Practical No: 1

Basics of R

Using R execute basic commands, array, list and frames.

```
x = c(9,20,21,25,30,50)
y = c(3,5,6,7,8,9)
cat("the variable x=",x,"\n")
cat("the variable y=",y,"\n")
cat("addition of x and y is",x+y,"\n")
cat("product of x and y is",x*y,"\n")
cat("mod of x and y is",x%%y,"\n")
cat("division of x and y is",x%/%y,"\n")
cat("Is x greater than y",x<y,"\n")
cat("Is x and y are equal",x==y,"\n")
cat("Is x and y are not equal",x!=y,"\n")
print(x&y)
print(x|y)
print(x&&y)
print(x||y)
z = 4:10
w=5
u=-6
print(w%in%z)
print(u%in%z)
t=seq(0,50,5)
print(t)
```

LIST

```
C = list("this","that","there")
D = list(1,2,3)
#Converting list to vector
print(c(C,D))
```

Matrices

```
B = matrix(c(1,2,3,4,5,6),nrow=2,ncol=3,byrow = FALSE)
A = list("good morning",c(11,22,3),1,B)
print(A[4])
```

Array

```
p=c(1,2,3)
q=c(10,11,12,13,14,15)
column.names = c("col1","col2","col3")
row.names = c("row1","row2","row3")
matrix.names = c("matrix1","matrix2")
R = array(c(p,q),dim = c(3,3,2),dimnames = list(row.names,column.names,matrix.names))
print(R)
```

Print the element the third row of the second matrix

```
print(R[3,,2])
print(R[2,1,])
```

Frames

```
K = data.frame(Name = c("Ravi","sagar","mohit","Reena"),marks = c(20,30,40,50),stream
= c("sci","arts","com","arts"))
print(K)
```

#R code for plotting various plots

Line Plots

```
x=c(7,12,28,3,41)
y=c(14,7,6,19,2)
png(file=”line_chart_2lines.jpg”)
plot(x,type=”o”,col=”red”,xlab=”month”,
ylab=”rainfall”,main=”Rainfall chart”)
lines(y,type=”o”,col=”blue”)
dev.off()
```

Barplot

```
x=c(12,14,15,17,20,34)
png(file="barplot.jpg")
barplot(x,col="blue",names.arg=c("m","t","w","th","F","s"),
xlab="days",ylab="students",main="Attendance")
dev.off()
```

Box plots

```
input = mtcars[,c('mpg','cyl')]
print(head(input))
png(file="boxplot.jpg")
boxplot(mpg~cyl,data=mtcars,xlab="no of cyclilnder",
        ylab="mileage", main='mileage data')
dev.off()
```

Histogram

```
v = c(19, 23, 11, 5, 16, 21, 32,
      14, 19, 27, 39)

# Create the histogram.
hist(v, xlab = "No.of Articles ",
    col = "green", border = "black")
```

Pie charts

```
x = c(34,56,76,66)
y = c("data science","oracle","Rtool","DBA")
```

```
pct = round(x/sum(x)*100)
lbls = paste(y,pct)
lbls=paste(lbls,"%",sep = " ")
pie(x,labels = lbls, col = rainbow(length(lbls)),
main = "subject of computer science")
```

Plots

```
x=c(7,12,28,3,41)
y=c(14,7,6,19,2)
png(file=”line_chart_2lines.jpg”)
plot(x,type=”o”,col=”red”,xlab=”month”,
ylab=”rainfall”,main=”Rainfall chart”)
lines(y,type=”o”,col=”blue”)
dev.off()
```

```
x=c(12,14,15,17,20,34)
png(file="barplot.jpg")
barplot(x,col="blue",names.arg=c("m","t","w","th","F","s"),
xlab="days",ylab="students",main="Attendance")
dev.off()
```

```
x=c(20,50,34,78,90)
png(file="line plot.jpg")
```

```
plot(x,type="o")
dev.off()
```

