ML_Assignment_1

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:

Loading Tech-Sales-Reps.csv into R data frame

```
MyFile<-read.csv("Tech-Sales-Reps.csv")
str(MyFile)
```

```
## 'data.frame':    21993 obs. of  11 variables:
##  $ Sales_Rep  : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ Business   : chr  "Hardware" "Hardware" "Software" "Hardware" ...
##  $ Age        : int  59 52 47 61 39 28 25 51 34 38 ...
##  $ Female     : int  1 0 1 0 0 0 1 1 0 1 ...
##  $ Years      : int  2 10 1 2 1 6 1 10 4 1 ...
##  $ College    : chr  "Yes" "Yes" "Yes" "Yes" ...
##  $ Personality: chr  "Diplomat" "Diplomat" "Explorer" "Diplomat" ...
##  $ Certficates: int  1 4 1 3 5 1 5 0 2 5 ...
##  $ Feedback   : num  2.01 3.64 3.88 2.7 3.44 2.43 3.3 2.15 2.91 1.23 ...
##  $ Salary     : int  70200 133000 52600 96000 122000 60000 68000 43800 92000 73400 ...
##  $ NPS        : int  5 10 8 6 7 6 6 5 7 6 ...
```

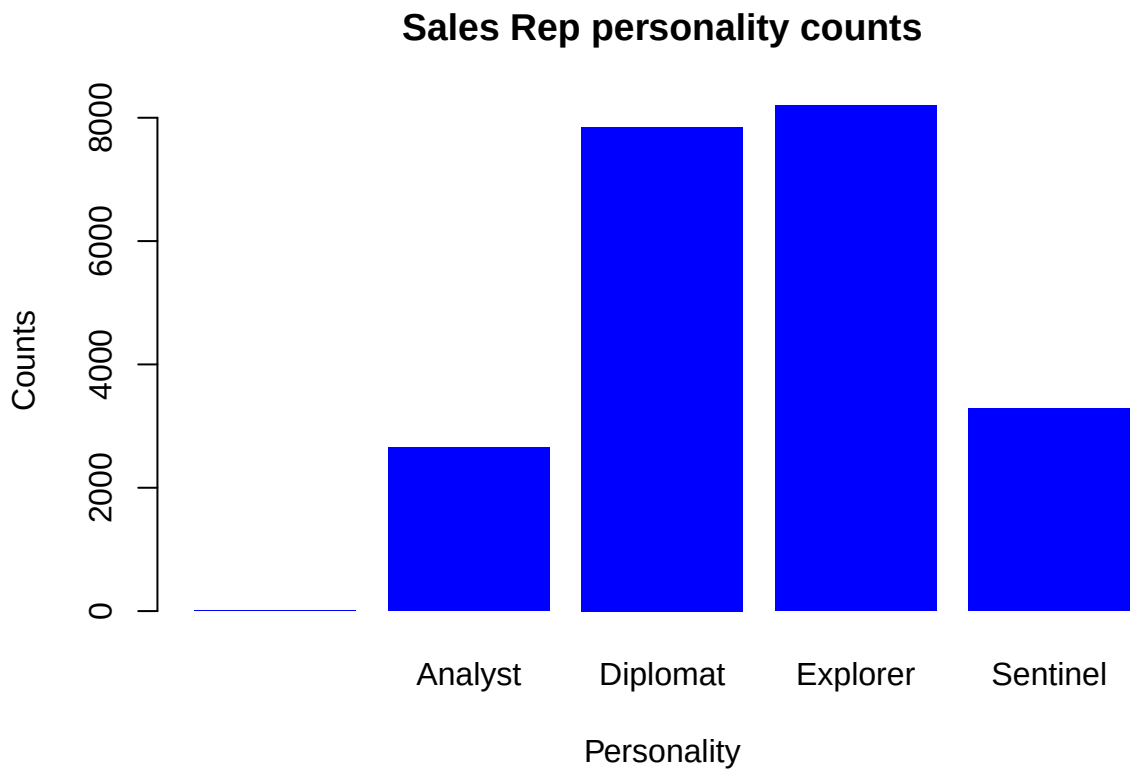
Creating frequency table on qualatitive/catagorical data column "Personality"

```
FTable<-table(MyFile$Personality)
FTable
```

```
##
##           Analyst Diplomat Explorer Sentinel
##        3     2659     7849     8200     3282
```

