## Practical 1

#1.Create Vector

fruits<-c("apple","banana","mango","pineapple"); print(fruits)

#class of vector print(class(fruits)) num<-c(1,4,5,7,8,9);

print(num) #class of vetor print(class(num))

#2.List

list1<-list(c(1:3),fruits) print(list1)

#3.Matrices M=matrix(c(1:9),nrow=3,byrow=TRUE) print(M) K=matrix(c(1:12),nrow=4,ncol=3,byrow=TRUE) print(K)

#4.Array

A<-array(c("yes","no","true","false"),dim=c(4)) print(A)

B<-array(c("yes","no","true","false"),dim=c(3,4)) print(B)

C<-array(c("yes","no","true","false"),dim=c(3,4,2)) print(C)

#5.Factor

apple\_colors<-c("red","blue","green") factor\_apple<-factor(apple\_colors) #convert in sequence order print(factor\_apple) print(nlevels(factor\_apple))

#6.Data frames my<-data.frame(

name=c("veer","nilesh","pooja","habat"),

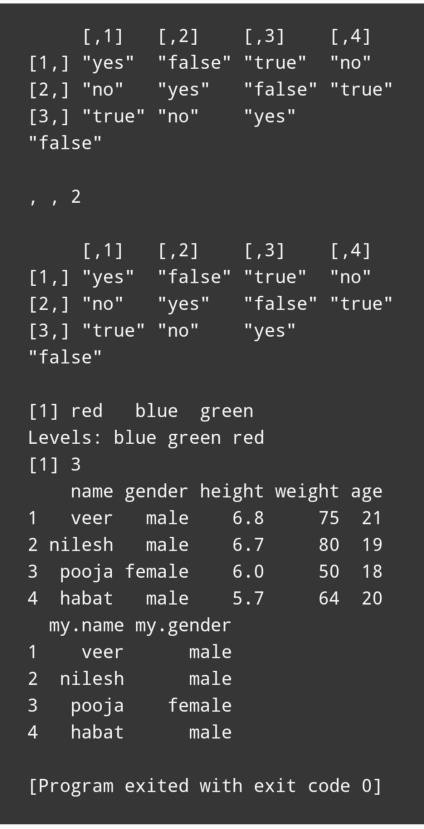
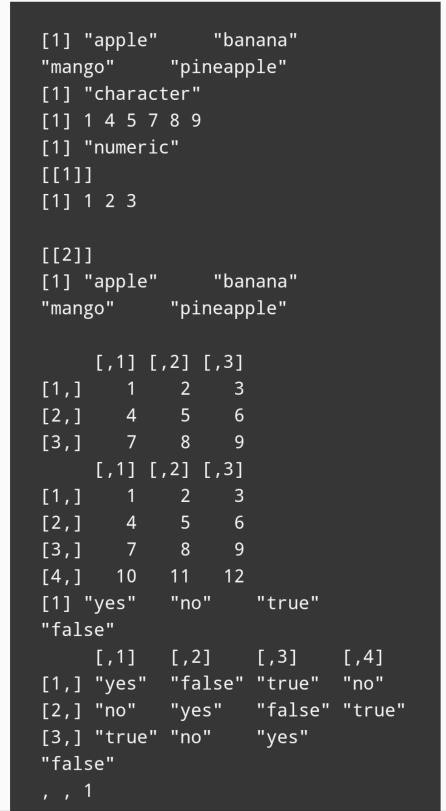
gender=c("male","male","female","male"), height=c(6.8,6.7,6,5.7),

weight=c(75,80,50,64), age=c(21,19,18,20)

)

print(my) data.frame(my$name,my$gender)

## Output:-



**Practical 3**

#1.find mean of 12,7,3,4,2,1,18,2,54,-21,8,-5

x<-c(12,7,3,4,2,1,18,2,54,-21,8,-5)

result.mean<-mean(x)

print(result.mean)

#2.find median of 12,7,3,4,2,1,18,2,54,-21,8,-5

x<-c(12,7,3,4,2,1,18,2,54,-21,8,-5)

result.median<-median(x) print(result.median)

#3.find mode of 2,1,2,3,12,3,4,1,5,5,3,2,2

getmode=function(v)

{

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

}

v<-c(2,1,2,3,12,3,4,1,5,5,3,2,2)

result.mode<-getmode(v) print(result.mode)

#find range of the eruption duration in the data set faithful duration=faithful$eruption

min(duration) max(duration)

result.range=max(duration)-min(duration) print(result.range)

#find standard deviation of the eruption duration in the data set faithful duration=faithful$eruption

