Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 343

Practical No :1

Aim : Using R Execute the basic commands.

```
Console   Terminal ×   Background Jobs ×
R 4.3.0 · ~/
> #Kunal M. Mistri- 3432
> #Practical 1 - Basic Commands
> #Assignment Operator
> x=5
> x
[1] 5
> #Concatenate Function -
> y=c(4,5,6,8)
> y
[1] 4 5 6 8
> #Scan Function
> s=scan()
1: 2
2: 3
3: 4
4: 5
5: 6
6:
Read 5 items
> s
[1] 2 3 4 5 6
> #Assigning Method
> z=1:10
> z
 [1]  1  2  3  4  5  6  7  8  9 10
> #Sequence
> a=seq(1,20,5)
> a
[1]  1  6 11 16
> m=c(25,23,21,14,18,19,15,16,12,10)
> m
 [1] 25 23 21 14 18 19 15 16 12 10
> max(m)
[1] 25
> min(m)
```

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 343

Console Terminal Background Jobs

R 4.3.0 · ~/

```
> min(m)
[1] 10
> length(m)
[1] 10
> sum(m)
[1] 173
> prod(m)
[1] 1.665075e+12
> sqrt(m)
 [1] 5.000000 4.795832 4.582576 3.741657 4.242641 4.358899
 [7] 3.872983 4.000000 3.464102 3.162278
> sqrt(sum(m))
[1] 13.15295
> mean(m)
[1] 17.3
> median(m)
[1] 17
> mode(m)
[1] "numeric"
> sd(m)
[1] 4.808557
> var(m)
[1] 23.12222
> #cumulative freqeuncy
> cumsum(m)
 [1]  25  48  69  83 101 120 135 151 163 173
> #sorting in Ascending Order
> sort(m)
 [1] 10 12 14 15 16 18 19 21 23 25
> #sorting in descending Order
> sort(m,decreasing = T)
 [1] 25 23 21 19 18 16 15 14 12 10
> summary(m)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  10.00   14.25   17.00   17.30   20.50   25.00 
```

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 343

Q2 Find the max, min , length, product, sum, square root , mean , median, SD, Variance , sort , summary

Console Terminal Background Jobs

R 4.3.0 · ~/

```
> w=scan()
1: 32
2: 42
3: 45
4: 25
5: 56
6: 43
7: 65
8: 34
9: 43
10: 45
11:
Read 10 items
> w
 [1] 32 42 45 25 56 43 65 34 43 45
> w
 [1] 32 42 45 25 56 43 65 34 43 45
> min(w)
[1] 25
> max(w)
[1] 65
> length(w)
[1] 10
> prod(w)
[1] 1.556975e+16
> sum(w)
[1] 430
> sqrt(w)
 [1] 5.656854 6.480741 6.708204 5.000000 7.483315 6.557439 8.062258 5.830952
 [9] 6.557439 6.708204
> sqrt(sum(w))
[1] 20.73644
> mean(w)
[1] 43
> median(w)
```

```
> median(w)
[1] 43
> mode(w)
[1] "numeric"
> sd(w)
[1] 11.48913
> var(w)
[1] 132
> #cumulative freqeuncy
> cumsum(w)
 [1]  32  74 119 144 200 243 308 342 385 430
> #sorting in Ascending Order
> sort(w)
 [1] 25 32 34 42 43 43 45 45 56 65
> #sorting in descending Order
> sort(w,decreasing = T)
 [1] 65 56 45 45 43 43 42 34 32 25
> summary(w)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
     25      36      43      43      45      65 
>
```

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 343

R Console

```
> #Data Structure In R
> num_vec=c(3,4,6,9)
> num_vec
[1] 3 4 6 9
> char_vec=c("Apple","Mango","Orange")
> char_vec
[1] "Apple"  "Mango"  "Orange"
> #To Create Matrix
> mat=matrix(1:9,nrow=3,ncol=3)
> mat
     [,1] [,2] [,3]
[1,]    1    4    7
[2,]    2    5    8
[3,]    3    6    9
> mat1=matrix(c(12,23,53,14,36,25,78,64,45),nrow=3,ncol=3)
> mat1
     [,1] [,2] [,3]
[1,]   12   14   78
[2,]   23   36   64
[3,]   53   25   45
> mat2=matrix(c(12,23,53,14,36,25),nrow=2,ncol=3)
> mat2
     [,1] [,2] [,3]
[1,]   12   53   36
[2,]   23   14   25
> mat3=matrix(c(12,23,53,14,36,25),nrow=3,ncol=2)
> mat3
     [,1] [,2]
[1,]   12   14
[2,]   23   36
[3,]   53   25
```

```
> #TO Create List
> my_list=list(4,"Mango",TRUE,14+5i)
> my_list
[[1]]
[1] 4

[[2]]
[1] "Mango"

[[3]]
[1] TRUE

[[4]]
[1] 14+5i

> #To Create Data Frame
> df=data.frame(c(1,2,3,4),c("apple","mango"),c(TURE,FALSE,TRUE,FALSE))
Error: object 'TURE' not found
> df=data.frame(c(1,2,3,4),c("apple","mango"),c(TURE,FALSE,TURE,FALSE))
Error: object 'TURE' not found
> df=data.frame(c(1,2,3,4),c("apple","mango"),c(TRUE,FALSE,TRUE,FALSE))
> df
  c.1..2..3..4. c..apple....mango.. c.TRUE..FALSE..TRUE..FALSE.
1             1               apple                        TRUE
2             2               mango                       FALSE
3             3               apple                        TRUE
4             4               mango                       FALSE
```

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 343

Q3 .Using following data construct a matrix of order 3x3 and perform element wise multiplication, matrix multiplication, transpose of matrix, diagonal of matrix, inverse of matrix, and rank of matrix

R Console

```
> b=c(3,5,7,3,4,2,7,3,9)
> b
[1] 3 5 7 3 4 2 7 3 9
> B=matrix(data=b,nrow=3,ncol=3)
> B
     [,1] [,2] [,3]
[1,]    3    3    7
[2,]    5    4    3
[3,]    7    2    9
> #Element wise Multiplication
> B*B
     [,1] [,2] [,3]
[1,]    9    9   49
[2,]   25   16    9
[3,]   49    4   81
> #matrix Multiplication
> B%*%B
     [,1] [,2] [,3]
[1,]   73   35   93
[2,]   56   37   74
[3,]   94   47  136
> #inverse matrix
> solve(B)
           [,1]       [,2]        [,3]
[1,] -0.2777778  0.1203704  0.17592593
[2,]  0.2222222  0.2037037 -0.24074074
[3,]  0.1666667 -0.1388889  0.02777778
> #joining two Matrix together
> cbind(B,B)
     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    3    3    7    3    3    7
[2,]    5    4    3    5    4    3
[3,]    7    2    9    7    2    9
> rbind(B,B)
     [,1] [,2] [,3]
[1,]    3    3    7
[2,]    5    4    3
[3,]    7    2    9
[4,]    3    3    7
[5,]    5    4    3
[6,]    7    2    9
> #finding Diagonal element
> diag(B)
[1] 3 4 9
> #Transpose of Matrix
> t(B)
     [,1] [,2] [,3]
[1,]    3    5    7
[2,]    3    4    2
[3,]    7    3    9
> #rank of Matrix
> qr(B)$rank
[1] 3
> 
```

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

PRACTICAL NO : 2

Aim : Import the data from excel/.csv to find mean, mode, median, standard deviation, variance, length, sum, sort, range, square root, summary, quantile

Q.1. Import the data from excel/.csv
to find mean,median,mode,SD,variance,length,sum,sort,range,sqrt,summary,quantile.

A1 | fx | x

	A	B	C	D
1	x			
2	12			
3	13			
4	17			
5	22			
6	25			
7	12			
8	10			
9	5			
10	17			
11	12			
12				

```
> #Kunal Mistri
> #3432
> data=read.csv(file.choose())
> data
    x
1  12
2  13
3  17
4  22
5  25
6  12
7  10
8   5
9  17
10 12
```

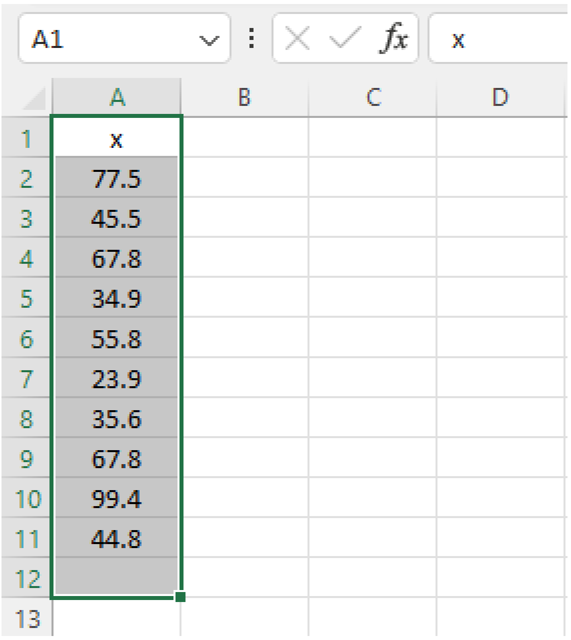
Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

```
> y=data$x
> y
 [1] 12 13 17 22 25 12 10  5 17 12
> length(y)
[1] 10
> sum(y)
[1] 145
> mean(y)
[1] 14.5
> median(y)
[1] 12.5
> max(y)
[1] 25
> min(y)
[1] 5
> mode(y)
[1] "numeric"
> sd(y)
[1] 5.87367
> sort(y)
 [1]  5 10 12 12 12 13 17 17 22 25

> var(y)
[1] 34.5
> sort(y,decreasing=TRUE)
 [1] 25 22 17 17 13 12 12 12 10  5
> sqrt(y)
 [1] 3.464102 3.605551 4.123106 4.690416 5.000000 3.464102 3.162278 2.236068 4.123106 3.464102
> range(y)
[1]  5 25
> summary(y)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
    5.0    12.0    12.5    14.5    17.0    25.0
> q1=quantile(y,0.25)
> q1
25%
 12
> q2=quantile(y,0.50)
> q2
 50%
12.5
> q3=quantile(y,0.75)
> q3
75%
 17
```

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

Q.2. Import the data from excel/.csv
to find mean,median,mode,SD,variance,length,sum,sort,range,sqrt,summary,quantile.



