**Ex1**. ARITHMETIC MEAN

a) Write suitable R code to compute the average of the following values. 12,7,3,4.2,18,2,54,-21,8,-5 b) Compute the mean after applying the trim option and removing 3 values from eachend. c)Compute the mean of the following vector . (12,7,3,4.2,18,2,54,-21,8,-5,NA)

#If there are missing values, then the mean function returns NA.

# Find mean dropping NA values. #To drop the missing values from the calculation use na.rm = TRUE

**PROGRAM:**

v1<-c(12,7,3,4.2,18,2,54,-21,8,-5)

s<-(sum(v1))

n<-length(v1)

cat("the average of data is:",s/n, "\n ")

tm<-mean(v1,trim = 0.3)

cat("the trimmed meam of data is:" , tm ,"\n ")

v2<-c(12,7,3,4.2,18,2,54,-21,8,-5,NA)

result.mean <- mean(v2)

cat("without droping NA:",result.mean,"\n")

result.mean <- mean(v2,na.rm = TRUE)

cat("without droping NA:",result.mean)

**Ex2**: MEDIAN Write suitable R code to compute the median of the following values. 12,7,3,4.2,18,2,54,-21,8,-5

**PROGRAM:**

v1<-c(12,7,3,4.2,18,2,54,-21,8,-5)

cat("sorted vector is : \n")

print(sort(v1))

s<-median(v1)

cat("the median of the data is :",s)

**Ex3**: Calculate the mode for the following numeric as well as character data set in R. (2,1,2,3,1,2,3,4,1,5,5,3,2,3) , ("o","it","the","it","it")

**PROGRAM:**

v1<-c(2,1,2,3,1,2,3,4,1,5,5,3,2,3)

v2<-c("o","it","the","it","it")

u1<-unique(v1)

cat("the mode of 1st vec:",u1[which.max(tabulate(match(v1, u1)))],"\n")

u2<-unique(v2)

cat("the mode of 1st vec:",u2[which.max(tabulate(match(v2, u2)))],"\n")

**Ex4**: Download mpg dataset which contains Fuel economy data from 1999 and 2008 for 38 popular models of car from the URL given below.

<https://vincentarelbundock.github.io/Rdatasets/datasets>.

html Answer the following queries

1. Find the car which gives maximum city miles per gallon
2. ii) Find the cars which gives minimum disp in compact and subcompact class

**PROGRAM:**

library(dplyr)

fuel<-read.csv("mpg.csv")

cat("Car which gives maximum city miles per gallon is:",fuel $manufacturer[which.max(fuel $cty)],fuel $model[which.max(fuel $cty)])

df<-filter(fuel,fuel$class=="compact")

df1<-filter(df,df$displ==min(df$displ))

un1<-unique(df1$manufacturer)

cat("\n the cars which gives minimum disp in compact is :",un1)

df2<-filter(fuel,fuel$class=="subcompact")

df3<-filter(df2,df2$displ==min(df2$displ))

un2<-unique(df3$manufacturer)

cat("\n the cars which gives minimum disp in subcompact is :",un2)

**Ex5**: Use the same dataset as used in Exercise 4 and perform the following queries

1. Find the standard deviation of city milles per gallon
2. ii) Find the variance of highway milles per gallon

**PROGRAM:**

fuel<-read.csv("mpg.csv")

s<-sd(fuel $cty)

v<-var(fuel $hwy)

cat("the standard deviation of city of miles per gallon:",s)

cat("\n the standard deviation of highway of miles per gallon:",v)

**EX6**: Use the same dataset and perform the following queries

1. Find the range of the disp in the data set mpg
2. ii) Find the Quartile of the disp in the data set mpg
3. iii) Find the IQR of the disp column in the data set mpg

**PROGRAM:**

fuel<-read.csv("mpg.csv")

cat("Range=",range(fuel$displ))

cat("\nQuartile=",quantile(fuel$displ))

cat("\nInner Quartile Range=",IQR(fuel$displ))

**Ex7**: #Install Library library(e1071)

a. Find the skewness of city miles per mileage in the data set mpg ? Use qplot function and display the graph for the city miles per mileage column.

b. Find the kurtosis of city miles per mileage in the data set mpg Use qplot function and display the graph for the city miles per mileage column.