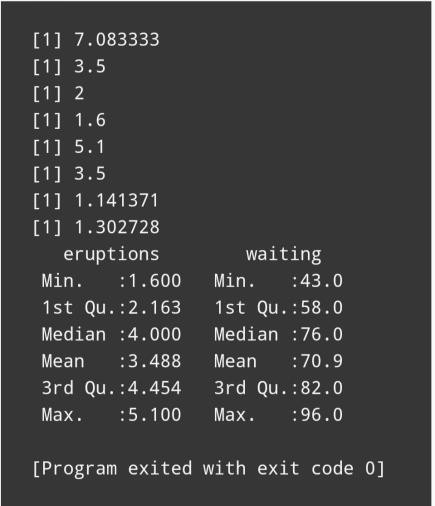
result.sd=sd(duration) print(result.sd)

#find variance of the eruption duration in the data set faithful duration=faithful$eruption

result.var=var(duration) print(result.var)

#find summary of the data set faithful result.summary=summary(faithful) print(result.summary)

## Output:-



**Practical No: 4**

salarydata<-data.frame( salaries\_low=c(25000,26000,27000,28000,29000,30000,31000), salaries\_high=c(25999,26999,27999,28999,29999,30999,32999), Numbers=c(8,10,16,14,10,5,2)

)

print(salarydata)

#salarydata<-salarydata[,cumNumbers:=cumsum(Numbers)] salarydata

mediangroup<-salarydata

{

((cumNumbers-Numbers)<=(max(cumNumbers)/2 & (cumNumbers)>=(max(cumNumbers)/2))

{

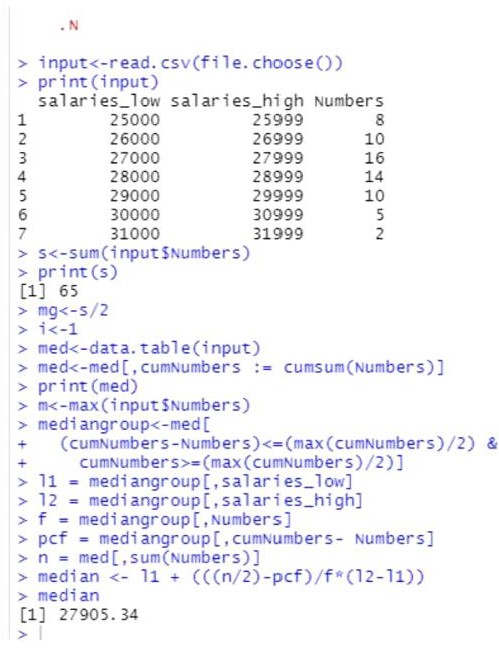
l1=mediangroup(,salaries\_low) l2=mediangroup(,salaries\_high) f=mediangroup(,Numbers) pcf=mediangroup(,cumNumbers-Numbers) n=salarydata(,sum(Numbers))

}

}

median<-l1+(((n/2)-pcf)/f\*(l2-l1)) median

## Output:-



**Practical No: 5**

salarydata<-data.frame(

salaries\_low=c(25000,26000,27000,28000,29000,30000,31000), salaries\_high=c(25999,26999,27999,28999,29999,30999,32999), Numbers=c(8,10,16,14,10,5,2)

)

print(salarydata)

salarydata[ , prevNumbers := shift(Numbers,1)] salarydata[ , nextNumbers := shift(Numbers,-1)] salarydata

#identifying mode group

modegroup <- salarydata[Numbers == max(Numbers)] modegroup

#creating the variables needed to calculate mode l1 = modegroup[,salaries\_low]

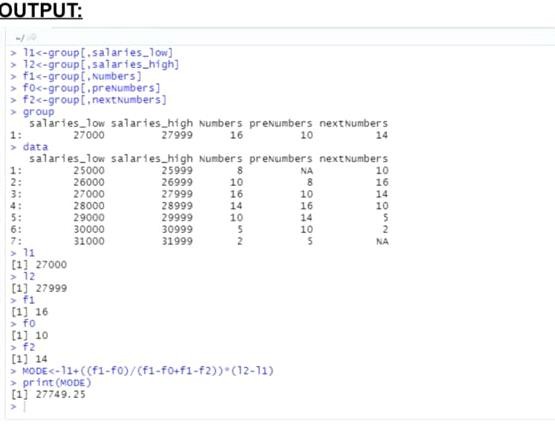
l2 = modegroup[,salaries\_high] f1 = modegroup[,Numbers]

f0 = modegroup[,prevNumbers] f2 = modegroup[,nextNumbers]

#calculating mode

groupmode <- l1 + ((f1-f0)/(f1-f0+f1-f2)\*(l2-l1)) groupmode

## Output:-



**Practical No: 6**

# a) 12, 6, 7, 3, 15, 10, 18, 5

# b) 9, 3, 8, 8, 9, 8, 9, 18.

input<-data.frame(

X1=c(12, 6, 7, 3, 15, 10, 18, 5),

X2=c(9, 3, 8, 8, 9, 8, 9, 18)

