

RWorksheet_Gallo#3b

2023-10-11

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

#1A. Write the codes

```
HouseholdData <- data.frame(  
  Respond_1 = c(1:20),  
  
  Sex = c("Female","Female","Male","Female","Female","Female","Female","Female","Female","Female","Female","Male",  
  
  FatherOccupation = c("Farmer","Others","Others","Others","Farmer","Driver","Others","Farmer","Farmer",  
  
  PersonAtHome = 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("Wood","Semi-Concrete","Concrete","Wood","Wood","Concrete","Concrete","Wood","Semi-  
)  
HouseholdData
```

##	Respond_1	Sex	FatherOccupation	PersonAtHome	SiblingsAtSchool
## 1	1	Female	Farmer	5	6
## 2	2	Female	Others	7	4
## 3	3	Male	Others	3	4
## 4	4	Female	Others	8	1
## 5	5	Female	Farmer	5	2
## 6	6	Female	Driver	9	1
## 7	7	Female	Others	6	5
## 8	8	Female	Farmer	7	3
## 9	9	Female	Farmer	8	1
## 10	10	Female	Farmer	4	2
## 11	11	Male	Others	7	3
## 12	12	Female	Driver	5	2
## 13	13	Female	Farmer	4	5
## 14	14	Female	Others	7	5
## 15	15	Female	Others	8	2
## 16	16	Female	Farmer	8	1
## 17	17	Female	Others	3	2
## 18	18	Female	Farmer	11	5
## 19	19	Male	Driver	7	3
## 20	20	Female	Farmer	6	2
##	TypesOfHouses				
## 1	Wood				
## 2	Semi-Concrete				
## 3	Concrete				
## 4	Wood				
## 5	Wood				
## 6	Concrete				
## 7	Concrete				
## 8	Wood				
## 9	Semi-Concrete				
## 10	Concrete				

```
## 11 Semi-Concrete
## 12      Concrete
## 13 Semi-Concrete
## 14 Semi-Concrete
## 15      Concrete
## 16      Concrete
## 17      Concrete
## 18      Concrete
## 19      Concrete
## 20 Semi-Concrete
```

#1b. Describe the data. Get the structure or the summary of the data

```
summary(HouseholdData)
```

```
##      Respond_1      Sex      FatherOccupation      PersonAtHome
## Min.   : 1.00   Length:20      Length:20      Min.   : 3.0
## 1st Qu.: 5.75   Class :character  Class :character 1st Qu.: 5.0
## Median :10.50   Mode  :character  Mode  :character Median : 7.0
## Mean   :10.50                                     Mean   : 6.4
## 3rd Qu.:15.25                                     3rd Qu.: 8.0
## Max.    :20.00                                     Max.    :11.0
## SiblingsAtSchool TypesOfHouses
## Min.   :1.00      Length:20
## 1st Qu.:2.00      Class :character
## Median :2.50      Mode  :character
## Mean   :2.95
## 3rd Qu.:4.25
## Max.    :6.00
```

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

```
MeanNumSib <- mean(HouseholdData$SiblingsAtSchool)
IsMean5 <- MeanNumSib == 5
print(IsMean5)
```

```
## [1] FALSE
```

False

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

```
ftrac <- HouseholdData[1:2, ]
print(ftrac)
```

```
##      Respond_1      Sex      FatherOccupation      PersonAtHome      SiblingsAtSchool      TypesOfHouses
## 1           1 Female           Farmer           5           6           Wood
## 2           2 Female           Others           7           4 Semi-Concrete
```

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

```
selrowcol <- HouseholdData[c(3, 5), c(2, 4)]
print(selrowcol)
```

```
##      Sex      PersonAtHome
## 3   Male           3
## 5 Female           5
```

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

```
types_houses <- HouseholdData$TypeOfHouses
```

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

```
MaleFarmers <- HouseholdData[HouseholdData$Sex == "Male" & HouseholdData$FatherOccupation == "Farmer", ]  
print(MaleFarmers)
```

```
## [1] Respond_1      Sex      FatherOccupation PersonAtHome  
## [5] SiblingsAtSchool TypesOfHouses  
## <0 rows> (or 0-length row.names)
```