**PROGRAM:**

df<-read.csv("mpg.csv")

skew<-skewness(df$cty)

print(skew)

qqplot(skew,df$cty)

kur<-kurtosis(df$cty)

print(kur)

qqplot(kur,df$cty)

**Ex8:**

reference status gender testNewOrFollowUp

1 KRXH Accepted Female New

2 KRPT Accepted Male New

3 FHRA Rejected Male New

4 CZKK Accepted Female New

5 CQTN Rejected Female New

6 PZXW Accepted Female Follow-up

7 SZRZ Rejected Male New

8 RMZE Rejected Female New

9 STNX Accepted Female New

10 TMDW Accepted Female New

1. Load the dataset and Create a data frame and name it as dataframe1
2. Load the function for crosstab

**PROGRAM:**

data <- matrix(c("KRXH", "Accepted", "Female", "Test1", "New",

"KRPT", "Accepted", "Male", "Test1", "New",

"FHRA", "Rejected", "Male", "Test2", "New",

"CZKK", "Accepted", "Female", "Test3", "New",

"CQTN", "Rejected", "Female", "Test1", "New",

"PZXW", "Accepted", "Female", "Test4", "Follow-up",

"SZRZ", "Rejected", "Male", "Test4", "New",

"RMZE", "Rejected", "Female", "Test2", "New",

"STNX", "Accepted", "Female", "Test3", "New",

"TMDW", "Accepted", "Female", "Test1", "New"), ncol=5, byrow=TRUE)

dataframe1 <- data.frame(reference=data[,1], status=data[,2], gender=data[,3], testNewOrFollowUp=data[,5])

print(dataframe1)

**Ex9**:

i) Use Two Categorical Variables and Discover the relationships within a dataset

ii) Next, using the xtabs() function, apply two variables from “dataframe1 “, to create a table delineating the relationship between the “Reference” category, and the “Status” category.

1. Save the file in the name of dataframe2.

**PROGRAM:**

data <- matrix(c("KRXH", "Accepted", "Female", "Test1", "New",

"KRPT", "Accepted", "Male", "Test1", "New",

"FHRA", "Rejected", "Male", "Test2", "New",

"CZKK", "Accepted", "Female", "Test3", "New",

"CQTN", "Rejected", "Female", "Test1", "New",

"PZXW", "Accepted", "Female", "Test4", "Follow-up",

"SZRZ", "Rejected", "Male", "Test4", "New",

"RMZE", "Rejected", "Female", "Test2", "New",

"STNX", "Accepted", "Female", "Test3", "New",

"TMDW", "Accepted", "Female", "Test1", "New"), ncol=5, byrow=TRUE)

dataframe1 <- data.frame(reference=data[,1], status=data[,2], gender=data[,3], testNewOrFollowUp=data[,5])

print(dataframe1)

library(stats)

status\_gender\_table <- xtabs(~ status + gender, data = dataframe1)

print(status\_gender\_table)

dataframe2 <- xtabs(~reference+status, data=dataframe1)

print(dataframe2)

table <- xtabs(~reference+status, data=dataframe1)

print(table)

save(dataframe2, file="dataframe2.RData")

**Ex13**: 13. Write a program for creating a pie-chart in R using the input vector(21,62,10,53). Provide labels for the chart as ‘London’, ‘New York’, ‘Singapore’, ‘Mumbai’. Add a title to the chart as ‘city pie-chart’ and add a legend at the top right corner of the chart.

**PROGRAM:**

library(plotrix)

x <- c(21, 62, 10, 53)

labels <- c("London", "New York", "Singapore", "Mumbai")

pie(x,labels,main="City pie chart",col = rainbow(length(x)))

legend("topright", c("Mumbai", "Pune", "Chennai", "Bangalore"),cex = 0.5, fill = rainbow(length(x)))

**Ex14:** . Create a 3D Pie Chart for the dataset “political Knowledge” with suitable labels,colours and a legend at the top right corner of the chart.

**PROGRAM:**

library(plotrix)

x<- c(10, 30, 120, 50, 20)

lbs<-c("YSRCP", "DMK", "TDP", "BJP", "CONGRESS")

colouring<-c("red", "orange", "yellow", "blue", "green")

result <- pie3D(x, main="POLITICAL KNOWLEDGE OF PARTY ELECTED RATE", labels=lbs, col=colouring)

legend("topright", c("YSRCP", "DMK", "TDP", "BJP", "CONGRESS"),cex = 0.5, fill =colouring)

print(result)

**Ex15:** Write a program for creating a bar chart using the vectors H=c(7,12,28,3,41) and M=c(“mar”, “apr”, “may”, “jun”, “jul”). Add a title to the chart as “Revenue chart”.

**PROGRAM:**

H <- c(7,12,28,3,41)

M <- c("Mar","Apr","May","Jun","Jul")

barplot(H,names.arg=M,xlab="Month",ylab="Revenue",col="blue",main="Revenuechart", border="red")

**Ex16**: Make a histogram for the “AirPassengers“dataset, start at 100 on the x-axis, and from values 200 to 700, make the bins 200 wide

**PROGRAM:**

x<-AirPassengers

hist(x, main="Histogram for Air Passengers", xlab="Passengers", border="blue",

col="green",xlim=c(100,700),las=1,w=200)

**Ex17**: Create a Boxplot graph for the relation between "mpg"(miles per galloon) and "cyl"(number of Cylinders) for the dataset "mtcars" available in R Environment.

**PROGRAM:**

x<-mtcars

boxplot(mpg ~ cyl, data =x, xlab = "Number of Cylinders", ylab = "Miles Per Gallon",

main = "Mileage Data")