)

print("Standard Deviaton of X1 and X2 are : ")

sd(input$X1) sd(input$X2)

print("Variance of X1 and X2 are : ") var(input$X1)

var(input$X2)

# OUTPUT:-

## Practical No: 7

input<-data.frame( X=c(0,1,2,3,4),

F1=c(10,5,2,2,1),

F2=c(1,2,14,2,1),

F3=c(1,2,5,5,10)

)

getmode <- function(v) { uniqv <- unique(v)

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

}

mean1<-mean(rep(input$X,input$F1)) median1<-median(rep(input$X,input$F1)) mode1<-getmode(rep(input$X,input$F1)) sd1<-sd(rep(input$X,input$F1)) Pearsonfirst<-((mean1-mode1)/sd1) Pearsonsecond<-(3\*(mean1-median1)/sd1)

print("Pearson's first coefficients of skewness with x and f1 is : ") print(Pearsonfirst)

print("Pearson's second coefficients of skewness with x and f1 is : ") print(Pearsonsecond)

mean1<-mean(rep(input$X,input$F2)) median1<-median(rep(input$X,input$F2))

mode1<-getmode(rep(input$X,input$F2)) sd1<-sd(rep(input$X,input$F2)) Pearsonfirst<-((mean1-mode1)/sd1) Pearsonsecond<-(3\*(mean1-median1)/sd1)

print("Pearson's first coefficients of skewness with x and f2 is : ") print(Pearsonfirst)

print("Pearson's second coefficients of skewness with x and f2 is : ") print(Pearsonsecond)

mean1<-mean(rep(input$X,input$F3)) median1<-median(rep(input$X,input$F3)) mode1<-getmode(rep(input$X,input$F3)) sd1<-sd(rep(input$X,input$F3)) Pearsonfirst<-((mean1-mode1)/sd1) Pearsonsecond<-(3\*(mean1-median1)/sd1)

print("Pearson's first coefficients of skewness with x and f3 is : ") print(Pearsonfirst)

print("Pearson's second coefficients of skewness with x and f3 is : ") print(Pearsonsecond)

# OUTPUT:-

## Practical No: 8

Me<-data.frame(

Spades=402, Diamonds=358,Clubs=273, Hearts=467

)

attach(Me)

result<-chisq.test(Me) print(result) if(result$p.value>=0.05){

print("Null Hypothesis is Accepted")

}else{

print("Null Hypothesis is Rejected")

}

# OUTPUT:-

## Practical No: 9

input<-data.frame( Finance = c(12,7), Sales = c(38,19),

HR= c(5,3),Technology= c(8,1)) attach(input)

result<-chisq.test(input) print(result) if(result$p.value>=0.05){

print("Null Hypothesis is Accepted")

}else{

print("Null Hypothesis is Rejected")

}

# OUTPUT:-

## Practical No: 10

# (a) What percentage play in fewer than 750 games? # (b) What percentage play in more than 2000 games?

# (c) Find the 90th percentile for the number of games played during a career.

print("What percentage play in fewer than 750 games") pa<-pnorm(750, mean = 1500, sd = 350)

Percenta <- pa\*100 print(Percenta)

print("What percentage play in more than 2000 games")

pb<-pnorm(2000, mean = 1500, sd = 350, lower.tail = FALSE)

Percentb <- pb\*100 print(Percentb)

print("the 90th percentile for the number of games played during a career") p05<-round(qnorm(0.05,mean = 1500, sd = 350),0)

p95<-round(qnorm(0.95,mean = 1500, sd = 350),0) cat("Range for 90 Percentile is : ",p05,"-",p95)

## Output:-

**Practical No: 11**

prac<-data.frame( H=c(70,63,72,60,66,70,74,65,62,67,65,68), W=c(155,150,180,135,156,168,178,160,132,145,139,152)

)

print("a) H as the independent variable") reg<-lm(W ~ H,data = prac)

print(reg)

print("b) H as the dependent variable")

reg<-lm(H ~ W,data = prac) print(reg)

# OUTPUT:-

## Practical No: 12

a) Graph the data and show the least-squares regression line. # (b) Find and plot the trend line for the data.

# (c) Estimate the value of total agricultural exports in the # year 2006.

input<-data.frame( YEAR=c(2000,2001,2002,2003,2004,2005), TOTAL=c(51246,53659,53115,59364,61383,62958)

)

reg<-lm(TOTAL ~ YEAR,data = input) print(reg)

plot(input$YEAR,input$TOTAL,type = "p", col = "blue", pch = 16, cex = 1.3,xlab = "Year",ylab

= "Total Value",main = "total agricultural exports") abline(reg,col = "red")

print("Estimate the value of total agricultural exports in the year 2006.") newdata = data.frame(YEAR=2006)

predict(reg, newdata)

**OUTPUT:-**

