Assignment 1 Machine Learning

getwd()

## [1] "/Users/taliadagan/Desktop/Business Analytics 1"

setwd("/Users/taliadagan/Desktop/Business Analytics 1")  
  
#Import the dataset into R  
Myfile <- read.csv("Tech Sales Reps 2.csv")  
head(Myfile)

## Sales\_Rep Business Age Female Years College Personality Certficates Feedback  
## 1 21832 Software 29 1 2 Yes Diplomat 3 1.21  
## 2 21833 Software 37 1 1 Yes Diplomat 1 2.55  
## 3 21834 Hardware 60 0 2 Yes Diplomat 2 2.46  
## 4 21835 Software 38 0 1 Yes Explorer 0 3.50  
## 5 21836 Hardware 33 0 2 Yes Diplomat 3 2.70  
## 6 21837 Software 36 0 2 No Diplomat 2 1.95  
## Salary NPS  
## 1 59200 5  
## 2 47600 3  
## 3 62000 4  
## 4 76000 7  
## 5 104000 8  
## 6 54000 5

#Print out descriptive statistics for a selection of quantitative and categorical variables.

FTable <- table (Myfile$College)  
FTable

##   
## No Yes   
## 30 119

FTable <- table (Myfile$Personality)  
FTable

##   
## Analyst Diplomat Explorer Sentinel   
## 13 66 50 20

FTable <- table (Myfile$Business)  
FTable

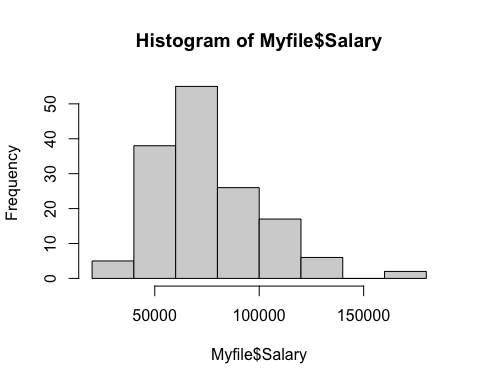
##   
## Hardware Software   
## 70 79

summary(Myfile)

## Sales\_Rep Business Age Female   
## Min. :21832 Length:149 Min. :21.00 Min. :0.0000   
## 1st Qu.:21869 Class :character 1st Qu.:31.00 1st Qu.:0.0000   
## Median :21906 Mode :character Median :40.00 Median :0.0000   
## Mean :21906 Mean :40.65 Mean :0.3557   
## 3rd Qu.:21943 3rd Qu.:49.00 3rd Qu.:1.0000   
## Max. :21980 Max. :65.00 Max. :1.0000   
## Years College Personality Certficates   
## Min. : 1.000 Length:149 Length:149 Min. :0.000   
## 1st Qu.: 1.000 Class :character Class :character 1st Qu.:2.000   
## Median : 2.000 Mode :character Mode :character Median :3.000   
## Mean : 2.255 Mean :2.805   
## 3rd Qu.: 2.000 3rd Qu.:4.000   
## Max. :12.000 Max. :6.000   
## Feedback Salary NPS   
## Min. :1.140 Min. : 27000 Min. : 2.000   
## 1st Qu.:2.140 1st Qu.: 58800 1st Qu.: 5.000   
## Median :2.890 Median : 70600 Median : 7.000   
## Mean :2.734 Mean : 75569 Mean : 6.544   
## 3rd Qu.:3.320 3rd Qu.: 88200 3rd Qu.: 8.000   
## Max. :4.000 Max. :177000 Max. :10.000

#Transform at least one variable. It doesn’t matter what the transformation is.

Myfile$Female=as.factor(Myfile$Female)  
  
#Plot at least one quantitative variable, and one scatterplot  
hist ( Myfile$Salary )



plot(Myfile$Age,Myfile$Salary)

