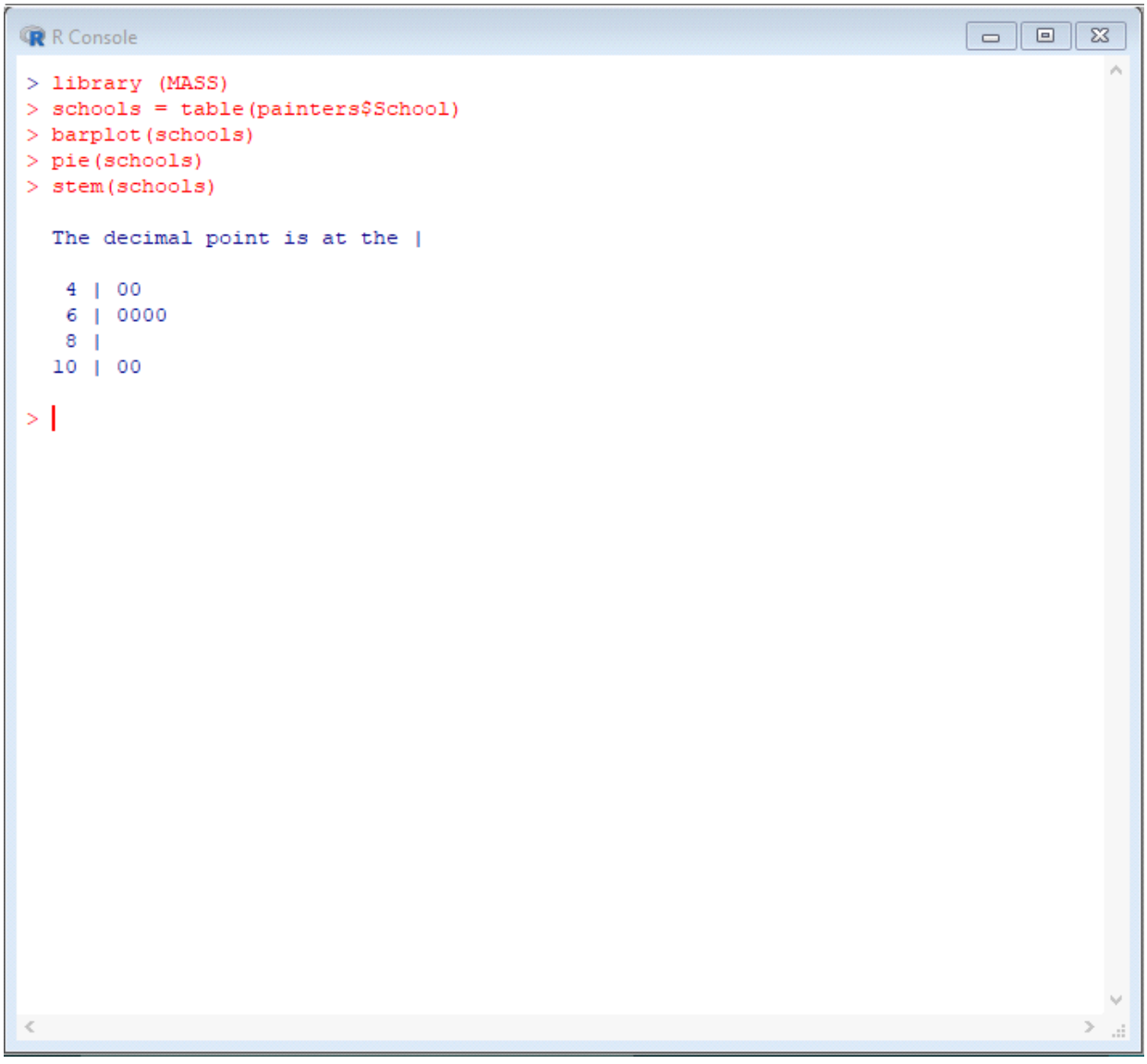
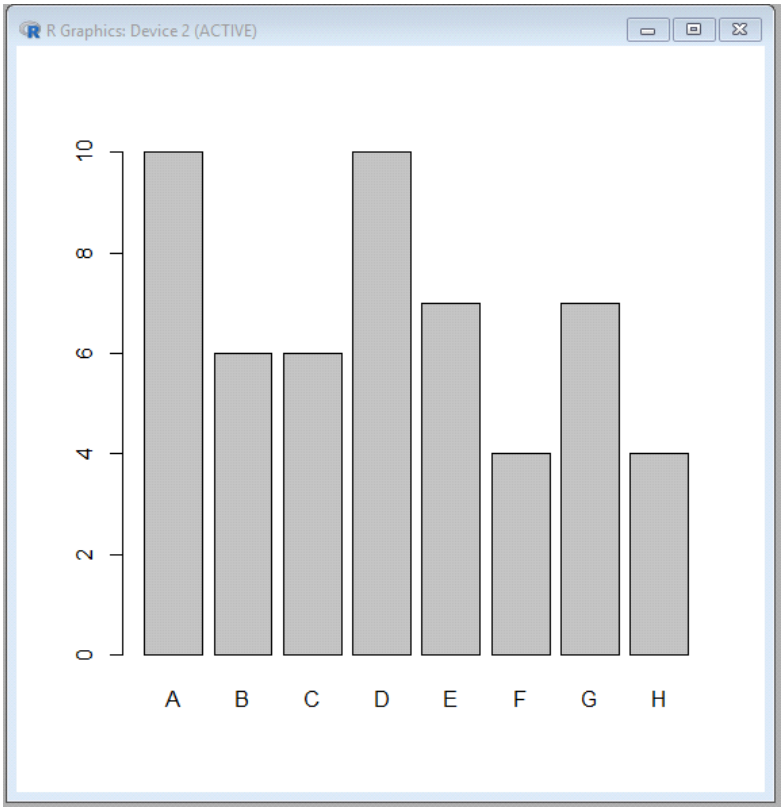
PRACTICAL 1

