row with 2nd and 4th column.
#Write the codes and its result.

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

```
##   Sex PersonAtHome
## 3   1            3
## 5   2            5
```

#1 F.Select the variable types of houses then store the vector that results as types_houses.
#Write the codes.

```r
typesHouses <- householdData$HouseType
typesHouses
```

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

#1 G. Select only all Males respondent that their father occupation was farmer.
#Write the codes and its output.

```r
householdData[householdData$Sex == 1 & householdData$FathersOccupation == "farmer", ]
```

```
## [1] Respondents      Sex              FathersOccupation PersonAtHome
## [5] SiblingsAtSchool HouseType
## <0 rows> (or 0-length row.names)
```

```r
householdData$FathersOccupation
```

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

#1 H. Select only all females respondent that have greater than or equal to 5 number of siblings attend
#Write the codes and its outputs.

```r
female <- householdData[householdData$SiblingsAtSchool >= 5,]
female
```

```
##   Respondents Sex FathersOccupation PersonAtHome SiblingsAtSchool HouseType
## 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
```

```r
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
otherRespondent <- c(1:10)
otherSex <- c("Male", "Female", "Female", "Male", "Male", "Female", "Female", "Male", "Female", "
otherFathersOccupuation <- c(1,2,3,3,1,2,2,3,1,3)
otherPersonsAtHome <- c(5,7,3,8,6,4,4,2,11,6)
otherSiblingsAtSchool <- c(2,3,0,5,2,3,1,2,6,2)
```

```r
    otherTypeOfHouse <- c("Wood", "Congrete", "Congrete", "Wood", "Semi-congrete", "Semi-congrete", "

    householdData <- data.frame(
        Respondents = otherRespondent,
        Sex = otherSex,
        FatherOccupation = otherFathersOccupution,
        PersonAtHome =otherPersonsAtHome,
        SiblingsAtSchool = otherSiblingsAtSchool,
        HouseType = otherTypeOfHouse
)

    write.csv(householdData, file = "householdData.csv")
```

#3 A. Import the csv file into the R environment. Write the codes.

```r
    householdData <- read.csv("householdData.csv")
    head(householdData)
```

```
##   X Respondents    Sex FatherOccupation PersonAtHome SiblingsAtSchool
## 1 1           1   Male                1            5                2
## 2 2           2 Female                2            7                3
## 3 3           3 Female                3            3                0
## 4 4           4   Male                3            8                5
## 5 5           5   Male                1            6                2
## 6 6           6 Female                2            4                3
##       HouseType
## 1          Wood
## 2      Congrete
## 3      Congrete
## 4          Wood
## 5 Semi-congrete
## 6 Semi-congrete
```

#3 B. Convert the Sex into factor using factor() function and change it into integer.
#[Legend:Male = 1 and Female = 2]. Write the R codes and its output.

```r
    householdData$Sex <- factor(householdData$Sex, levels = c("Male", "Female"))
    householdData$Sex <- as.integer(householdData$Sex)

    householdData$Sex
```

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

#OUTPUT 1 2 2 1 1 2 2 1 2 1

#3 C. Convert the Type of Houses into factor and change it into integer.
#[Legend: Wood= 1; Congrete = 2; Semi-Congrete = 3]. Write the R codes and its output.

```r
    householdData$HouseType <- factor(householdData$HouseType, levels = c("Wood", "Congrete", "Semi-c
    householdData$HouseType <- as.integer(householdData$HouseType)

    householdData$HouseType
```

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

#3 D. On father's occupation, factor it as Farmer = 1; Driver = 2; and Others = 3.
#What is the R code and its output?

```
householdData$FatherOccupation <- factor(householdData$FatherOccupation, levels = c(1,2,3), labels
householdData$FatherOccupation
```

```
## [1] Farmer, Driver, Others, Others, Farmer, Driver, Driver, Others, Farmer,
## [10] Others,
## Levels: Farmer, Driver, Others,
```

#3 E. Select only all females respondent that has a father whose occupation is driver.
#Write the codes and its output.

```
female <- householdData[householdData$Sex == 2 & householdData$FatherOccupation == "Driver",]
female
```

```
## [1] X               Respondents   Sex             FatherOccupation
## [5] PersonAtHome    SiblingsAtSchool HouseType
## <0 rows> (or 0-length row.names)
```

#3 F. Select the respondents that have greater than or equal to 5 number of siblings attending school.
#Write the codes and its output.

```
five <- householdData[householdData$SiblingsAtSchool >= 5,]
five
```

```
##   X Respondents Sex FatherOccupation PersonAtHome SiblingsAtSchool HouseType
## 4 4           4   1          Others,            8                5         1
## 9 9           9   2          Farmer,           11                6         3
```

#4. Interpret the graph.

#2022
#On July 14, the sentiments stated in tweets were as follows: 2500 unfavorable, 1500 neutral, and a ran

#On July 15, the number of negative tweets climbed from 2500 to 4000 compared to the previous day. The
#On July 17, the number of negative sentiments fell to between 3000 and 3500, while neutral sentiments
#On July 18, the number of unfavorable attitudes remained unchanged from the previous day, with a range
#On July 20, the quantity of negative attitudes in tweets fell from the previous day to over 2500. Posi
#On July 21, the number of tweets with negative feelings increased from the previous day, reaching over

#The particular reasons for these shifts in mood could only be explained with additional context and st