Worksheet - 3b

Angel Janica Marie De Jesus

2022-11-11

##1. A.Write the codes.

```
Respondents <- c(seq(1,20))
Sex <- c(2,2,1,2,2,2,2,2,2,2,1,2,2,2,2,2,2,2,1,2)
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)

df <- data.frame(Respondents, Sex, FathersOccupation, PersonsAtHome, SiblingsAtSchool, TypesOfHouses)

print (df)
```

```
##    Respondents 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
```

```
names(df)
```

```
## [1] "Respondents"       "Sex"               "FathersOccupation"
## [4] "PersonsAtHome"     "SiblingsAtSchool"  "TypesOfHouses"
```

##B. Describe the data: ##The data given are the respondents information according to sex, parent occupation, people at home, siblings at school and the type of houses.

##Get the structure or the summary of the data

```
summary(df)
```

```
##   Respondents         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. Is the mean number of siblings attending is 5?

```
mean(SiblingsAtSchool)
```

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

```
##No
```

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

```
sbst1 <- df[c(1,2),c(1:6)]
sbst1
```

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

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

```
sbst2 <- df[c(3,5),c(2,4)]
sbst2
```

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

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

```
types_houses <- (df$'Types Of Houses')
types_houses
```

```
## NULL
```

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

```
Male <- subset(df,Sex == '1' & FathersOccupation == '1')
Male
```

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

```
MaleFarm <- Male[c(2:3)]
MaleFarm
```

```
## [1] Sex               FathersOccupation
## <0 rows> (or 0-length row.names)
```

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

```
Female <- subset(df,Sex == '2' & SiblingsAtSchool >= '5')
Female
```

```
##    Respondents 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
```

```
FemaleSib <- Female[c(2,5)]
FemaleSib
```

```
##    Sex SiblingsAtSchool
## 1    2                6
## 7    2                5
## 13   2                5
## 14   2                5
## 18   2                5
```

##2. Write a R program to create an empty data frame. Using the following codes:

```
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. Describe the results : data frame has 5 variables which are the Ints,Doubles,Characters,Logical and Factors. But the result on the tables has no data available since there is no inputted data on each variable.

##3.Interpret the graph: The graph is titled as Figure 1: Sentiments of Tweets per day - Donald trump. The x = Day of Date / Sentiments consisting of dates from July 14, 15, 17,18, 20, 21, 2020. The y = Count of Sheet1 from 0 to 4000. The graph used is a bar graph consisting of 3 sentiments(Negative, Neutral and Positive).A legend is placed at the upper right corner indicating specific colors of every sentiments such as Negative = Red, Neutral= yellow and positive = blue.