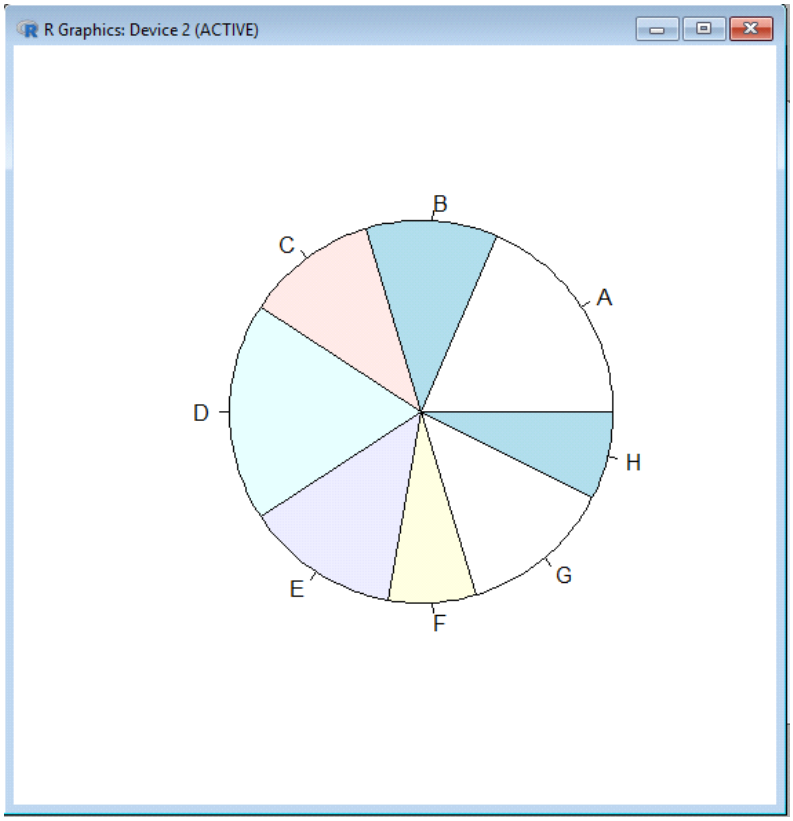
Aim:- Data presentation.

```
library (MASS)
schools = table(painters$School)
barplot(schools)
pie(schools)
stem(schools)
painters
```