A1 | fx | x

	A	B	C	D
1	x			
2	77.5			
3	45.5			
4	67.8			
5	34.9			
6	55.8			
7	23.9			
8	35.6			
9	67.8			
10	99.4			
11	44.8			
12				
13				

```
>
> data=read.csv(file.choose())
> data
      x
1  77.5
2  45.5
3  67.8
4  34.9
5  55.8
6  23.9
7  35.6
8  67.8
9  99.4
10 44.8
>
```

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

```
> y=data$x
> y
 [1] 77.5 45.5 67.8 34.9 55.8 23.9 35.6 67.8 99.4 44.8
> length(y)
[1] 10
> sum(y)
[1] 553
> mean(y)
[1] 55.3
> median(y)
[1] 50.65
> mode(y)
[1] "numeric"
> min(y)
[1] 23.9
> max(y)
[1] 99.4
> sd(y)
[1] 22.96592
> var(y)
[1] 527.4333
> sort(y)
 [1] 23.9 34.9 35.6 44.8 45.5 55.8 67.8 67.8 77.5 99.4
> sort(y,decreasing=TRUE)
 [1] 99.4 77.5 67.8 67.8 55.8 45.5 44.8 35.6 34.9 23.9
> range(y)
[1] 23.9 99.4
> sqrt(y)
 [1] 8.803408 6.745369 8.234076 5.907622 7.469940 4.888763 5.966574 8.234076 9.969955 6.693280
> summary(y)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  23.90   37.90   50.65   55.30   67.80   99.40 
> q1=quantile(y,0.25)
> q1
 25% 
37.9 
> q2=quantile(y,0.50)
> q2
  50% 
50.65 
```

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

Q.3. Import the data from excel/.csv
to find mean,median,mode,SD,variance,length,sum,sort,range,sqrt,summary,quantile.

A1 sr.no

	A	B	C	D	E
1	sr.no	Marks	Height		
2	1	20	143		
3	2	12	148		
4	3	32	141.5		
5	4	27	155		
6	5	40	159		
7	6	49	153		
8	7	19	157		
9	8	27	149		
10	9	38	152.5		
11	10	32	165		
12					
13					

```
>
> data=read.csv(file.choose())
> data
   sr.no Marks Height
1      1    20  143.0
2      2    12  148.0
3      3    32  141.5
4      4    27  155.0
5      5    40  159.0
6      6    49  153.0
7      7    19  157.0
8      8    27  149.0
9      9    38  152.5
10    10    32  165.0
>
> m=data$Marks
> m
 [1] 20 12 32 27 40 49 19 27 38 32
> h=data$Height
> h
 [1] 143.0 148.0 141.5 155.0 159.0 153.0 157.0 149.0 152.5 165.0
```

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

```
> m;h
 [1] 20 12 32 27 40 49 19 27 38 32
 [1] 143.0 148.0 141.5 155.0 159.0 153.0 157.0 149.0 152.5 165.0
> length(m);length(h)
[1] 10
[1] 10
> sum(m);sum(h)
[1] 296
[1] 1523
> mean(m);mean(h)
[1] 29.6
[1] 152.3
> median(m);median(h)
[1] 29.5
[1] 152.75
> mode(m);mode(h)
[1] "numeric"
[1] "numeric"
> sd(m);sd(h)
[1] 11.02724
[1] 7.215724
> var(m);var(h)
[1] 121.6
[1] 52.06667
> min(m);min(h)
[1] 12
[1] 141.5
> max(m);max(h)
[1] 49
[1] 165
> range(m);range(h)
[1] 12 49
[1] 141.5 165.0
> sort(m);sort(h)
 [1] 12 19 20 27 27 32 32 38 40 49
 [1] 141.5 143.0 148.0 149.0 152.5 153.0 155.0 157.0 159.0 165.0
> sort(m,decreasing=TRUE);sort(h,decreasing=TRUE)
 [1] 49 40 38 32 32 27 27 20 19 12
 [1] 165.0 159.0 157.0 155.0 153.0 152.5 149.0 148.0 143.0 141.5
> sqrt(m);sqrt(h)
 [1] 4.472136 3.464102 5.656854 5.196152 6.324555 7.000000 4.358899 5.196152 6.164414 5.656854
 [1] 11.95826 12.16553 11.89538 12.44990 12.60952 12.36932 12.52996 12.20656 12.34909 12.84523
```

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

Q.4. Given the following data about avg rainfall in every month in the year of 2017
calculate arithmetic mean,mode,median,first quantile,56th percentile,3rd decile for above data.

A1 | month

	A	B	C	D	E
1	month	Rainfall(in mm)			
2	jan	10			
3	feb	10			
4	mar	10			
5	april	10			
6	may	10			
7	june	560			
8	july	640			
9	aug	520			
10	sept	320			
11	oct	90			
12	nov	20			
13	dec	10			
14					
15					

```
>
> data=read.csv(file.choose())
> data
   month Rainfall.in.mm.
1    jan              10
2    feb              10
3    mar              10
4  april              10
5    may              10
6   june             560
7   july             640
8    aug             520
9   sept             320
10   oct              90
11   nov              20
12   dec              10
>
> R=data$Rainfall
> R
 [1]  10  10  10  10  10 560 640 520 320  90  20  10
```

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

```
> mean(R)
[1] 184.1667
> median(R)
[1] 15
> mode=function(x){
+ uniqx=unique(x)
+ uniqx[which.max(tabulate(match(x,uniqx)))]
+ }
> mode(R)
[1] 10
> quantile(R,.25)
25%
 10
> quantile(R,.56)
 56%
31.2
> quantile(R,.3)
30%
 10
```

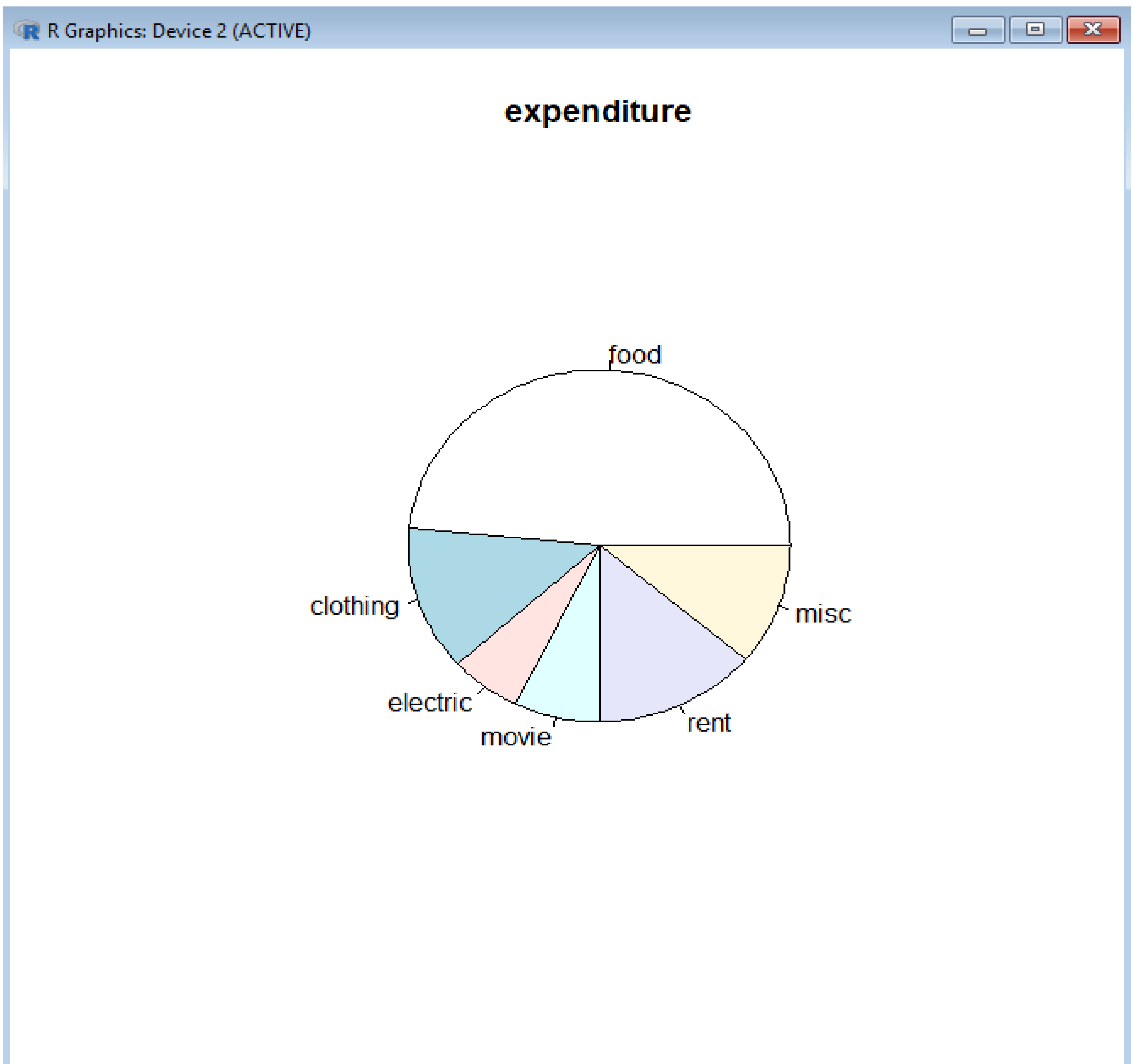
Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