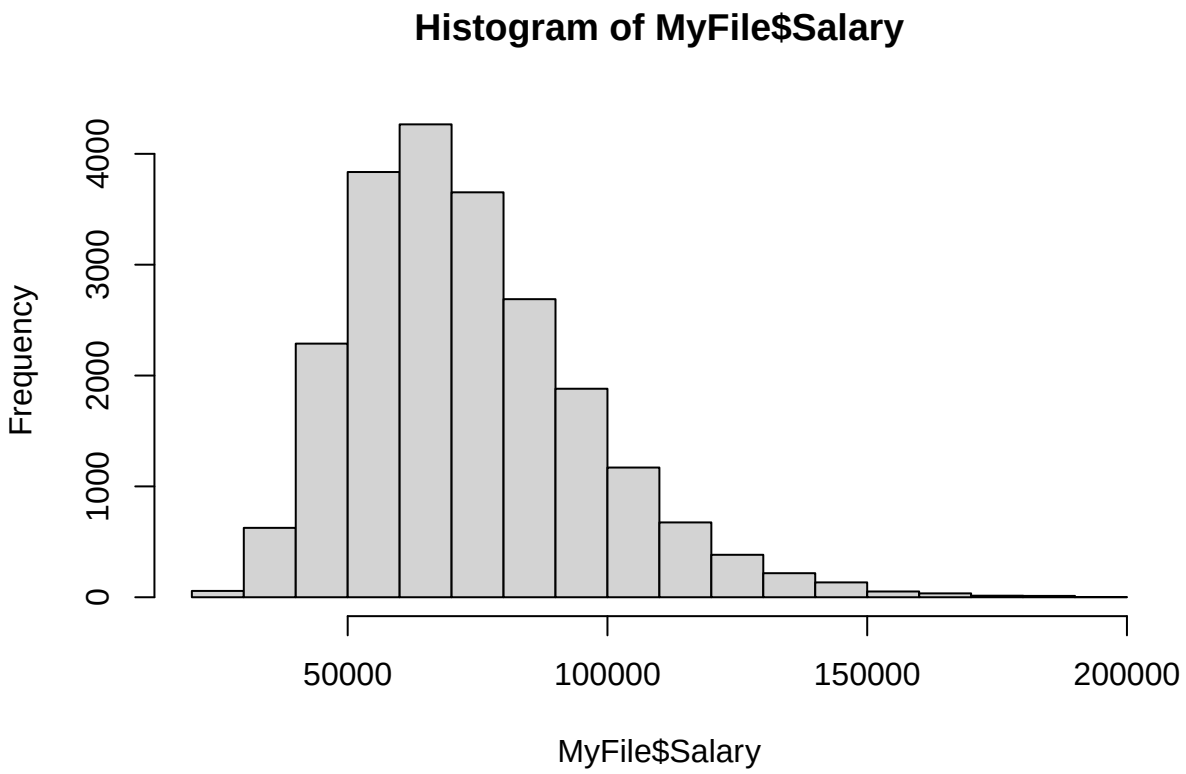
plotting bar chart out of the 'Personality' qualitative data

```
barplot ( FTable , main="Sales Rep personality counts",  xlab = "Personality", ylab ="Counts", col = "bl
```

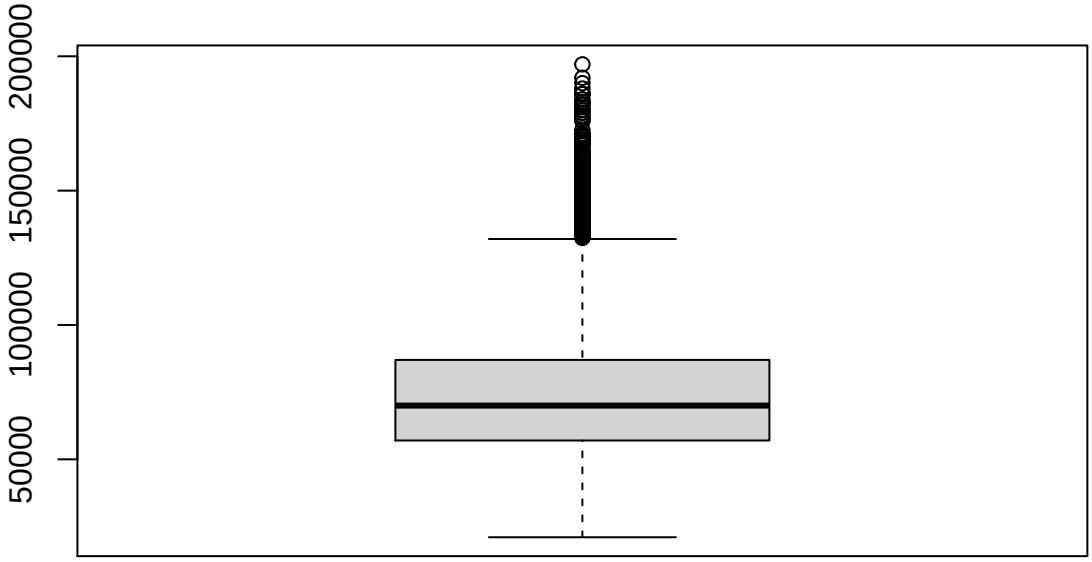


Creating graph on Quantitative data on “Salary” column

```
hist(MyFile$Salary)
```




```
boxplot(MyFile$Salary)
```