Output:-




R Graphics: Device 2 (ACTIVE)
0
2
4
6
8
10
A
B
C
D
E
F
G
H


R Graphics: Device 2 (ACTIVE)
B
C
A
D
H
G
E
F

PRACTICAL 2.a

Aim:- Data entry using, functions, c(), scan (), Creating vectors, Mathematical Operations: ** +/-/*/ / ^ , exp, log, log10, etc, creating vector of text type, useful functions: data, frame, matrix operations, seq(), split() etc.

#Practical 2.a

```
x = scan("",what = "int")
x
v1 = c(3,8,4,5,0,11)
v2 = c(4,11,0,8,1,2)
result.add = v1 + v2
result.add
result.sub = v1 - v2
result.sub
result.mul = v1 * v2
result.mul
result.div = v1 / v2
result.div
exp(v1)
exp(v2)
log10(v1)
log10(v2)
v1 = c(5:13)
v1
v2 = c(6.6:12.6)
v2
```

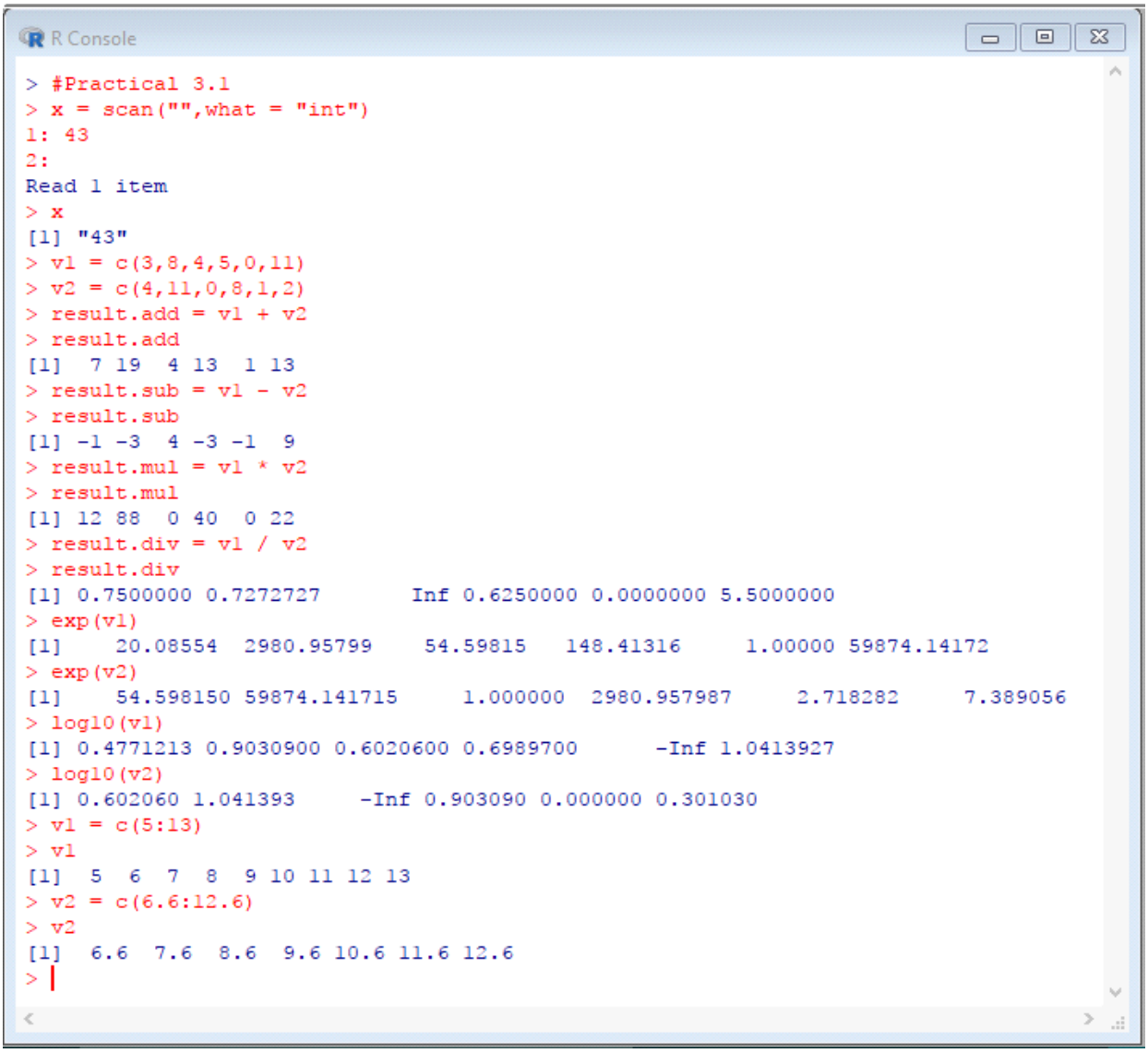
#Practical 2.b

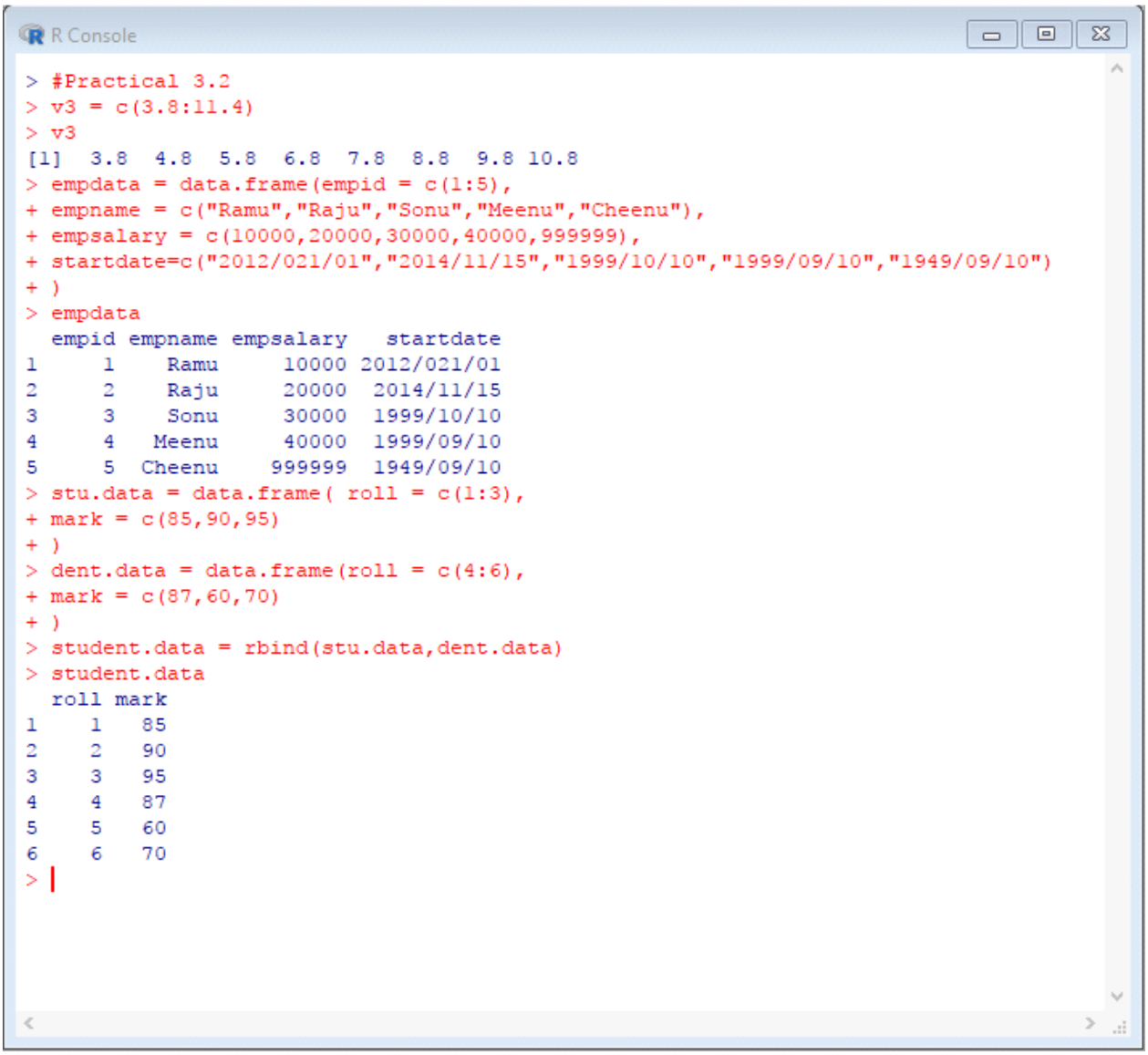
```
 v3 = c(3.8:11.4)
 v3
empdata = data.frame(empid = c(1:5),
empname = c("Ramu","Raju","Sonu","Meenu","Cheenu"),
empsalary = c(10000,20000,30000,40000,999999),
startdate=c("2012/021/01","2014/11/15","1999/10/10","1999/09/10","1949/09/10") )
empdata
stu.data = data.frame( roll = c(1:3),
mark = c(85,90,95)
)
dent.data = data.frame(roll = c(4:6),
mark = c(87,60,70)
)
student.data = rbind(stu.data,dent.data)
student.data
```