PRACTICAL NO : 3

Aim : Perform R program for making Diagrams

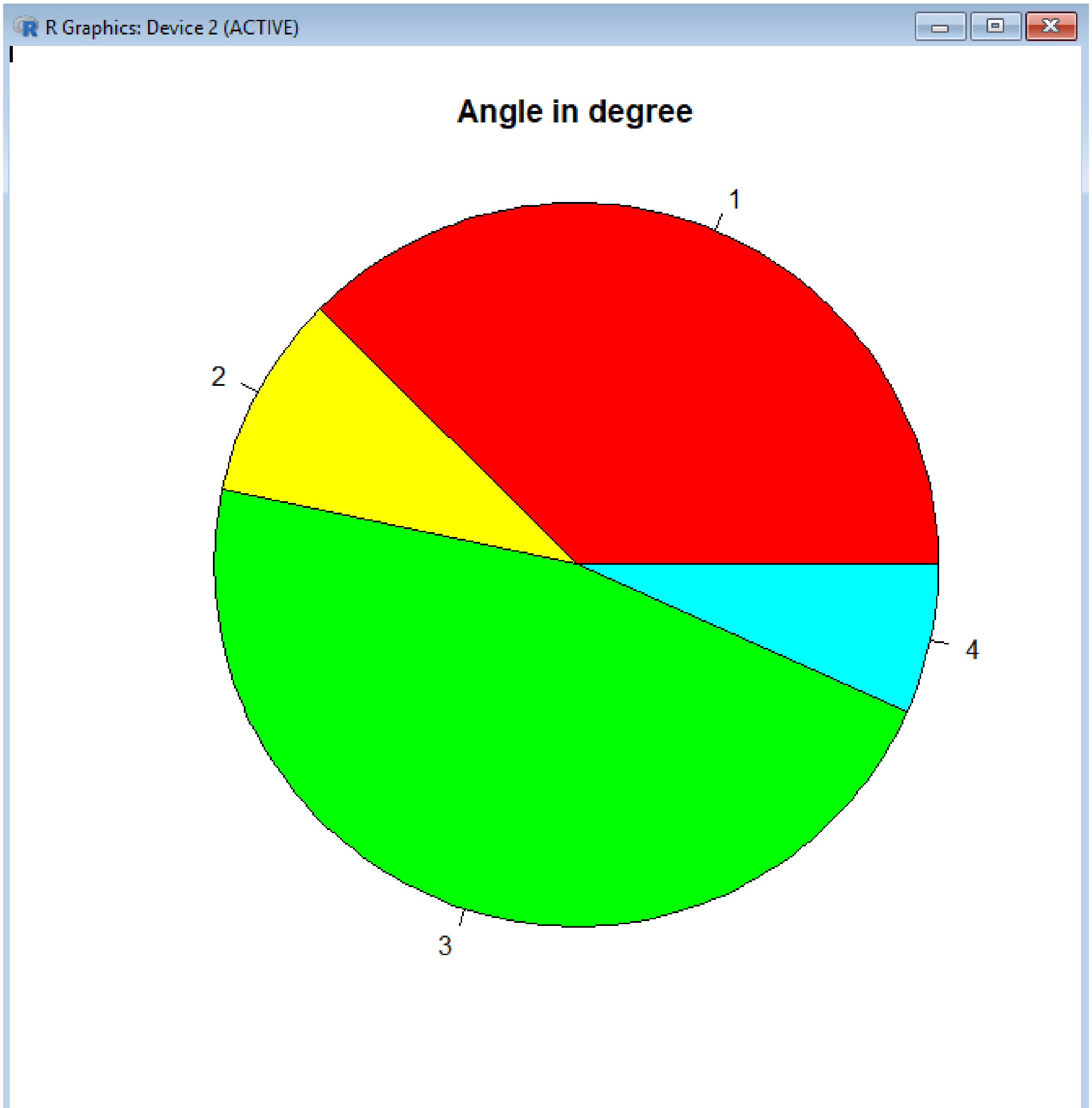
Q. Represent the following data by a pie diagram.

```
> #Kunal Mistri
> #3432
> itm=c("food","clothing","electric","movie","rent","misc")
> exp=c(87,24,11,13,25,20)
> pie(exp,main="expenditure",labels=itm,radius=1,color=rainbow(length(exp)))
```



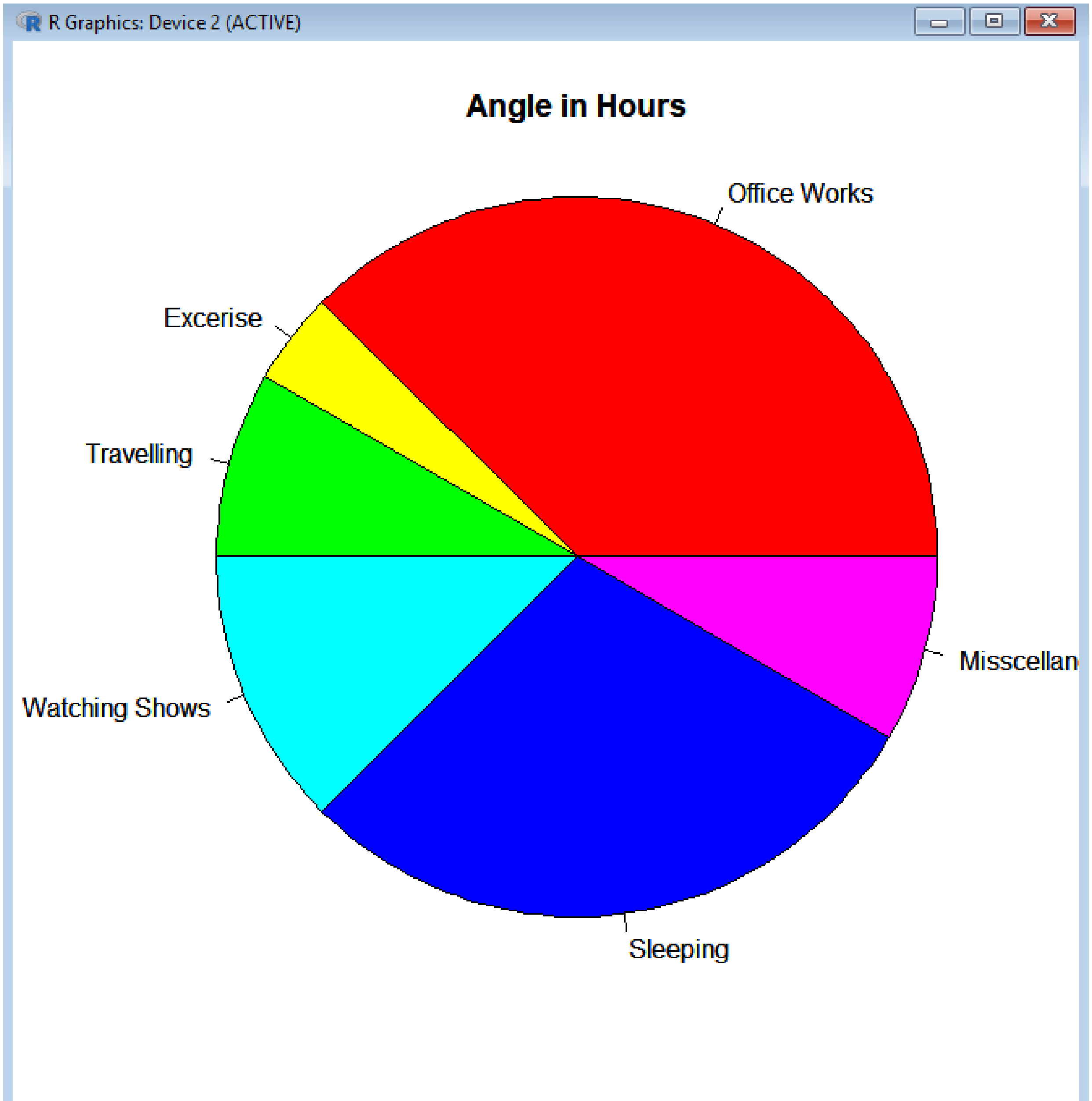
Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

```
> src = c("Govt Canals","Tanks","Wells and Tubewells","others")
> milhect = c(135,33,168,24)
>
> pie (milhect, main = "Angle in degree",lablels = src, radius = 1, col=rainbow(length(exp)))
```



Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

```
> Activity = c("Office Works","Excerise","Travelling","Watching Shows","Sleeping","Misscellaneous")
> Noofhrs = c(9,1,2,3,7,2)
> pie (Noofhrs, main = "Angle in Hours",labels = Activity, radius = 1, col=rainbow(length(Noofhrs)))
```



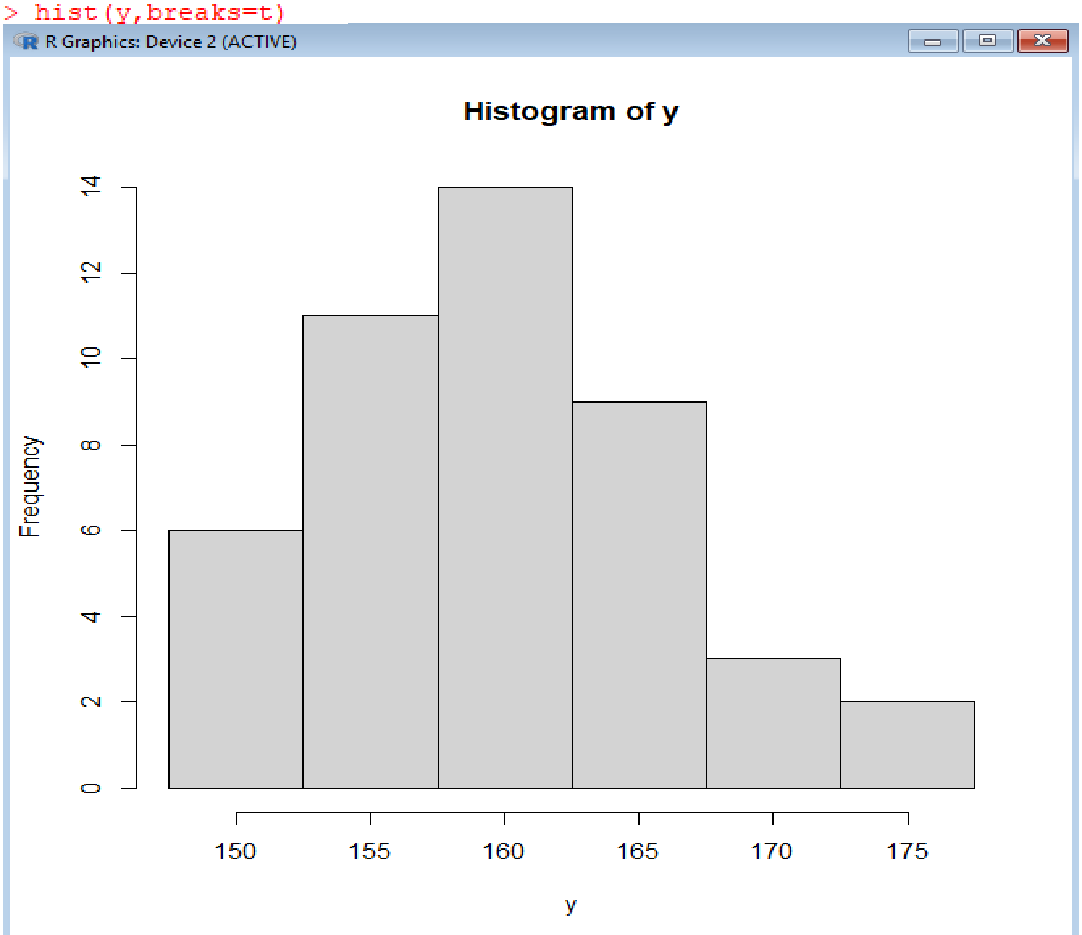
Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

