Rworksheet#3b

Jhon Albert Paclibar

2024-10-07

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

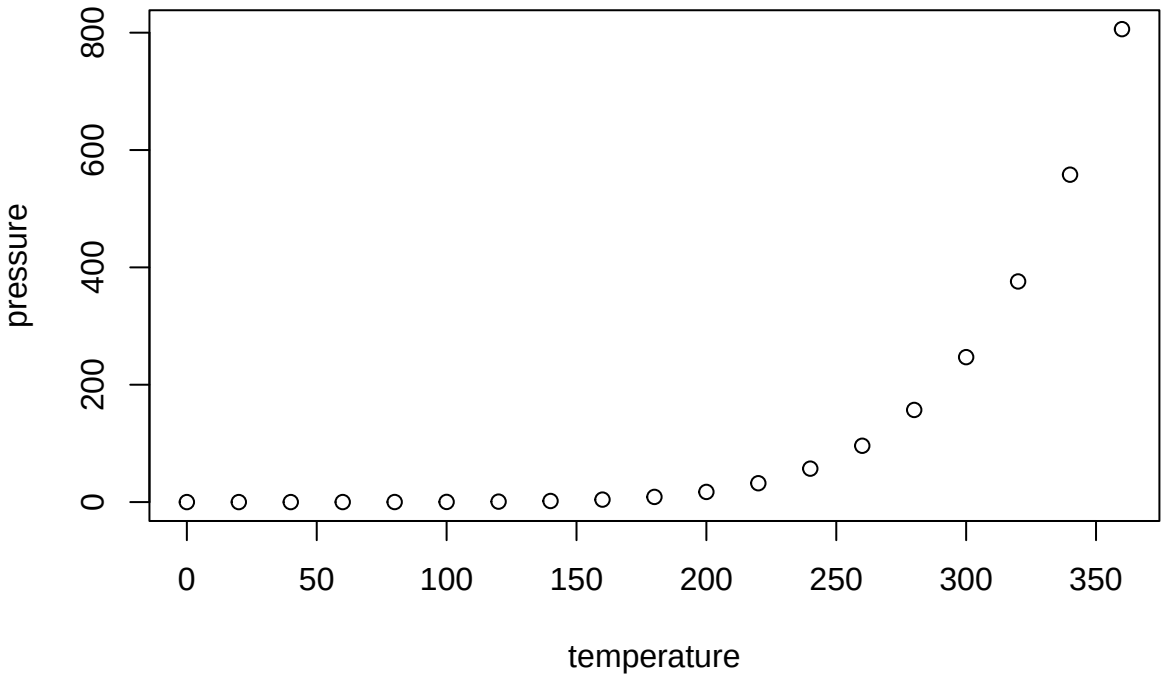
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed           dist       
##  Min.   : 4.0   Min.   :  2.00  
##  1st Qu.:12.0   1st Qu.: 26.00  
##  Median :15.0   Median : 36.00  
##  Mean   :15.4   Mean   : 42.98  
##  3rd Qu.:19.0   3rd Qu.: 56.00  
##  Max.   :25.0   Max.   :120.00
```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

#1. Create a data frame using the table below. #a

```
data <- data.frame(
Respondents = 1:20,
Sex = c(2, 1, 2, 2, 1, 1, 2, 2, 2, 1, 1, 2, 2, 1, 2, 2, 1, 2, 1, 2),
FathersOccupation = c(1, 2, 3, 1, 2, 1, 3, 2, 3, 3, 1, 3, 2, 1, 3, 1, 3, 3, 1, 1),
PersonsAtHome = c(5, 7, 3, 5, 5, 3, 6, 6, 7, 7, 3, 7, 4, 7, 8, 8, 3, 11, 8, 6),
SiblingsAtSchool = c(6, 4, 3, 2, 3, 3, 5, 5, 4, 5, 3, 7, 5, 2, 1, 3, 1, 5, 3, 2),
TypesOfHouses = c(1, 2, 3, 1, 3, 1, 3, 3, 3, 1, 3, 3, 3, 1, 3, 3, 3, 3, 3, 2)
)
data
```

```
##    Respondents Sex FathersOccupation PersonsAtHome SiblingsAtSchool
## 1            1   2                 1            5                6
## 2            2   1                 2            7                4
## 3            3   2                 3            3                3
## 4            4   2                 1            5                2
## 5            5   1                 2            5                3
## 6            6   1                 1            3                3
## 7            7   2                 3            6                5
## 8            8   2                 2            6                5
## 9            9   2                 3            7                4
## 10          10   1                 3            7                5
## 11          11   1                 1            3                3
## 12          12   2                 3            7                7
## 13          13   2                 2            4                5
## 14          14   1                 1            7                2
## 15          15   2                 3            8                1
## 16          16   2                 1            8                3
## 17          17   1                 3            3                1
## 18          18   2                 3           11                5
## 19          19   1                 1            8                3
## 20          20   2                 1            6                2
##    TypesOfHouses
## 1              1
## 2              2
## 3              3
## 4              1
## 5              3
## 6              1
## 7              3
## 8              3
## 9              3
## 10             1
## 11             3
## 12             3
## 13             3
## 14             1
## 15             3
## 16             3
## 17             3
## 18             3
```