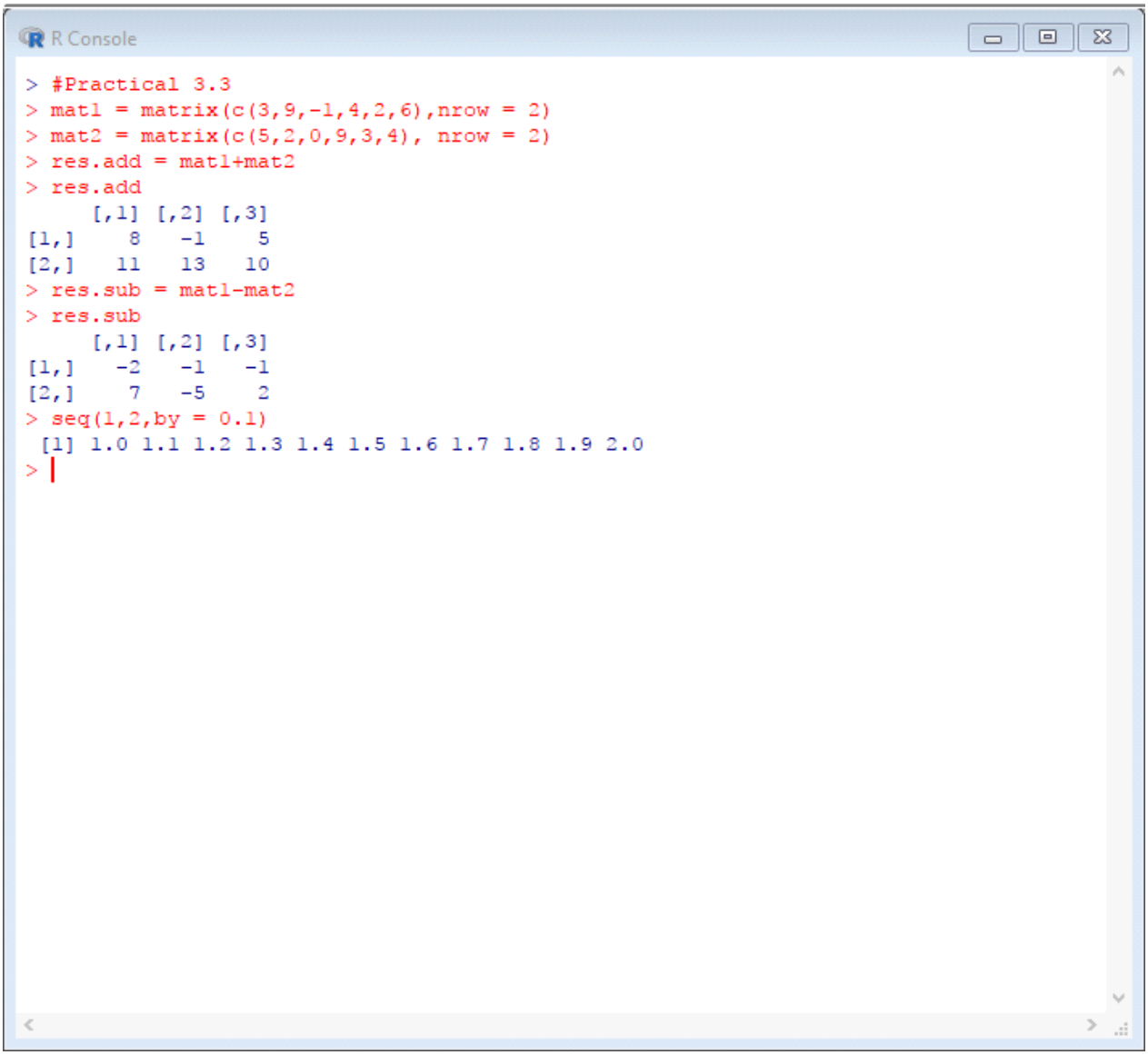
#Practical 2.c

```
mat1 = matrix(c(3,9,-1,4,2,6),nrow = 2)
mat2 = matrix(c(5,2,0,9,3,4), nrow = 2)
res.add = mat1+mat2
res.add
res.sub = mat1-mat2
res.sub
seq(1,2,by = 0.1)
```