PRACTICAL NO : 4

Aim : Perform R program for making Graphs

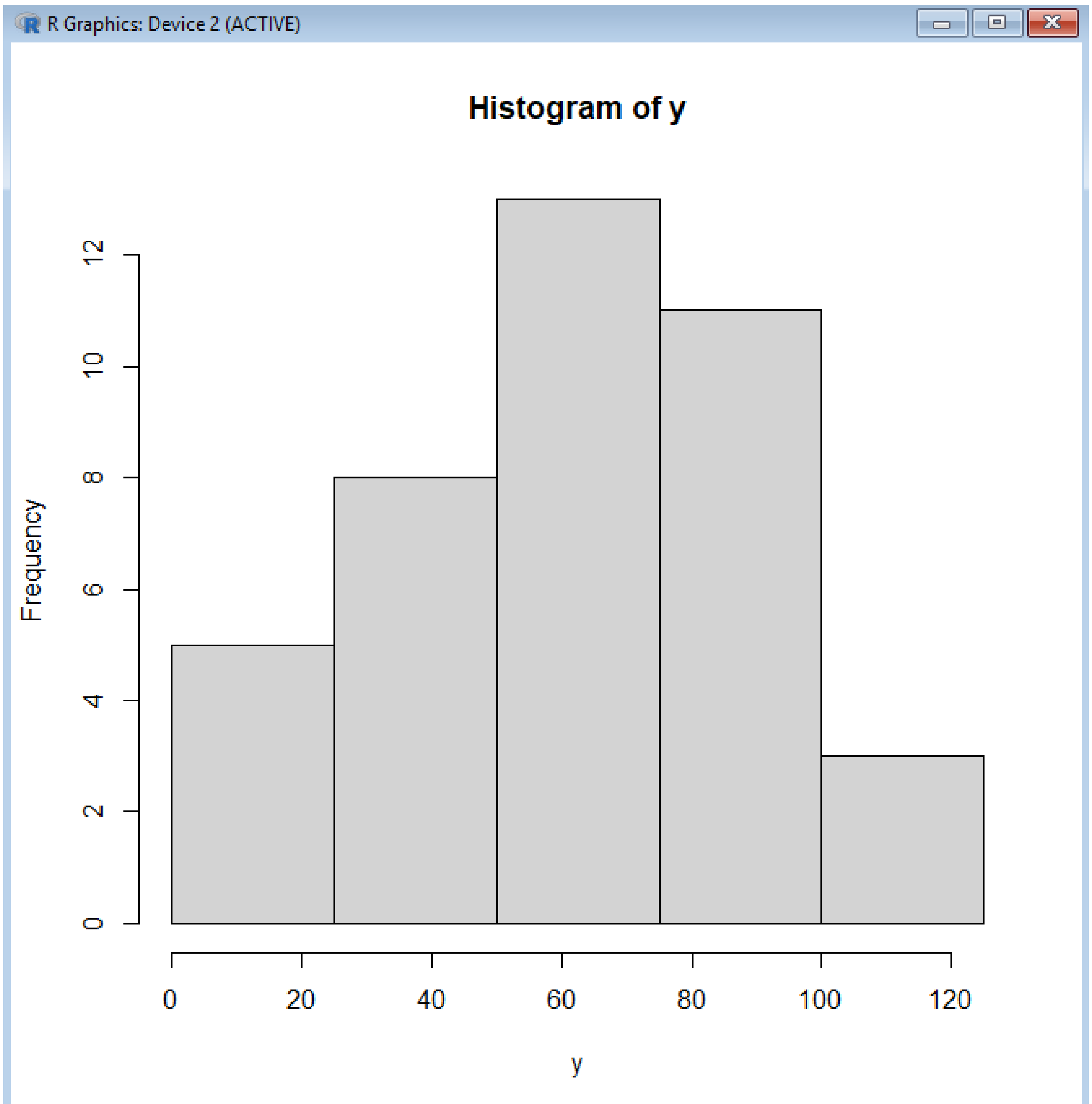
Q. Represent the following data by a histogram, frequency polygon and Ogive.

```
> #Kunal Mistri
> #3432
> x<-seq(150,175,5)
> f<-c(6,11,14,9,3,2)
> y<-rep(x,f)
> hist(y)
> t=seq(147.5,177.5,5)
> hist(y,breaks=t)
```



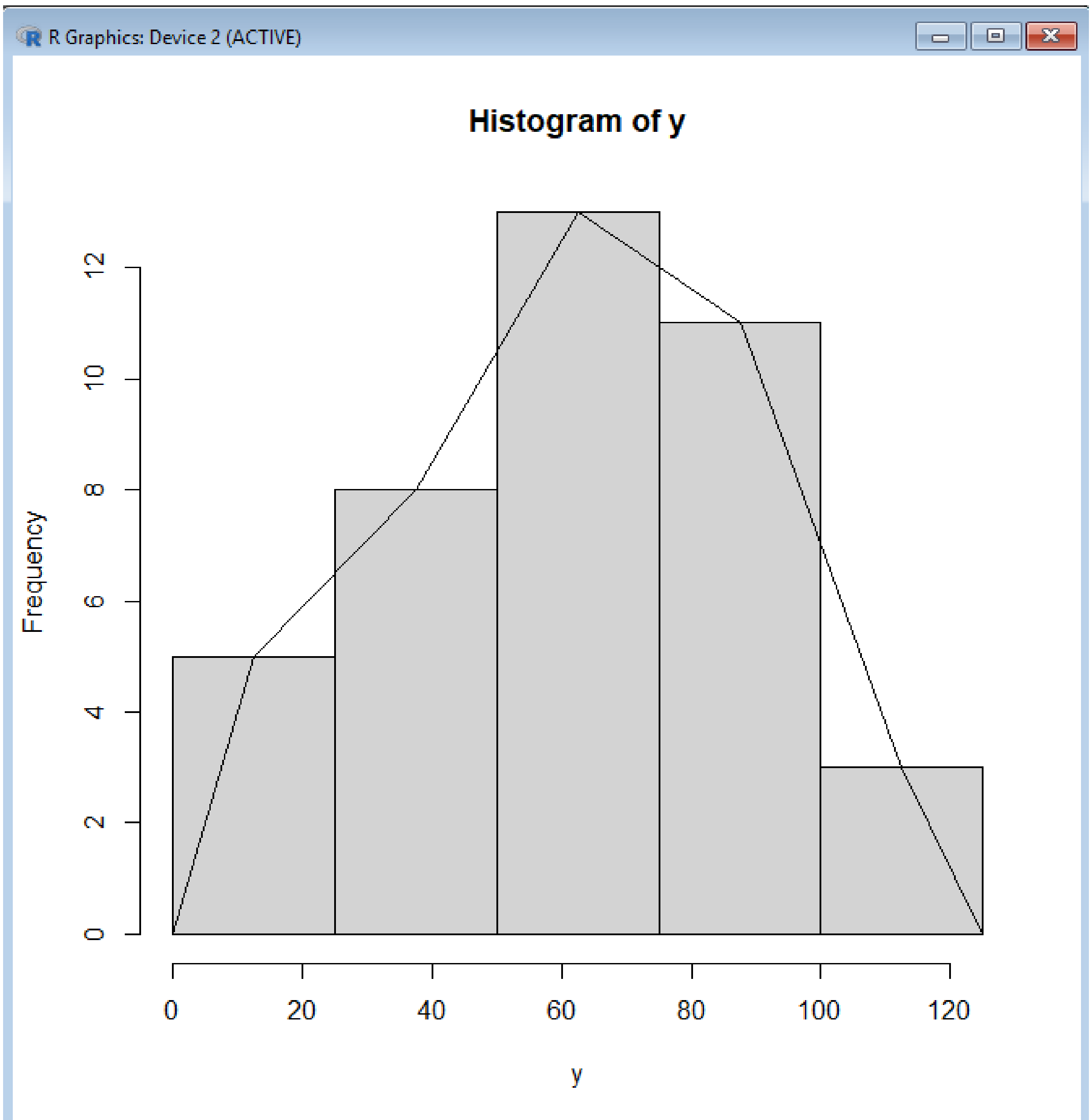
Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

```
> mindx= seq(12.5,112.5,25)
> f=c(5,8,13,11,3)
> cls_limit=seq(0,125,25)
> y=rep(mindx,f)
> hist(y)
> hist(y,breaks = cls_limit)
```



