Assignment 1

Hanish

1/30/2022

#Reference  
  
#https://www.kaggle.com/tejashvi14/employee-future-prediction  
  
data1 = read.csv("C:/Users/Hanish Bhogadi/Desktop/Employees data.csv")  
data1

## Education JoiningYear City PaymentTier Age Gender EverBenched  
## 1 Bachelors 2017 Bangalore 3 34 Male No  
## 2 Bachelors 2013 Pune 1 28 Female No  
## 3 Bachelors 2014 New Delhi 3 38 Female No  
## 4 Masters 2016 Bangalore 3 27 Male No  
## 5 Masters 2017 Pune 3 24 Male Yes  
## 6 Bachelors 2016 Bangalore 3 22 Male No  
## 7 Bachelors 2015 New Delhi 3 38 Male No  
## 8 Bachelors 2016 Bangalore 3 34 Female No  
## 9 Bachelors 2016 Pune 3 23 Male No  
## 10 Masters 2017 New Delhi 2 37 Male No  
## 11 Masters 2012 Bangalore 3 27 Male No  
## 12 Bachelors 2016 Pune 3 34 Male No  
## 13 Bachelors 2018 Pune 3 32 Male Yes  
## 14 Bachelors 2016 Bangalore 3 39 Male No  
## ExperienceInCurrentDomain LeaveOrNot  
## 1 0 0  
## 2 3 1  
## 3 2 0  
## 4 5 1  
## 5 2 1  
## 6 0 0  
## 7 0 0  
## 8 2 1  
## 9 1 0  
## 10 2 0  
## 11 5 1  
## 12 3 0  
## 13 5 1  
## 14 2 0

#ABOVE IS THE DATA USED FOR THE ASSIGNMENT  
#Quantitative Descriptive Statistics  
  
  
mean(data1$Age)

## [1] 31.21429

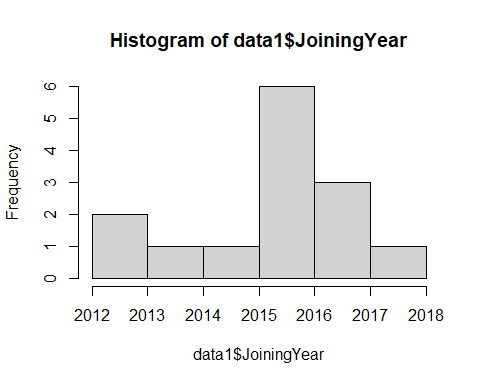
#ABOVE IS THE MEAN OF THE AGE COLUMN IN THE DATA  
#Categorical Variables  
table(data1$City)

##   
## Bangalore New Delhi Pune   
## 6 3 5

data1$Age\_mode = median(data1$Age)\* median(data1$Age)\*median(data1$Age) - mean(data1$Age)\* mean(data1$Age)  
data1$Age\_mode

## [1] 34962.67 34962.67 34962.67 34962.67 34962.67 34962.67 34962.67 34962.67  
## [9] 34962.67 34962.67 34962.67 34962.67 34962.67 34962.67

#Below is the Histogram of Quantitative Variable  
hist(data1$JoiningYear)



#Below is the Scatter plot   
x <- data1$Age  
y <- data1$JoiningYear  
plot(x,y, main = "Age and year", xlab = "Age", ylab = "Year")