Descriptive Statistics for Quantitative data

```
summary(MyFile$Salary)
```

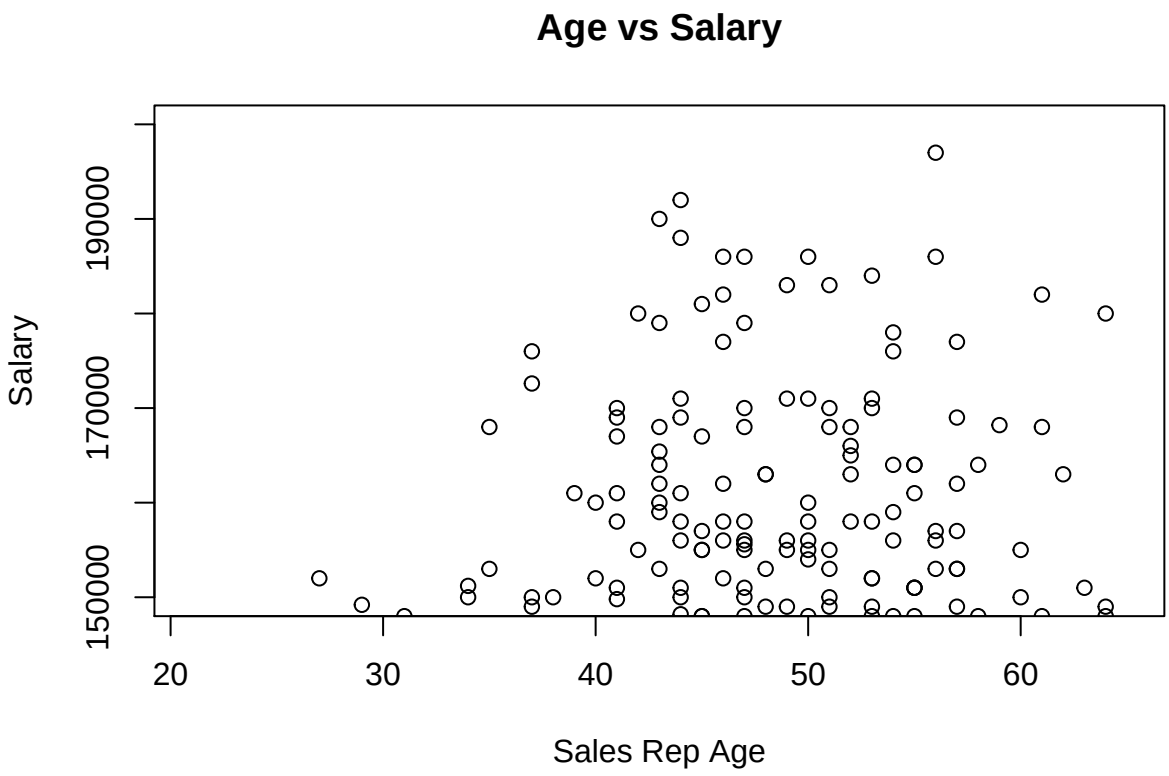
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   21000   57000   70000   73674   87000  197000       3
```

Creating Scatterplot

```
input<-MyFile[,c('Age','Salary')]
head(input)
```

```
##   Age Salary
## 1  59  70200
## 2  52 133000
## 3  47  52600
## 4  61  96000
## 5  39 122000
## 6  28  60000
```

```
plot(x=input$Age, y=input$Salary, xlab = "Sales Rep Age", ylab = "Salary",
        ylim = c(150000,200000),
   main = "Age vs Salary" )
```



Data Transformation of Salary column into 1000/K representation

```
MyFile2<-transform(MyFile, SalaryinK= MyFile$Salary/1000)
head(MyFile2)
```

```
##   Sales_Rep Business Age Female Years College Personality Certficates Feedback
## 1         1 Hardware  59      1     2     Yes    Diplomat           1     2.01
## 2         2 Hardware  52      0    10     Yes    Diplomat           4     3.64
## 3         3 Software  47      1     1     Yes    Explorer           1     3.88
## 4         4 Hardware  61      0     2     Yes    Diplomat           3     2.70
## 5         5 Software  39      0     1      No    Diplomat           5     3.44
## 6         6 Hardware  28      0     6     Yes    Explorer           1     2.43
##   Salary NPS SalaryinK
## 1  70200   5      70.2
## 2 133000  10     133.0
## 3  52600   8      52.6
## 4  96000   6      96.0
## 5 122000   7     122.0
## 6  60000   6      60.0
```