Assignment 1

Praneeth Simha

2023-01-29

##Downloaded the dataset from Kagle website

airline <- read.csv("C:/Users/ADMIN/Downloads/airline.csv")

##Descriptive Statictics

summary(airline)

## X id Gender Customer.Type   
## Min. : 0 Min. : 1 Length:103904 Length:103904   
## 1st Qu.: 25976 1st Qu.: 32534 Class :character Class :character   
## Median : 51952 Median : 64857 Mode :character Mode :character   
## Mean : 51952 Mean : 64924   
## 3rd Qu.: 77927 3rd Qu.: 97368   
## Max. :103903 Max. :129880   
##   
## Age Type.of.Travel Class Flight.Distance  
## Min. : 7.00 Length:103904 Length:103904 Min. : 31   
## 1st Qu.:27.00 Class :character Class :character 1st Qu.: 414   
## Median :40.00 Mode :character Mode :character Median : 843   
## Mean :39.38 Mean :1189   
## 3rd Qu.:51.00 3rd Qu.:1743   
## Max. :85.00 Max. :4983   
##   
## Inflight.wifi.service Departure.Arrival.time.convenient Ease.of.Online.booking  
## Min. :0.00 Min. :0.00 Min. :0.000   
## 1st Qu.:2.00 1st Qu.:2.00 1st Qu.:2.000   
## Median :3.00 Median :3.00 Median :3.000   
## Mean :2.73 Mean :3.06 Mean :2.757   
## 3rd Qu.:4.00 3rd Qu.:4.00 3rd Qu.:4.000   
## Max. :5.00 Max. :5.00 Max. :5.000   
##   
## Gate.location Food.and.drink Online.boarding Seat.comfort   
## Min. :0.000 Min. :0.000 Min. :0.00 Min. :0.000   
## 1st Qu.:2.000 1st Qu.:2.000 1st Qu.:2.00 1st Qu.:2.000   
## Median :3.000 Median :3.000 Median :3.00 Median :4.000   
## Mean :2.977 Mean :3.202 Mean :3.25 Mean :3.439   
## 3rd Qu.:4.000 3rd Qu.:4.000 3rd Qu.:4.00 3rd Qu.:5.000   
## Max. :5.000 Max. :5.000 Max. :5.00 Max. :5.000   
##   
## Inflight.entertainment On.board.service Leg.room.service Baggage.handling  
## Min. :0.000 Min. :0.000 Min. :0.000 Min. :1.000   
## 1st Qu.:2.000 1st Qu.:2.000 1st Qu.:2.000 1st Qu.:3.000   
## Median :4.000 Median :4.000 Median :4.000 Median :4.000   
## Mean :3.358 Mean :3.382 Mean :3.351 Mean :3.632   
## 3rd Qu.:4.000 3rd Qu.:4.000 3rd Qu.:4.000 3rd Qu.:5.000   
## Max. :5.000 Max. :5.000 Max. :5.000 Max. :5.000   
##   
## Checkin.service Inflight.service Cleanliness Departure.Delay.in.Minutes  
## Min. :0.000 Min. :0.00 Min. :0.000 Min. : 0.00   
## 1st Qu.:3.000 1st Qu.:3.00 1st Qu.:2.000 1st Qu.: 0.00   
## Median :3.000 Median :4.00 Median :3.000 Median : 0.00   
## Mean :3.304 Mean :3.64 Mean :3.286 Mean : 14.82   
## 3rd Qu.:4.000 3rd Qu.:5.00 3rd Qu.:4.000 3rd Qu.: 12.00   
## Max. :5.000 Max. :5.00 Max. :5.000 Max. :1592.00   
##   
## Arrival.Delay.in.Minutes satisfaction   
## Min. : 0.00 Length:103904   
## 1st Qu.: 0.00 Class :character   
## Median : 0.00 Mode :character   
## Mean : 15.18   
## 3rd Qu.: 13.00   
## Max. :1584.00   
## NA's :310

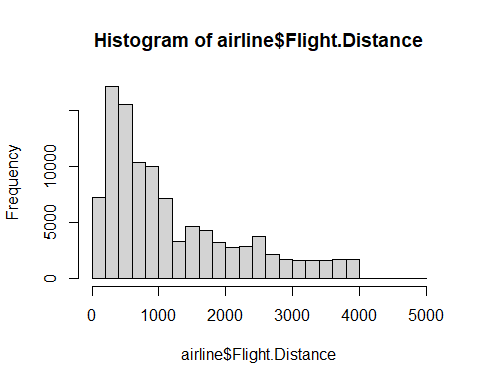
##Transformation of Age Varibale

airline$Age <- mean(airline$Age)/sd(airline$Age)  
head(airline$Age)

## [1] 2.605346 2.605346 2.605346 2.605346 2.605346 2.605346

##.Histogram and Scatterplot

hist(airline$Flight.Distance)



library(ggplot2)  
ggplot(airline) +  
 aes(x = Arrival.Delay.in.Minutes, y =Departure.Delay.in.Minutes ) +  
 geom\_point(shape = "circle", size = 1.5, colour ="violet") +  
 theme\_minimal()

## Warning: Removed 310 rows containing missing values (`geom\_point()`).