Box plots

```
input = mtcars[,c('mpg','cyl')]
print(head(input))
png(file="boxplot.jpg")
boxplot(mpg~cyl,data=mtcars,xlab="no of cyclilnder",
        ylab="mileage", main='mileage data')
dev.off()
```

Histogram

```
v = c(19, 23, 11, 5, 16, 21, 32,
      14, 19, 27, 39)

# Create the histogram.
hist(v, xlab = "No.of Articles ",
    col = "green", border = "black")
```

Pie charts

```
x = c(34,56,76,66)
y = c("data science","oracle","Rtool","DBA")
pct = round(x/sum(x)*100)
lbls = paste(y,pct)
lbls=paste(lbls,"%",sep = " ")
pie(x,labels = lbls, col = rainbow(length(lbls)),
main = "subject of computer science")
```

Practical No:2

Matrix Operation

Create matrix using R and perform the operation addition multiplication, transpose and inverse

 # defining matrices

```
names_row = c("R1","R2","R3")
names_col = c("C1","C2")
A = matrix(c(1,2,3,4,5,6),nrow = 3,ncol = 2,byrow = TRUE,dimnames =
list(names_row,names_col))
print(A)
print(A[3,2])
print(A[1,1])
print(A[2,])
```

addition of matrix

```
names_row = c("R1","R2","R3")
names_col = c("C1","C2")
A = matrix(c(1,2,3,4,5,6),nrow = 3,ncol = 2,byrow = TRUE,dimnames =
list(names_row,names_col))
print(A)
B= matrix(c(1,0,2,-4,1,3),nrow = 3,ncol = 2,byrow = TRUE,dimnames =
list(names_row,names_col))
print(B)
print(A+B)
```

Determinant of the matrix

```
A = matrix(c(1,2,3,4,5,6,7,8,9),nrow = 3,ncol = 3,byrow = TRUE)
print(A)
print( det(A))
```

```
print(t(A))
```

R program to multiply two matrices

Multiplilcation of matrix

```
B = matrix(c(1, 2, 3, 4, 5, 6), nrow = 3, ncol = 2)
print(B)

# Creating 2nd Matrix
C = matrix(c(7, 8, 9, 10, 11, 12), nrow = 2, ncol = 3)
print(C)
if (ncol(B) == nrow(C)){
  print(B%*%C)
}else{
  print("multiplication is not possible")
}
```

transpose of matrix

```
A = matrix(c(12,2,3,4,5,6),nrow = 3,ncol = 2,byrow = FALSE)
print (" The given matrix A is ")
print(A)
print(" The transpose of the matrix A is ")
print(t(A))
```

Practical :3

Using R execute the Statistical function

```
v=c(1,2,3,7,3,5,7,-1,2,2,3,2,3,4,5,2)
mean(v)
median(v)
getmode = function(v) {
   uniqv = unique(v)
   uniqv[which.max(tabulate(match(v, uniqv)))]
}
print(getmode(v))
```

```
print(getwd())
```

```
R = max(v) - min(v)
cat(" Range =",R,"\n")
Q.D = (quantile(v)- quantile(v))/2
cat(" Quartile deviation  =",Q.D,"\n")
M = mean(v)
print(M)
M.D = sum(abs(v-M))/length(v)
cat(" mean deviation  =",M.D,"\n")
cat(' the standard deviation =',sd(vX$marks),"\n")
C.V = (sd(v)/M)*100
cat(' the coefficient of variance=',C.V,"\n")
```

Practical NO:4

Using R import the data from Excel/.csv and perform all measure of central tendency

```
data(mtcars)
```

```
head(mtcars,6)
print(summary(mtcars$mpg))
D=quantile(mtcars$cyl,seq(0.25,0.75,0.25))
print(D)
```

```
X= read.csv("filename.csv")
print(head(X,8))

  cat('The mean of the Age",mean(X$Age))

getmode = function(v) {
   uniqv = unique(v)
   uniqv[which.max(tabulate(match(v, uniqv)))]
}
cat('The mode of the countries',getmode(X$Country))
print(summary(X$Age))
```