```
## 19             3
## 20             2
```

#b

```
str(data)
```

```
## 'data.frame':    20 obs. of  6 variables:
##  $ Respondents      : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ Sex              : num  2 1 2 2 1 1 2 2 2 1 ...
##  $ FathersOccupation: num  1 2 3 1 2 1 3 2 3 3 ...
##  $ PersonsAtHome    : num  5 7 3 5 5 3 6 6 7 7 ...
##  $ SiblingsAtSchool : num  6 4 3 2 3 3 5 5 4 5 ...
##  $ TypesOfHouses    : num  1 2 3 1 3 1 3 3 3 1 ...
```

```
summary(data)
```

```
##   Respondents          Sex      FathersOccupation PersonsAtHome
##  Min.   : 1.00   Min.   :1.0   Min.   :1         Min.   : 3.00
##  1st Qu.: 5.75   1st Qu.:1.0   1st Qu.:1         1st Qu.: 4.75
##  Median :10.50   Median :2.0   Median :2         Median : 6.00
##  Mean   :10.50   Mean   :1.6   Mean   :2         Mean   : 5.95
##  3rd Qu.:15.25   3rd Qu.:2.0   3rd Qu.:3         3rd Qu.: 7.00
##  Max.   :20.00   Max.   :2.0   Max.   :3         Max.   :11.00
##  SiblingsAtSchool TypesOfHouses
##  Min.   :1.00     Min.   :1.00
##  1st Qu.:2.75     1st Qu.:1.75
##  Median :3.00     Median :3.00
##  Mean   :3.60     Mean   :2.40
##  3rd Qu.:5.00     3rd Qu.:3.00
##  Max.   :7.00     Max.   :3.00
```

#c

```
mean(data$SiblingsAtSchool)
```

```
## [1] 3.6
```

#d

```
data[1:2, ]
```

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

#e

```
data[c(3, 5), c(2, 4)]
```

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

#f

```
types_houses <- data$TypesOfHouses
types_houses
```

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

#g

```
MFarmers <- subset(data, Sex == 1 & FathersOccupation == 1)
MFarmers
```

```
##    Respondents Sex FathersOccupation PersonsAtHome SiblingsAtSchool
## 6            6   1                 1            3                3
## 11          11   1                 1            3                3
## 14          14   1                 1            7                2
## 19          19   1                 1            8                3
##    TypesOfHouses
## 6              1
## 11             3
## 14             1
## 19             3
```

#h

```
FemaleSiblings <- subset(data, Sex == 2 & SiblingsAtSchool >= 5)
FemaleSiblings
```

```
##    Respondents Sex FathersOccupation PersonsAtHome SiblingsAtSchool
## 1            1   2                 1            5                6
## 7            7   2                 3            6                5
## 8            8   2                 2            6                5
## 12          12   2                 3            7                7
## 13          13   2                 2            4                5
## 18          18   2                 3           11                5
##    TypesOfHouses
## 1              1
## 7              3
## 8              3
## 12             3
## 13             3
## 18             3
```

#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: #The data frame is empty, it also contains 5 columns which is integers, doubles, characters, logicals, and factors. #3. Create a .csv file of this. Save it as HouseholdData.csv

```
houseData <- read.csv("HouseholdData.csv")
```

#b

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

```
##    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
##    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
```

#c

```
houseData$Type.of.Houses <- factor(houseData$Type.of.Houses, levels = c("Wood", "Semi-Congrete", "Congr
houseData$Type.of.Houses <- as.integer(houseData$Type.of.Houses)
houseData
```

```
##    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
##    Type.of.Houses
## 1               1
## 2               3
## 3               3
## 4               1
```

```
## 5               2
## 6               2
## 7               1
## 8               2
## 9               2
## 10              3
```

#d

houseData$Fathers.Occupation <- factor(HouseDataFathers.Occupation, levels = c("Farmer", "Driver", "Others"),labels = c(1,2,3)) houseData

#e

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

```
## [1] Respondents        Sex                Fathers.Occupation Persons.at.Home
## [5] Siblings.at.School Type.of.Houses
## <0 rows> (or 0-length row.names)
```

#f

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

```
##   Respondents Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 4           4   1                  3               8                  5
## 9           9   2                  1              11                  6
##   Type.of.Houses
## 4              1
## 9              2
```

#4. Interpret the graph. #The graph shows the number of sentiments of tweets from July 14 to July 21 year 2020. The graph shows the sentiments of the people wether they are negative, positive, or neutral about their tweet.