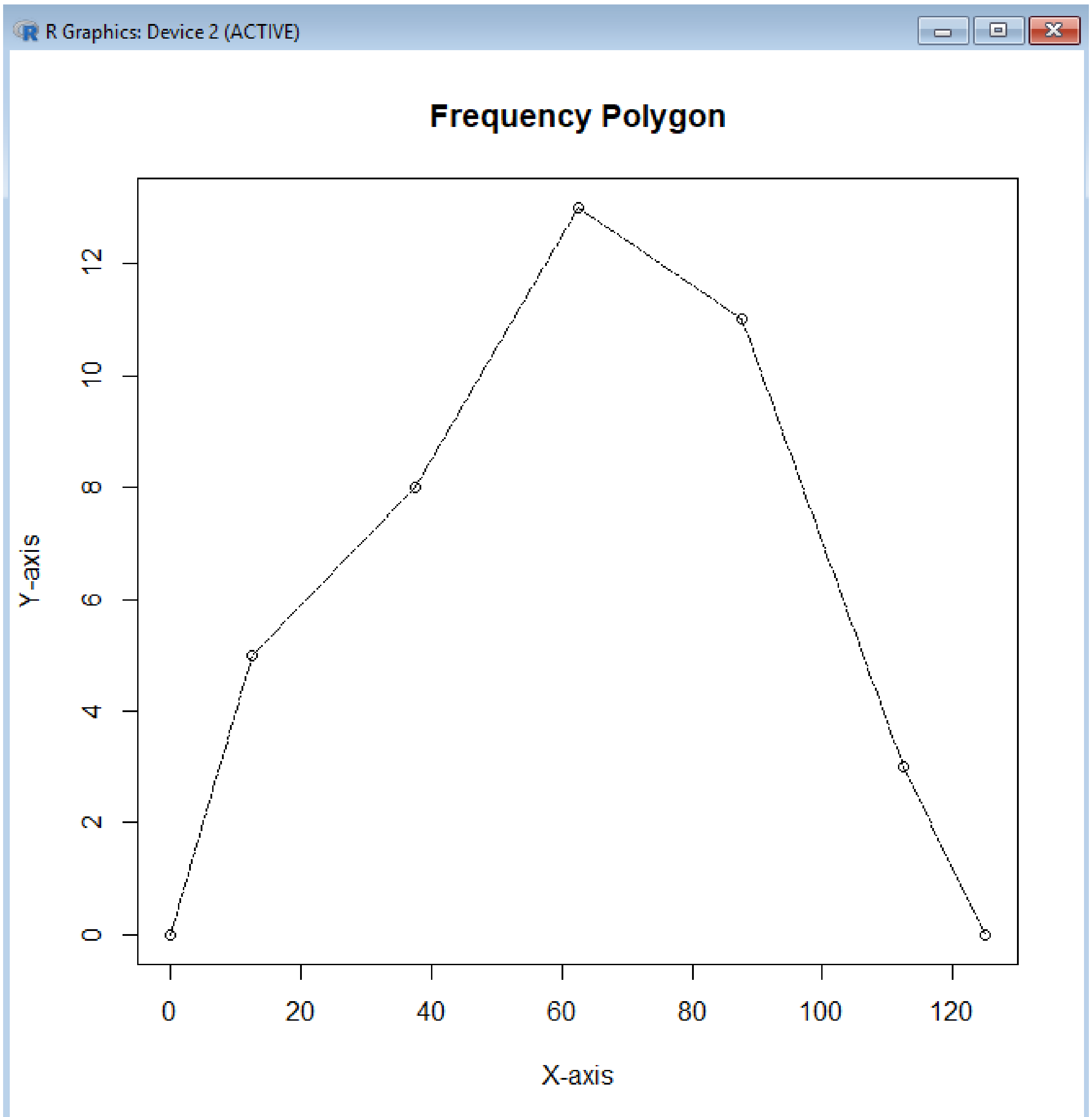
Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

```
> lb=seq(0,100,25)
> ub=seq(25,125,25)
> mindx=(lb+ub)/2
> f=c(5,8,13,11,3)
> x0=c(0,mindx,125)
> f0=c(0,f,0)
> y=rep(mindx,f)
> bks=seq(0,125,25)
> hist(y,breaks = bks)
> lines(x0,f0)
```



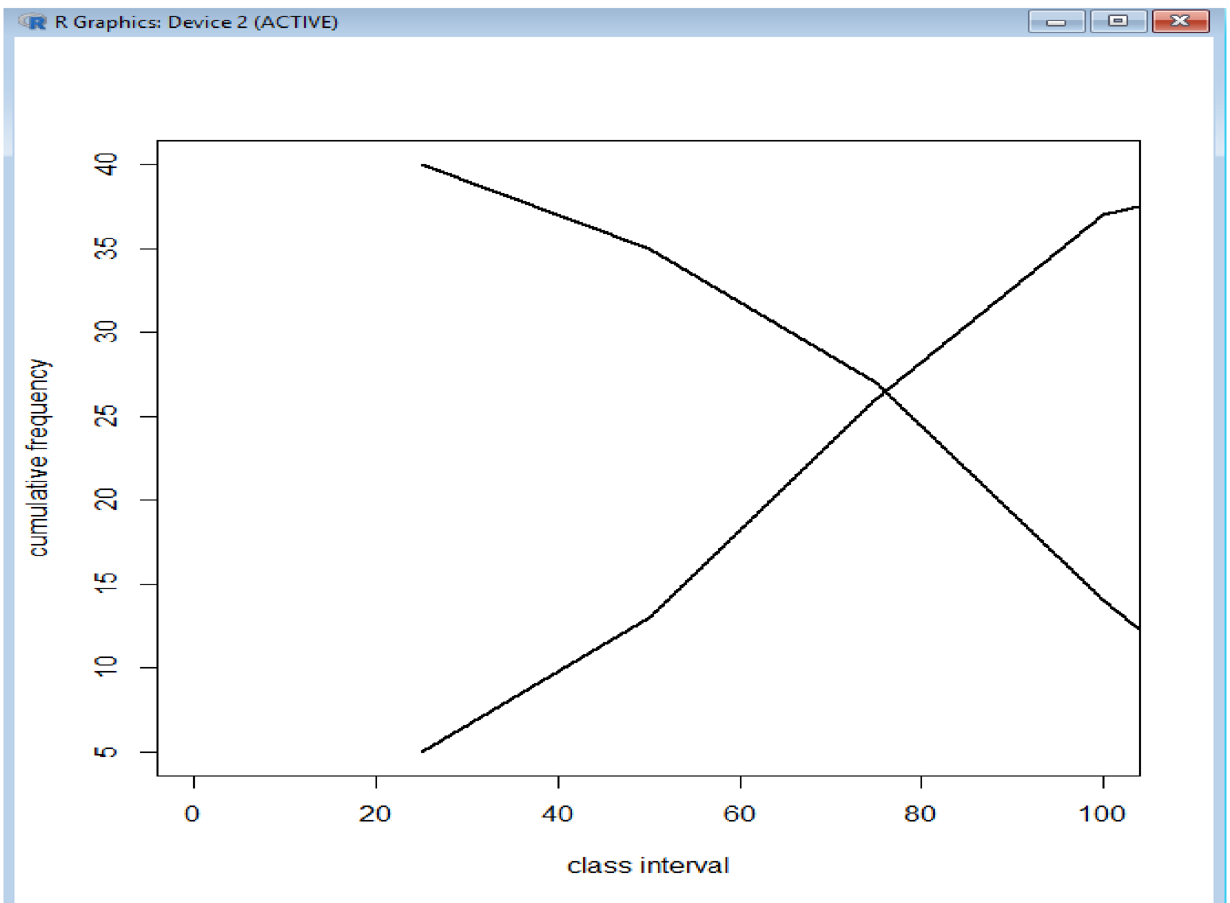
Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

```
>
> lb=seq(0,100,25)
> ub=seq(25,125,25)
> mindx=(lb+ub)/2
> f=c(5,8,13,11,3)
> x0=c(0,mindx,125)
> f0=c(0,f,0)
> y=rep(mindx,f)
> bks=seq(0,125,25)
> hist(y,breaks = bks)
> lines(x0,f0)
> plot(x0,f0,main="Frequency Polygon",xlab="X-axis",ylab="Y-axis",type="o",lty=6,xlim=range(min(x0),max(x0)))
```



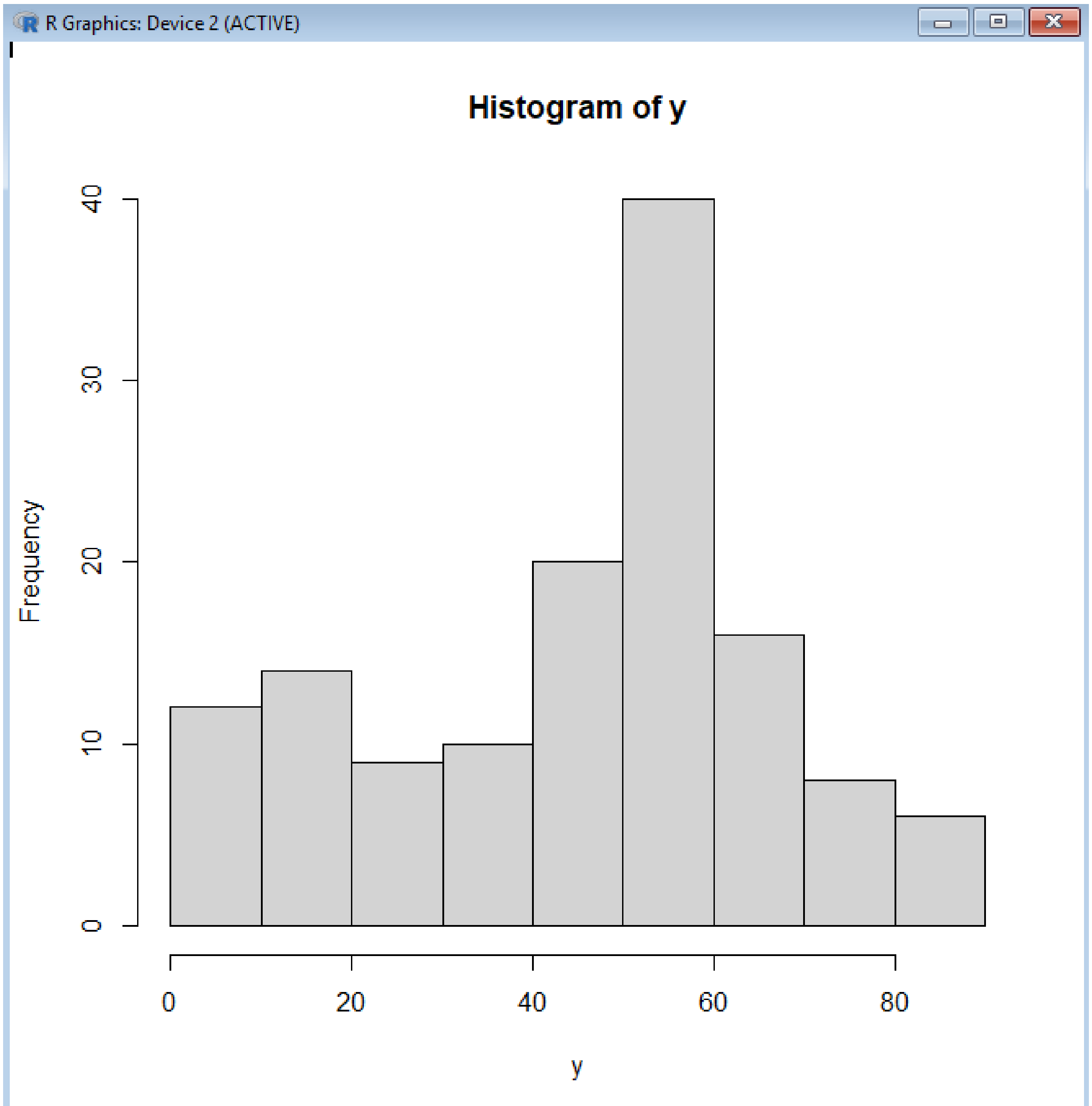
Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

```
>
> f-c(5,8,13,11,3)
[1] 0 0 0 0 0
> lc=cumsum(f)
> lc
[1]  5 13 26 37 40
> uc=1:5
> uc
[1] 1 2 3 4 5
> for(i in 5:1)
+ {uc[i]=sum(f[5:i])}
> uc
[1] 40 35 27 14  3
> lbx=seq(0,100,25)
> lbx
[1]   0  25  50  75 100
> lbx=seq(25,125,25)
> ubx=seq(25,125,25)
> ubx
[1]  25  50  75 100 125
> plot(ubx,lc,type="1",xlim=c(0,100),xlab="class interval",ylab="cumulative frequency",lwd=2)
Error in plot.xy(xy, type, ...) : invalid plot type '1'
> plot(ubx,lc,type="l",xlim=c(0,100),xlab="class interval",ylab="cumulative frequency",lwd=2)
> lines(lbx,uc,type="l",xlim=c(0,100),xlab="class interval",ylab="cumulative frequency",lwd=2)
```



Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

```
> midx=seq(5,85,10)
> f=c(12,14,9,10,20,40,16,8,6)
> cls_limit=seq(0,90,10)
> y=rep(midx,f)
> hist(y)
> hist(y,breaks = cls_limit)
```



Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

Practical No 5

Q1. A DIE IS TURNED UP132 TIMES WITH THE FOLLOWING RESULTSNO OF TURNED UP: 1 2 3 4 5 6 FREQUENCY:16 20 25 14 29 28 WITH EXPECTED FREQUENCIES 1/6 FOR EACH NO IS THE DIE UNBIASED

```
> #Kunal Mistri
> #3432
```

```
> x = c(16,20,25,14,29,28)
> ch.test=chisq.test(x,p=c(1/6,1/6,1/6,1/6,1/6,1/6))
> ch.test

        Chi-squared test for given probabilities

data:  x
X-squared = 9, df = 5, p-value = 0.1091
```

#since p value >0.05 we accept H0

Q2. We collected wild tulips and found that 81 were red, 50 were yellow and27 were white..Suppose that, in the region where you collected the data, the ratio of red, yellow and white tulip is 3:2:1 (3+2+1 = 6). This means that the expected proportion is: 3/6 (= 1/2) for red,2/6 (= 1/3) for yellow, 1/6 for white. Test the hypothesis using chi square test of goodnessof fit.

```
> #Q2
> x =c(81,50,27)
> ch.test=chisq.test(x,p=c(1/2,1/3,1/6))
> ch.test

        Chi-squared test for given probabilities

data:  x
X-squared = 0.20253, df = 2, p-value = 0.9037
```

#The p-value of the test is 0.9037, which is greater than the significance level alpha=0.05. We can conclude that the observed proportions are not significantly different from the expected proportions.

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

CHI-SQUARE TEST FOR INDEPENDENCE OF ATTRIBUTE

Q1.Use the following data to test whether the attributes-condition at home(A)and condition of child(B)are independent

A B	CLEAN	DIRTY
CLEAN	70	50
FAIRLY CLEAN	80	20
DIRTY	35	45

```
> #H0:the condition at home and condition of child are independent
> #H1:the condition at home and condition of child are not independent
> #l.o.s=5%
```

```
> #Q3
> x=c(70,80,35,50,20,45)
> y = matrix(x,nrow = 3,ncol=2)
> y
     [,1] [,2]
[1,]   70   50
[2,]   80   20
[3,]   35   45
> chisq=chisq.test(y,correct = F)
> chisq

        Pearson's Chi-squared test

data:  y
X-squared = 25.646, df = 2, p-value = 2.698e-06
```

#pvalue<0.05 reject HO

Q2.TWO RESEARCHERS ADOPTED DIFFERENT SAMPLING TECHNIQUES WHILE INVESTIGATING THE SAME GROUP OF STUDENTSTO FIND THE NUMBER OF STUDENTS FALLING IN DIFFERENT INTELLIGENCE LEVEL.THE RESULTS ARE AS FOLLOWS

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

Q2.TWO RESEARCHERS ADOPTED DIFFERENT SAMPLING TECHNIQUES WHILE INVESTIGATING THE SAME GROUP OF STUDENTSTO FIND THE NUMBER OF STUDENTS FALLING IN DIFFERENT INTELLIGENCE LEVEL.THE RESULTS ARE AS FOLLOWS

Researchers	Intelligence level			
	Below avg	Avg	Above avg	Genius
X	86	60	44	10
Y	40	33	25	5

```
> #Case2
> x=c(86,40,60,33,44,25,10,5)
> y=matrix(x,nrow=2,ncol=4)
> y
     [,1] [,2] [,3] [,4]
[1,]   86   60   44   10
[2,]   40   33   25    5
> chisq=chisq.test(y,correct=F)
> chisq

        Pearson's Chi-squared test

data:  y
X-squared = 0.5327, df = 3, p-value = 0.9117
```

>#pvalue>0.05 accept HO

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

PRACTICAL-6

AIM: Perform an R program on z-test- One population mean, Two population means, One population proportion, Two population proportion.

CASE 1: One population mean

Q1. Test the hypothesis H0:M=10 against H1:M!=10. A random sample of size 400 is drawn and the given mean is 10.2 and Standard Deviation 2.25. Use 5% Level of Significance.

```
> #Kunal Mistri
> #3432

>
> #H0:m=10 vs H1:M1=10, LOS=0.05
> n=400
> m=10.2
> s=2.25
> M=10
> zcal=(m-M)/(s/sqrt(n))
> zcal
[1] 1.777778
> pvalue=2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.07544036
>
```

Q2. A jumbo orange is supposed to average 4 ounces in weight. For quality control, a sample of 100 oranges is taken and it is found that the average weight is 4.03 ounces and the sample standard deviation is 0.1 ounces. Use 5% Level of significance.

```
>
> #H0:M=4 vs H1:M!=4, LOS=0.05
> n=100
> m=4.03
> s=0.1
> M=4
> zcal=(m-M)/(s/sqrt(n))
> zcal
[1] 3
> pvalue=2*(1-pnorm(abs(acal)))
Error: object 'acal' not found
> pvalue=2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.002699796
```

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

CASE 2: Two population mean

Q1. Two Random samples of size 1000 and 2000 are drawn from two populations with the same Standard Deviation 2.5 gives means 67.5 and 68 respectively test the hypothesis. Use 5% Level of Significance.

```
> #H0:M1=M2 vs H1:M!=M2, LOS=0.05
> N1=1000
> N2=2000
> m1=67.5
> m2=68
> sd1=2.5
> sd2=2.5
> zcal=(m1-m2)/sqrt(sd1^2/N1+s2^2/N2)
Error: object 's2' not found
> zcal=(m1-m2)/sqrt(sd1^2/N1+sd2^2/N2)
> zcal
[1] -5.163978
> pvalue=2*(1=pnorm(abs(zcal)))
Error in 1 = pnorm(abs(zcal)) :
  invalid (do_set) left-hand side to assignment
> pvalue=2*(1-pnorm(abs(zcal)))
> pvalue
[1] 2.417564e-07
> #pvalue>0.05, we reject H0
```

Q2. The amount of a certain trace element in blood is known to vary with a standard deviation of 14.1 ppm (parts per million) for male blood donors and 9.5 ppm for female donors. Random samples of 75 male and 50 female donors yield concentration means of 128 and 33 ppm, respectively. What is the likelihood that the population means of concentration of the element are the same for men and women?

```
> # H0:M1=M2  H1:M1!=M2, LOS=5%
> N1=75
> N2=50
> M1=28
> M2=33
> SD1=14.1
> SD2=9.5
> zcal=(M1-M2)/sqrt(SD1^2/N1+SD2^2/N2)
> zcal
[1] -2.368684
> pvalue=2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.01785149
> # pvalue<0.05, reject H0
```

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

CASE 3: One population proportion

Q1. Experience has shown that 20% of a manufactured product is of the top quality. In one day's production of 400 articles only 50 are of top quality. Test the hypothesis that experience of 20% is wrong. Use 5% Level of Significance.

```
> #H0:P=0.2 vs H1:P!=0.2, LOS=0.05
> P=0.2
> Q=1-P
> Q
[1] 0.8
> n=400
> p=50/400
> p
[1] 0.125
> zcal=(p-P)/sqrt(P*Q/n)
> zcal
[1] -3.75
> pvalue=2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.0001768346
> #pvalue<0.05, we reject H0
```

Q2. A survey claims that 9 out of 10 doctors recommend aspirin for their patients with headaches. To test this claim, a random sample of 100 doctors are obtained. Out of these 100 doctors, 82 indicate that they recommend aspirin. Is this claim accurate? Use 5% Level of Significance.

```
> #H0:P=0.9 vs H1:P!=0.9, LOS=0.05
> P=0.9
> Q=1-P
> Q
[1] 0.1
> n=100
> p=82/100
> zcal=(p-P)/sqrt(P*Q/n)
> zcal
[1] -2.666667
> pvalue=2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.007660761
> #pvalue<0.05, we reject H0
```

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

CASE 4: Two population proportion

Q1. From each of two consignments of apples a sample of size 200 is drawn and the number of rotten counted test whether the proportion of rotten apples in the two consignments are different for - sample A size of 200 and number of rotten apples are 44 and sample B size of 200 and number of rotten apples are 30. Use 5% Level of significance.

```
> n1=200
> p1=44/200
> n2=200
> p2=30/200
> p=(n1*p1+n2*p2)/(n1+n2)
> q=1-p
> zcal=(p1-p2)/sqrt(p*q*(1/n1+1/n2))
> zcal
[1] 1.802741
> pvalue=2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.07142888
> #pvalue>0.05, we accept H0
```

Q2. A drug research experimental unit is testing two drugs newly developed to reduce blood pressure levels. The drugs are administered to two different sets of animals.
Drug Test Result -

	Animals Tested	Animals Responded
Drug1	600	350
Drug2	500	260

The research unit wants to test whether there is a difference between the efficiency of the said drug at 5% Level of Significance.