#1h. Select only all females respondent that have greater than or equal to 5 number of siblings attending

```
FemGT5Sib <- HouseholdData[HouseholdData$Sex == "Female" & HouseholdData$SiblingsAtSchool >= 5, ]  
print(FemGT5Sib)
```

```
##      Respond_1      Sex FatherOccupation PersonAtHome SiblingsAtSchool  
## 1             1 Female           Farmer             5             6  
## 7             7 Female           Others             6             5  
## 13            13 Female           Farmer             4             5  
## 14            14 Female           Others             7             5  
## 18            18 Female           Farmer            11             5  
##      TypesOfHouses  
## 1             Wood  
## 7             Concrete  
## 13 Semi-Concrete  
## 14 Semi-Concrete  
## 18             Concrete
```

#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
```

#2a. The data frame is empty

#3

```
HouseholdData <- data.frame(
```

```

Respondents = c(1:10),
Sex = c("Male", "Female", "Female", "Male", "Male", "Female", "Female", "Male", "Female", "Male"),
FathersOccupation = c(1, 2, 3, 3, 1, 2, 2, 3, 1, 3),
PersonAtHome = c(5, 7, 3, 8, 6, 4, 4, 2, 11, 6),
SiblingsAtSchool = c(2, 3, 0, 5, 2, 3, 1, 2, 6, 2),
TypesOfHouse = c("Wood", "Congrete", "Congrete", "Wood", "Semi-Congrete", "Semi-Congrete", "Wood", "Semi-Congrete", "Wood", "Semi-Congrete")
)
HouseholdData

```

```

##      Respondents      Sex FathersOccupation PersonAtHome SiblingsAtSchool
## 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
##      TypesOfHouse
## 1             Wood
## 2             Congrete
## 3             Congrete
## 4             Wood
## 5      Semi-Congrete
## 6      Semi-Congrete
## 7             Wood
## 8      Semi-Congrete
## 9      Semi-Congrete
## 10            Congrete

```

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

```

write.csv(HouseholdData, file = "HouseholdData.csv", row.names = FALSE)
ImporteData <- read.csv("HouseholdData.csv")

```

#3b.. Convert the Sex into factor using factor() function and change it into integer. [Legend: Male = 1 and Female = 2]

```

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

```

#3c.Convert the Type of Houses into factor and change it into integer. [Legend: Wood= 1; Congrete = 2; Semi-Congrete = 3]

```

ImporteData$TypesOfHouse <- factor(ImporteData$TypesOfHouse, levels = c("Wood", "Congrete", "Semi-Congrete"))
ImporteData$TypesOfHouse <- as.integer(ImporteData$TypesOfHouse)

```

```

ImporteData$TypesOfHouse <- factor(ImporteData$TypesOfHouse, levels = c("Wood", "Congrete", "Semi-Congrete"))
ImporteData$TypesOfHouse <- as.integer(ImporteData$TypesOfHouse)

```

#3d.On father's occupation, factor it as Farmer = 1; Driver = 2; and Others = 3. What is the R code and how to change it into integer?

```

ImporteData$FathersOccupation <- factor(ImporteData$FathersOccupation, levels = c("Farmer", "Driver", "Others"))
ImporteData$FathersOccupation <- as.integer(ImporteData$FathersOccupation)

```

```
#3e. Select only all females respondent that has a father whose occupation is driver. Write the codes and
FemaleDrivers <- ImporteData[ImporteData$Sex == 2 & ImporteData$FathersOccupation == 2, ]
print(FemaleDrivers)
```

```
##      Respondents Sex FathersOccupation PersonAtHome SiblingsAtSchool
## NA              NA NA              NA              NA              NA
## NA.1            NA NA              NA              NA              NA
## NA.2            NA NA              NA              NA              NA
## NA.3            NA NA              NA              NA              NA
## NA.4            NA NA              NA              NA              NA
##      TypesOfHouse
## NA              NA
## NA.1            NA
## NA.2            NA
## NA.3            NA
## NA.4            NA
```

```
#3f. Select the respondents that have greater than or equal to 5 number of siblings attending school. Write the codes and
GT5Sib <- ImporteData[ImporteData$SiblingsAtSchool >= 5, ]
print(GT5Sib)
```

```
##      Respondents Sex FathersOccupation PersonAtHome SiblingsAtSchool TypesOfHouse
## 4              4   1              NA              8              5              NA
## 9              9   2              NA              11             6              NA
```

#4. Interpret the graph

The graph in figure 3 represents the sentiments of people every day that has a major impact on our work