**1. a) Write R program to find R-Mean, Median & Mode with the sample data.**

**Source code:**

install.packages("datasets")

library(datasets)

airquality <- datasets::airquality

summary(airquality)

var(airquality$Solar.R,na.rm = T)

summary(airquality$Temp)

summary(airquality$Wind)

plot(airquality$Wind)

plot(airquality$Temp,airquality$Wind)

plot(airquality)

plot(airquality$Wind, type= "b") # p: points, l: lines,b: both

plot(airquality$Wind, xlab = 'ozone Concentration',

ylab = 'No of Instances', main = 'Ozone levels in NY city',

col = 'blue')

**# Horizontal bar plot**

barplot(airquality$Ozone, main = 'Ozone Concenteration in air',

ylab = 'ozone levels', col= 'blue',horiz = F)

**#Histogram**

hist(airquality$Temp)

hist(airquality$Temp,

main = 'Solar Radiation values in air',

xlab = 'Solar rad.', col='blue')

**#Single box plot**

boxplot(airquality$Temp,main="Boxplot")

**# Multiple box plots**

boxplot(airquality[,1:4],main='Multiple')

**Mean**

# Create a vector.

x <- c(12,7,3,4.2,18,2,54,-21,8,-5)

# Find Mean.

result.mean <- mean(x)

print(result.mean)

**Applying NA Option**

# Create a vector.

x <- c(12,7,3,4.2,18,2,54,-21,8,-5,NA)

# Find mean.

result.mean <- mean(x)

print(result.mean)

# Find mean dropping NA values.

result.mean <- mean(x,na.rm = TRUE)

print(result.mean)

## Median

# Create the vector.

x <- c(12,7,3,4.2,18,2,54,-21,8,-5)

# Find the median.

median.result <- median(x)

print(median.result)

**Mode**

# Create the function.

getmode <- function(v) {

uniqv <- unique(v)

uniqv[which.max(tabulate(match(v, uniqv)))]

}

# Create the vector with numbers.

v <- c(2,1,2,3,1,2,3,4,1,5,5,3,2,3)

# Calculate the mode using the user function.

result <- getmode(v)

print(result)

# Create the vector with characters.

charv <- c("o","it","the","it","it")

# Calculate the mode using the user function.

result <- getmode(charv)

print(result)

**b) Write R program to find Analysis and Covariance with the sample data and visualize the regression graphically.**

Consider the R built in data set mtcars. In it we observer that the field "am" represents the type of transmission (auto or manual).

It is a categorical variable with values 0 and 1.

The miles per gallon value(mpg) of a car can also depend on it besides the value of horse power("hp").

We study the effect of the value of "am" on the regression between "mpg" and "hp".

It is done by using the aov() function followed by the anova() function to compare the multiple regressions.

**INPUT DATA**

input <- mtcars[,c("am","mpg","hp")]

print(head(input))

We create a regression model taking "hp" as the predictor variable and "mpg"

as the response variable taking into account the interaction between "am" and "hp".

**ANCOVA ANALYSIS**

**MODEL WITH INTERACTION BETWEEN CATEGORICAL VARIABLE AND PREDICTOR VARIABLE**

# Get the dataset.

input <- mtcars

# Create the regression model.

result <- aov(mpg~hp\*am,data = input)

print(summary(result))

**MODEL WITHOUT INTERACTION BETWEEN CATEGORICAL VARIABLE AND PREDICTOR VARIABLE**

# Get the dataset.

input <- mtcars

# Create the regression model.

result <- aov(mpg~hp+am,data = input)

print(summary(result))

**Comparing Two Models**

# Get the dataset.

input <- mtcars

# Create the regression models.

result1 <- aov(mpg~hp\*am,data = input)

result2 <- aov(mpg~hp+am,data = input)

# Compare the two models.

print(anova(result1,result2))