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

```
> #H0:P1=P2 vs P1!=P2, LOS=5%
> n1=600
> n2=500
> p1=350/600
> p2=260/500
> p=(n1*p1+n2*p2)/(n1+n2)
> q=1-p
> zcal=(p1-p2)/sqrt(p*q*(1/n1+1/n2))
> zcal
[1] 2.104391
> pvalue=2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.03534439
> #pvalue<0.05, we reject H0
```

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

Practical no 7

Case 1

Q1) The following data refers to the amount of coffee(in ounces) filled by a machine in six randomly picked jars: 15.7,15.9,16.3,16.2,15.7 and 15.9. Is the true mean amount of coffee in a jars is 16 ounces ? use los =5%

Output

```
> #Kunal Mistri
> #3432
```

```
x = c(15.7,15.9,16.3,16.2,15.7,15.9)
t.test(x,mu=16)
```

```
>  t.test(x,mu=16)

        One Sample t-test

data:  x
t = -0.48795, df = 5, p-value = 0.6462
alternative hypothesis: true mean is not equal to 16
95 percent confidence interval:
 15.68659 16.21341
sample estimates:
mean of x
    15.95
```

#p-value > los, Accept H0.

Hence we conclude that the true mean amount of coffee in a jar is 16 ounces.

Q2) Raju restaurant near the railway station at Falna has been having average sales of 500 tea cups per day. Because of the development of a bus stand nearby , it expects to increase its sales . During the first 12 days after the start of the bus stand , the daily sales were as under :550,570,490,615,505,580,570,460,600,580,530,526. On the basis of this information , can one conclude that Raju's restaurtant's sales have increased ?

```
>  sales = c(550,570,490,615,505,580,570,460,600,580,530,526)
>  t.test(sales,mu=500,alternative="greater")

        One Sample t-test

data:  sales
t = 3.5614, df = 11, p-value = 0.002231
alternative hypothesis: true mean is greater than 500
95 percent confidence interval:
 523.7954      Inf
sample estimates:
mean of x
      548

>  # Hence we conclude Raju's restaurtant's sales have increased
```

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

Q3) An IQ test was administered to 5 persons before and after they were trained . The results are as follows :

Candidate: 1 2 3 4 5

Before: 110 120 123 132 125

After:120 118 125 136 121

Test whether there is any change in IQ after the training Program at 1% Los

```
#Q3
# H0:M1=M2 Vs H1:M1!=M2, los=0.01
Before=c(110,120,123,132,125)
After=c(120,118,125,136,121)
t.test(Before,After,paired=T)

        Paired t-test

ata:  Before and After
 = -0.8165, df = 4, p-value = 0.4601
lternative hypothesis: true mean difference is not equal to 0
5 percent confidence interval:
-8.800874  4.800874
ample estimates:
ean difference
            -2

# p value > los; Accept H0
# Hence we conclude that there is no change in IQ after the training
#program at 1% los .
```

Q4) The sales data of a product in six shops before and after a special promotional campaign are as under:

Shops : A B C D E F

Before : 53 28 31 48 50 42

After : 58 29 30 55 56 45

Can the significance be judges to be success at 5% level of significance

```
> #Q4
> #H0:M1=M2 Vs H1:M1<M2,los=0.05
> Before=c(53,28,31,48,50,42)
> After=c(58,29,30,55,56,45)
> t.test(Before,After,paired=T,alternative="less")

        Paired t-test

data:  Before and After
t = -2.7815, df = 5, p-value = 0.01942
alternative hypothesis: true mean difference is less than 0
95 percent confidence interval:
       -Inf -0.9644531
sample estimates:
mean difference
           -3.5

> #p value < los; Reject H0
> # Hence we conclude that the campaign can be judged to be success
 at
> #5% level of significance .
```

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

Q5) A group of seven week old chickens reared on a high protein diet weigh 12, 15,11,16,14,14 and 16 ounces ; a second group of five chickens similarly treated except that they except that they receive a low protein diet weigh 8,10,14,10 and 13 0unces .
Test at 5% los whether there is significant evidence that additional protein has increased the weight of chicken .

```
> #Q5
> #H0:M1=M2 vs H1:M1>M2,los=0.05
> High_protien=c(12,15,11,16,14,14,16)
> Low_protien=c(8,10,14,10,13)
> t.test(High_protien,Low_protien,alternative="greater",var.equal=T)

        Two Sample t-test

data:  High_protien and Low_protien
t = 2.3888, df = 10, p-value = 0.01902
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 0.7238325       Inf
sample estimates:
mean of x mean of y
       14        11

> # p value < los ; Reject H0
> # Hence we conclude that there is significant evidence that additional
 protein .
```

Q6. A group of 5 patients treated with medicine A weight 42,39,48,60 and 41 kg second group of 7 patients from the same hospital treated with medicine B weight 38,42,56,64,68,69 and 62 kgs. Do you agree with the claim that medicine B increase the weight significantly?

	X2				
42	38	-4	16	-19	361
39	42	-7	49	-15	255
48	56	2			
60	64				
41	68				
	69				
	62				

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

```
> #Q6 # Weights of patients treated with Medicine A
> medicine_A <- c(42, 39, 48, 60, 41)
> # Weights of patients treated with Medicine B
> medicine_B <- c(38, 42, 56, 64, 68, 69, 62)
> # Perform unpaired t-test
> t_test_result <- t.test(medicine_B, medicine_A, alternative = "greater")
> # Print the result
> print(t_test_result)

        Welch Two Sample t-test

data:  medicine_B and medicine_A
t = 1.8195, df = 9.9996, p-value = 0.04943
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 0.04278476        Inf
sample estimates:
mean of x mean of y 
       57        46 
```

Practical 8

Sign Test

SIGN TEST

Q1.THE FOLLOWING DATA GIVES THE WEIGHTS (INKG) OF 40 STUDENTS IN A RANDOM SAMPLE

46,49,57,64,46,67,54,48,69,61,57,54,50,48,65,61,66,54,50,48,59,62,
47,49,47,55,59,63,53,56,67,49,60,64,53,50,48,51,52,54

USE SIGN TEST TO TEST WHETHER THE MEDIAN WEIGHT IN THE POPULATION IS 50 KG AGAINST ALTERNATIVE IT IS GREATER THAN

50KG

```
> #Kunal Mistri
> #3432
> #H0: M=50,H1>50,LOS=5%
> x=c(46,49,57,64,46,67,54,48,69,61,57,54,50,48,65,61,66,54,50,48,59,62,47,49,47,55,59,63,53,56,67,4
9,60,64,53,50,48,51,52,54)
> M=50
> SP=length(x[x>M])
> SN=length(x[x<M])
> n=SN+SP
> PVALUE=pbinom(SN,n,0.5)
> l.o.s=0.05
> if(PVALUE>l.o.s)(cat("accept H0"))else(cat("reject H0"))
reject H0NULL
```

```
> #QUESTION2
> #h0:M=41,H1>41,LOS=5%
> x=c(25,29,52,48,57,39,45,36,30,49,28,39,44,63,32,65,42)
> M=41
> SP=length(x[x>M])
> SN=length(x[x<M])
> n=SN+SP
> PVALUE=pbinom(SN,n,0.5)
> l.o.s=0.05
> if(PVALUE>l.o.s)(cat("accept H0"))else(cat("reject H0"))
accept H0NULL
```

WILCOXON SIGNED RANK TEST

```
> #WILCOXON SIGNED RANK TEST
> #QUESRION 1
> #H0:M=20,h1M>20,LOS=5%
> x=c(15,17,24,25,20,21,32,28,12,25,24,26)
> wilcox.test(x,alter="greater",mu=20)

        Wilcoxon signed rank test with continuity
        correction

data:  x
V = 48.5, p-value = 0.09051
alternative hypothesis: true location is greater than 20

Warning messages:
1: In wilcox.test.default(x, alter = "greater", mu = 20) :
  cannot compute exact p-value with ties
2: In wilcox.test.default(x, alter = "greater", mu = 20) :
  cannot compute exact p-value with zeroes

> # pvalue=0.09051>0.05,accept H0
```

Q2.A random sample of 10 infants showed the following pulse rate: minute:110,121,125,122,112,117,129,114,124,127.Assuming that th distribution of pulse rates is symmetric. Is there any evidence to s the median pulse rate of infants is more than 120 beats per minut Wilcoxon Signed Rank Test At 5%l.o.s.

```
> #question-2
> #H0:M=120,h1M>120,LOS=5%
> x=c(110,121,125,122,112,117,129,114,124,127)
> wilcox.test(x,alter="greater",mu=120)

        Wilcoxon signed rank exact test

data:  x
V = 28, p-value = 0.5
alternative hypothesis: true location is greater than 120
```

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

Q3.TO DETERMINE THE EFFECTIVENESS OF A NEW TRAFFIC CONTROL SYSTEM THE NUMBER OF ACCIDENTS OCCURRED AT 12 DIFFERENT LOCATIONS DURING FOUR WEEKS BEFORE AND AFTER
THE INSTALLATION OF NEW SYSTEM WERE OBSERVED AND RECORDED DATA IS AS FOLLOWS LOCATION

LOCATION	1	2	3	4	5	6	7	8	9	10	11	12
BEFORE	3	5	2	3	3	3	0	4	1	6	4	1
AFTER	1	2	0	2	2	0	2	3	3	4	1	0

USE WILCOXON SIGN TEST AT5%L.O.S TO TEST WHETHER I
TRAFFIC CONTROL SYSTEM IS EFFECTIVE

```
> #H0:M1=M2,H1:M1>M2,L.O.S=5%
> x=c(3,5,2,3,3,3,0,4,1,6,4,1)
> y=c(1,2,0,2,2,0,2,3,3,4,1,0)
> d=x-y
> wilcox.test(d,mu=0,alter="greater")

        Wilcoxon signed rank test with continuity
        correction

data:  d
V = 64, p-value = 0.02573
alternative hypothesis: true location is greater than 0
```

```
Warning message:
In wilcox.test.default(d, mu = 0, alter = "greater") :
  cannot compute exact p-value with ties
```

#Reject H0

Q4.THE WEIGHTS (INKGS)OF THE PERSONS BEFORE AND AF
STOPPED SMOKING ARE AS FOLLOWS

PERSON	1	2	3	4	5
WEIGHTS BEFORE	65	75	75	62	72
WEIGHTS AFTER	72	82	72	66	73

USE WILCOXON SIGN TEST AT5% L.O.S TO TEST WHETHER
OF A PERSON INCREASES AFTER STOPPING THE SMOKING.

