**Team Unity**

Sai Rohith AP21110010219

Uday Sai Chaganti AP21110010200

Jithin Atluri AP21110010228

Ritesh Jadhav AP21110010260

Chaitanya Sai AP21110010253

Lab Assignment – 2

CSE 338L - Applied Data science Lab

Date: 30-01-2024

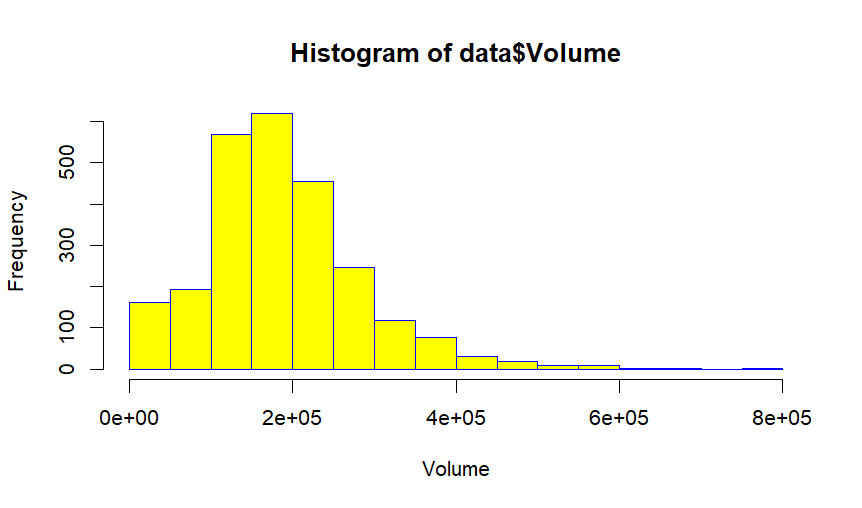
**Write R – Programming to plot various charts and graphs. You have to consider**

**minimum two popular data sets and draw all the statistical observations.**

**R Language Code**

***Goldstock (Kaggle Dataset)***

| library(readxl)  # Setting the working directory setwd("C:/Users/Nutakki/Desktop")  # Reading data from CSV file data <- read.csv("goldstock.csv")  # Printing data to check if it's read correctly print(data)  # Calculating mean mean\_value <- mean(data$Volume) print(mean\_value)  # Calculating median median\_value <- median(data$Volume) print(median\_value)  # Calculating mode getmode <- function(data)  {  uniqdata <- unique(data)  uniqdata[which.max(tabulate(match(data, uniqdata)))] }  mode\_value <- getmode(data$Volume) print(mode\_value)  # Calculating variance variance <- var(data$Volume) print(variance)  # Calculating standard deviation sd\_value <- sqrt(variance) print(sd\_value)  # Plotting histogram hist(data$Volume, xlab = "Volume", ylab = "Frequency", col = "yellow", border = "blue") |
| --- |



***Glassdoor (Kaggle Dataset)***

| library(readxl)  # Setting the working directory  setwd("C:/Users/Nutakki/Desktop")  # Reading data from CSV file  data <- read.csv("glassdoor.csv")  # Changing column name to "salary\_avg\_estimate"  colnames(data)[colnames(data) == "work\_life\_balance\_rating"] <- "salary\_avg\_estimate"  # Filter out non-numeric values  numeric\_salary <- as.numeric(as.character(data$salary\_avg\_estimate))  numeric\_salary <- numeric\_salary[!is.na(numeric\_salary)]  # Check if there are any NAs left after filtering  if (any(is.na(numeric\_salary))) {  cat("Warning: NAs were found after filtering non-numeric values.\n")  # Remove NAs from numeric\_salary  numeric\_salary <- numeric\_salary[!is.na(numeric\_salary)]  }  # Calculating mean  mean\_value <- mean(numeric\_salary)  print(mean\_value)  # Calculating median  median\_value <- median(numeric\_salary)  print(median\_value)  # Calculating mode  getmode <- function(data) {  uniqdata <- unique(data)  uniqdata[which.max(tabulate(match(data, uniqdata)))]  }  mode\_value <- getmode(numeric\_salary)  print(mode\_value)  # Calculating variance  variance <- var(numeric\_salary)  print(variance)  # Calculating standard deviation  sd\_value <- sqrt(variance)  print(sd\_value)  # Plotting histogram  hist(numeric\_salary, xlab = "Salary Average Estimate", ylab = "Frequency", col = "yellow", border = "blue") |
| --- |