Practical :5

Using R import the data from Excel/.csv and perform all measure of dispersion

```
print(getwd())
X= read.csv("filename.csv")
R = max(X$Age) - min(X$Age)
cat(" Range =",R,"\n")
Q.D = (quantile(X$Age,0.75)- quantile(X$Age,0.25))/2
cat(" Quartile deviation  =",Q.D,"\n")
M = mean(X$Age)
```

```
print(M)
M.D = sum(abs(X$Age-M))/length(X$Age)
cat(" mean deviation  =",M.D,"\n")
cat(' the standard deviation =',sd(X$marks),"\n")
cat(' the variance =',var(X$marks),"\n")
C.V = (sd(X)/M)*100
cat(' the coefficient of variance=',C.V,"\n")
cat('the covariance of M1 & M2='cov(X$M1,X$M2)
```

Practical No:5

Write R code to find the central and raw Moments for data 1,2,5,6,7,11,10,9,12

```
X = c(1,2,5,6,7,11,10,9,12)
# Raw moments
library(moments)
Y=moment(X, order = 3,central = FALSE)
print(Y)
#Central moment
Z = moment(X, order = 3,central = TRUE)
print(Z)
```

Find the R code to find the skewness using relative and absolute measure of the data 88,95,92,97,96,94,86,91,95,97,88,85,76,68

```
data = c(88, 95, 92, 97, 96, 97, 94, 86, 91, 95,
97, 88, 85, 76, 68)
cat (“ mean of the data”,mean(data))
cat (“ mean of the data”,median(data))
#Absolute measure of skewness
A.S = mean(data)-median(data)
cat(‘The Absolute  measure of skewness’,A.S)
#Relative measure
K.S=3*(mean(data)-median(data))/sd(data)
cat(‘The relative measure of skewness’,K.S)
```

Find the R code to find the skewness of the data using moments
88,95,92,97,96,94,86,91,95,97,88,85,76,68

```
library(moments)
S=skewness(data)
hist(data)
```

Write a R program to find the Kurtosis of the data.

```
data = c(88, 95, 92, 97, 96, 97, 94, 86, 91, 95, 97, 88, 85, 76, 68)
library (moments)
print ("kurtosis of the given data")
print(kurtosis(data))
```

Practical No.6

binomial distribution & Normal Distribution

Suppose there are twelve multiple choice questions in an English class quiz. Each question has five possible answers, and only one of them is correct.Write a R code to find the probability of having four or less correct answers if a student attempts to answer every question at random

```
x  = dbinom(4, size=12, prob=0.2)
sprintf("%s is the probability of attempting 4 question correctly",x)
y = pbinom(4, size=12, prob=0.2)
sprintf("%s is the probability of attempting less than equal to 4 ",y)
```

Write a R code to plot a binomial distribution with random variables taking values from 0 to 50 and probability of the success is 0.5.

```
x = seq(0,50,by = 1)
print(x)
y = dbinom(x,50,0.5)
png(file = "dbinom.png")
plot(x,y)
dev.off()
```

Write R code to find the number of heads that have a probability of 0.25 will come out when a coin is tossed 51 times .

```
x= qbinom(0.25,51,0.5)
sprintf("%s is the number of times head appears in 51 trial if probability is 0.25",x)
```

Write R code to find 8 random variables from a sample of 150 with probability of 0.4.

```
x = rbinom(8,150,0.4)
sprintf("%s  are the random numbers in the sample of 150  with probability 0.4",x)
```