Research

```
> #H0:M1=M2,H1:M1>M2,L.O.S=5%
> x=c(65,75,75,62,72)
> y=c(72,82,72,66,73)
> d=x-y
> wilcox.test(d,mu=0,alter="greater")

        Wilcoxon signed rank test with continuity
        correction

data:  d
V = 2, p-value = 0.948
alternative hypothesis: true location is greater than 0
```

```
Warning message:
In wilcox.test.default(d, mu = 0, alter = "greater") :
  cannot compute exact p-value with ties
```

```
> #accept H0
```

Q4.THE WEIGHTS (INKGS)OF THE PERSONS BEFORE AND AF
STOPPED SMOKING ARE AS FOLLOWS

PERSON	1	2	3	4	5
WEIGHTS BEFORE	65	75	75	62	72
WEIGHTS AFTER	72	82	72	66	73

USE WILCOXON SIGN TEST AT5% L.O.S TO TEST WHETHER
OF A PERSON INCREASES AFTER STOPPING THE SMOKING.

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

Practical No 9

Case 1 ONE WAY ANOVA

Q1.Q1.A DRUG COMPANY TESTED THREE FORMULATIONS OF A N RELIEF MEDICINE FOR MIGRAINE HEADACHE SUFFERS FOR THE EXPERIMENT 27 VOLUNTEERS WERE SELECTED AND 9 WERE RANDOMLY ASSIGN THE ONE OF THREE DRUGS FORMULATION THE SUBJECTS WERE INSTRUCTED TO TAKE THE DRUG DURING THEIR NEXT MIGRAINE HEADACHE AND REPORT THEIR PAIN ON THE SCALEOF 1TO 10

```
> #Kunal Mistri
> #3432
> #Q1 H0:m1=m2=m3
> #H1:at least two differ from each other
> drug = c(rep("A",9),rep("B",9),rep("C",9))
> pain=c(4,5,4,3,2,4,3,4,4,6,8,4,5,4,6,5,8,6,6,7,6,6,7,5,6,5,5)
> migraine=data.frame(pain,drug)
> migraine
   pain drug
1     4    A
2     5    A
3     4    A
4     3    A
5     2    A
6     4    A
7     3    A
8     4    A
9     4    A
10    6    B
11    8    B
12    4    B
13    5    B
14    4    B
15    6    B
16    5    B
17    8    B
18    6    B
19    6    C
20    7    C
21    6    C
22    6    C
23    7    C
24    5    C
25    6    C
26    5    C
```

Q2.THREE PROCESS A,B,C ARE TESTED TO SEE WHETHER THERE OTPUTS ARE EQUIVALENT THE FOLLOWING OBSERVATIONS OF OUTPUT ARE MADE CARRY OUT ANOVA AND STATE YOUR CONCLUSION

A:20,22,23,21,20,24,25,23

B:19,21,20,22,23

C:21,20,20,25,24,22,23

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

```
> #Q2
> #H0: three test outputs are equivalent
> #H1:at least 2 test are different form other
> tput = c(20,22,23,21,20,24,25,23,19,21,20,22,23,21,20,20,25,24,22,23)
> process=c(rep("A",8),rep("B",5),rep("C",7))
> data=data.frame(tput,process)
> data
   tput process
1    20       A
2    22       A
3    23       A
4    21       A
5    20       A
6    24       A
7    25       A
8    23       A
9    19       B
10   21       B
11   20       B
12   22       B
13   23       B
14   21       C
15   20       C
16   20       C
17   25       C
18   24       C
19   22       C
20   23       C
```

TWO WAY ANOVA

Q1. A TEA COMPANY APPOINTS FOUR SALESMAN 1,B,C,AND D AND OBSERVES THEIR SALES IN THREE

SEASONS-SUMMER, WINTER, AND MONSOON, THE FIGURES (IN LAKHS) OF SALES ARE GIVEN IN THE FOLLOWING TABLE

SEASON/SALES MA

Season Sales	A	B	C	D
Summer	36	32	21	30
Winter	24	25	20	22
Mansoon	20	10	19	15

Name: Kunal Manojkumar Mistri
Class: Msc.DA
Roll No: 3432

```
> #Q3
> f1=c(rep(1:3,rep(4,3)))
> f2=rep(c("A","B","C","D"),3)
> season=factor(f1)
> salesman=factor(f2)
> ans=aov(sales~season+salesman)
> summary(ans)
            Df Sum Sq Mean Sq F value Pr(>F)
season       2   2239    1119   0.381  0.699
salesman     3   4094    1365   0.464  0.718
Residuals    6  17645    2941
```

Q2 five differnet fertililizers are used and three types of seeds are grown the yeild obtained (INKGS) are tabulated below.

	FERT1	FERT2	FERT3	FERT4	FERT5
Seed1	110	100	107	104	102
Seed2	112	97	101	112	107
Seed3	97	87	99	101	98

```
> #Q4 five differnet fertililizers are used and three types of seeds are grown
 the yeild obtained (INKGS) are tabulated below
> yield=c(110,100,107,104,102,112,99,101,112,107,97,87,99,101,98)
> f1=c(1,1,1,1,1,2,2,2,2,2,3,3,3,3,3)
> f2=rep(c(1:5),3)
> seed=factor(f1)
> fertilizer=factor(f2)
> ans=aov(yield~seed+fertilizer)
> summary(ans)
            Df Sum Sq Mean Sq F value  Pr(>F)
seed         2  276.4  138.20  10.954 0.00512 **
fertilizer   4  228.3   57.07   4.523 0.03335 *
Residuals    8  100.9   12.62
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Practical 10

FRIEDMAN TEST

Q1.THE FOLLOWING shows the reaction time of five patients on fo
different drugs. Since each patient is measured on each of the fou
will use the Friedman Test to determine if the mean reaction time
between drugs.

PERSON=1,2,3,4,5,DRUG=1,2,3,4,

SCORE=30,28,16,34,14,18,10,22,24,20,18,30,38,34,20,44,26,28,14,

```
> #H0:the mean reaction time differs between drugs
> #H1: the mean reaction times does not differs between drugs
> data=data.frame(person=rep(1:5,each=4),drug=rep(c(1,2,3,4),times=5

score=c(30,28,16,34,14,18,10,22,24,20,18,30,38,34,20,44,26,28,14,30))
```

```
> #Q2
> #H0:significant difference between average time taken on this machine to complete the job
> #H1:signifcant difference
> X=c(2.9,3.0,2.5,2.6,3.2)
> Y=c(3.8,2.7,4.0,2.4)
> Z=c(2.8,3.4,3.7,2.2,2.0)
> kruskal.test(list(X,Y,Z))

        Kruskal-Wallis rank sum test

data:  list(X, Y, Z)
Kruskal-Wallis chi-squared = 0.77143, df = 2, p-value = 0.68

> #pvalue=0.68>0.05 we accept H0
> data
   person drug score
1       1    1    30
2       1    2    28
3       1    3    16
4       1    4    34
5       2    1    14
6       2    2    18
7       2    3    10
8       2    4    22
9       3    1    24
10      3    2    20
11      3    3    18
12      3    4    30
13      4    1    38
14      4    2    34
15      4    3    20
16      4    4    44
17      5    1    26
18      5    2    28
19      5    3    14
20      5    4    30
> friedman.test(y=data$score,groups=data$drug,blocks=data$person)

        Friedman rank sum test

data:  data$score, data$drug and data$person
Friedman chi-squared = 13.56, df = 3, p-value = 0.00357
```

Q2.TEST THE CLAIM AT 5% L.O.S THAT ALL THREE DRUGS SAME PROBABILITY DISTRIBUTION USE FRIEDMAN TEST

REACTION TIME FOR THREE DRUGS

SUBJECT	DRUGA	DRUGB	DRUGC
1	1.21	1.48	1.56
2	1.63	1.85	2.01
3	1.42	2.06	1.70
4	2.43	1.98	2.64
5	1.16	1.27	1.48
6	1.94	2.44	2.81

```
> #Ganesh Sawant
> #3448

> # Reaction time data for each drug across subjects
> drugA <- c(1.21, 1.63, 1.42, 2.43, 1.16, 1.94)
> drugB <- c(1.48, 1.85, 2.06, 1.98, 1.27, 2.44)
> drugC <- c(1.56, 2.01, 1.70, 2.64, 1.48, 2.81)
>
> # Combine data into a data frame
> reaction_time <- data.frame(
+   subject = factor(1:6),
+   drugA = drugA,
+   drugB = drugB,
+   drugC = drugC
+ )
>
> result <- friedman.test(as.matrix(reaction_time[,2:4]))
> print(result)

        Friedman rank sum test

data:  as.matrix(reaction_time[, 2:4])
Friedman chi-squared = 8.3333, df = 2, p-value = 0.0155

> #Reject H0
```

KRUSKAL WALLIS TEST

Q1 TEST THE HYPOTHESIS THAT THERE IS NO SIGNIFICANT DIFFERENCE BETWEEN MEANS OF ALL FOUR GROUPS

GROUP A:1,5,8,17,16

GROUP B:2,1,6,5,7,4

GROUP C:1,1,3,7,9

GROUP D:2,1,5,2,9,7

```
> #Q2
> #H0:no significance difference between means of 4 groups
> #H1:signifcance difference between means 4 group
> A=c(1,5,8,17,16)
> B=c(2,1,6,5,7,4)
> C=c(1,1,3,7,9)
> D=c(2,1,5,2,9,7)
> kruskal.test(list(A,B,C,D))

        Kruskal-Wallis rank sum test

data:  list(A, B, C, D)
Kruskal-Wallis chi-squared = 2.1699, df = 3, p-value = 0.5379

> #pvalue>=0.5379>0.05 we accept H0
```

Q2. THE TIME TAKEN TO COMPLETE A JOB ON THREE MACHINES ARE NOTED TEST THE HYPOTHESIS THAT THERE IS NO SIGNIFICANT DIFFERENCE BETWEEN AVERAGE TIME TAKEN ON THIS MACHINES TO COMPLETE THE JOB

MACHINE X:2.9,30,2.5,2.6,3.2

MACHINE Y:3.8,2.7,4.0,2.4

MACHINE Z:2.8,3.4,3.7,2.2,2.0

```
> #Q2
> #H0:significant difference between average time taken on this machine to complete the job
> #H1:signifcant difference
> X=c(2.9,3.0,2.5,2.6,3.2)
> Y=c(3.8,2.7,4.0,2.4)
> Z=c(2.8,3.4,3.7,2.2,2.0)
> kruskal.test(list(X,Y,Z))

        Kruskal-Wallis rank sum test

data:  list(X, Y, Z)
Kruskal-Wallis chi-squared = 0.77143, df = 2, p-value = 0.68

> #pvalue=0.68>0.05 we accept H0
```