## Machine Lerning Assingment

Importing dataset \_ Cost of US Airlines

Source (<https://vincentarelbundock.github.io/Rdatasets/articles/data.html>)

setwd(“~/Machine Learning Assingment”)

library(readr)

USAirlines <- read\_csv(“USAirlines.csv”)

## Descriptive statistics for quantitative and categorical variables)

summary(USAirlines$cost) Min. 1st Qu. Median Mean 3rd Qu. Max. 68978 292046 637001 1122524 1345968 4748320

summary(USAirlines$price) Min. 1st Qu. Median Mean 3rd Qu. Max. 103795 129848 357434 471683 849840 1015610

## Transforming Variable Price

summary(log10(USAirlines$price +1)) Min. 1st Qu. Median Mean 3rd Qu. Max. 5.016 5.113 5.553 5.546 5.929 6.007

summary(sqrt(USAirlines$price)) Min. 1st Qu. Median Mean 3rd Qu. Max. 322.2 360.3 597.9 641.5 921.9 1007.8

## Plotting quantitative variable cost

boxplot(USAirlines$cost)

hist(USAirlines$cost)

## scatterplot

plot(USAirlinesprice)

plot(USAirlinesprice, main = “Scatterplot for cost and price of Usairlines”)