Practical No.6
Normal distribution

Write a R code to plot a normal distribution with random variable taking values between -10 to 10 , mean=2.5 and s.d=2.

```
x = seq(-10, 10, by=0.1)
y = dnorm(x, mean = 2.5, sd = 2)
png(file="pnormExample.png")
plot(x, y)
dev.off()
```

Write R code to create a sequence of number between -10 to 10 incrementing by 0.2 and plot cumulative frequency distribution

```
x = seq(-10, 10, by=0.1)
y = pnorm(x, mean = 2.5, sd = 2)
png(file="pnormExample1.png")
plot(x, y)
dev.off()
```

Write R code to plot graph with function take the probability value and gives a number whose cumulative value matches the probability

```
x= seq(0,1,by =0.02)
```

```
y= qnorm(x,mean=2,sd=1)
png(file= “qnorm.png”)
plot(x,y)
dev.off()
```

Write a R code to generate a random number whose distribution is normal.

```
y= norm(50)
png(file= “rnorm.png”)
hist(y,main= “Normal distribution”)
dev.off()
```

Practical No :7

Fitting a Straight line

Using R fit a straight line on the following data of the number of hours studied and the corresponding exam score for 15 students in some. Find the equation of line and plot the graph

hours=(1, 2, 4, 5, 5, 6, 6, 7, 8, 10, 11, 11, 12, 12, 14),
score=(64, 66, 76, 73, 74, 81, 83, 82, 80, 88, 84, 82, 91, 93, 89)

```
df =data.frame(hours=c(1, 2, 4, 5, 5, 6, 6, 7, 8, 10,
11, 11, 12, 12, 14),
               score=c(64, 66, 76, 73, 74, 81, 83, 82, 80,
88, 84, 82, 91, 93, 89))

#view first six rows of data frame
head(df)
fit = lm(score~hours, data=df)
print(summary(fit))
plot(df$hours, df$score, pch=19,
xlab='hours', ylab='score')
abline(fit)
```

Practical No :8

Fitting a curve

Using R fit a curve on following data of the number of hours studied and the corresponding exam score for 15 students in some. Find the equation of line and plot the graph

hours=(1, 2, 4, 5, 5, 6, 6, 7, 8, 10, 11, 11, 12, 12, 14),
score=(64, 66, 76, 73, 74, 81, 83, 82, 80, 88, 84, 82, 91, 93, 89)

```
df =data.frame(hours=c(1, 2, 4, 5, 5, 6, 6, 7, 8, 10, 11, 11, 12, 12, 14),
               score=c(64, 66, 76, 73, 74, 81, 83, 82, 80, 88, 84, 82, 91, 93, 89))

#view first six rows of data frame
head(df)
fit = lm(score~hours+I(hours^2), data=df)
summary(fit)
plot(df$x, df$y, pch=19, xlab='hours',
ylab='score',col=’blue’)
lines(x_axis, predict(fit,
data.frame(hours=x_axis)), col='green')
```

Practical No :9

Hypothesis testing

#Write R Code for One sample test

#H0:mu=5

#H1:mu≠5

#$\alpha = 0.05$

 # Defining sample vector

```
x =rnorm(100)

# One Sample T-Test
t.test(x, mu = 5)
```

#write Rcode for the two sample test

#H0:mu1=mu2

#H1:mu1≠mu2

Defining sample vector

```
x =rnorm(100)
y =rnorm(100)

# Two Sample T-Test
t.test(x, y)

t.test(x, mu = 2, alternative = 'greater')
```

Practical 10

Chi Square test

Write R Code to perform Chi-square test to check the dependency between smoking habits and exercise practice.

```
library(MASS)
print(str(survey))
#Ho: Smoking habits and exercise habits are independent
#H1: there is some relationship between smoking habits and exercise
# Create a data frame from the main data set.
stu_data = data.frame(survey$Smoke,survey$Exer)

# Create a contingency table with the needed variables.
stu_data = table(survey$Smoke,survey$Exer)

print(stu_data)
print(chisq.test(stu_data))
```