Output:-




R Console
> #Practical 3.2
> v3 = c(3.8:11.4)
> v3
[1] 3.8 4.8 5.8 6.8 7.8 8.8 9.8 10.8
> empdata = data.frame(empid = c(1:5),
+ empname = c("Ramu","Raju","Sonu","Meenu","Cheenu"),
+ empsalary = c(10000,20000,30000,40000,999999),
+ startdate=c("2012/021/01","2014/11/15","1999/10/10","1999/09/10","1949/09/10")
+)
> empdata
  empid empname empsalary   startdate
1   1   Ramu   10000 2012/021/01
2   2   Raju   20000 2014/11/15
3   3   Sonu   30000 1999/10/10
4   4   Meenu   40000 1999/09/10
5   5 Cheenu   999999 1949/09/10
> stu.data = data.frame(roll = c(1:3),
+ mark = c(85,90,95)
+)
> dent.data = data.frame(roll = c(4:6),
+ mark = c(87,60,70)
+)
> student.data = rbind(stu.data,dent.data)
> student.data
  roll mark
1   1   85
2   2   90
3   3   95
4   4   87
5   5   60
6   6   70
>


R Console
> #Practical 3.3
> mat1 = matrix(c(3,9,-1,4,2,6),nrow = 2)
> mat2 = matrix(c(5,2,0,9,3,4), nrow = 2)
> res.add = mat1+mat2
> res.add
[,1] [,2] [,3]
[1,] 8 -1 5
[2,] 11 13 10
> res.sub = mat1-mat2
> res.sub
[,1] [,2] [,3]
[1,] -2 -1 -1
[2,] 7 -5 2
> seq(1,2,by = 0.1)
[1] 1.0 1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8 1.9 2.0
>

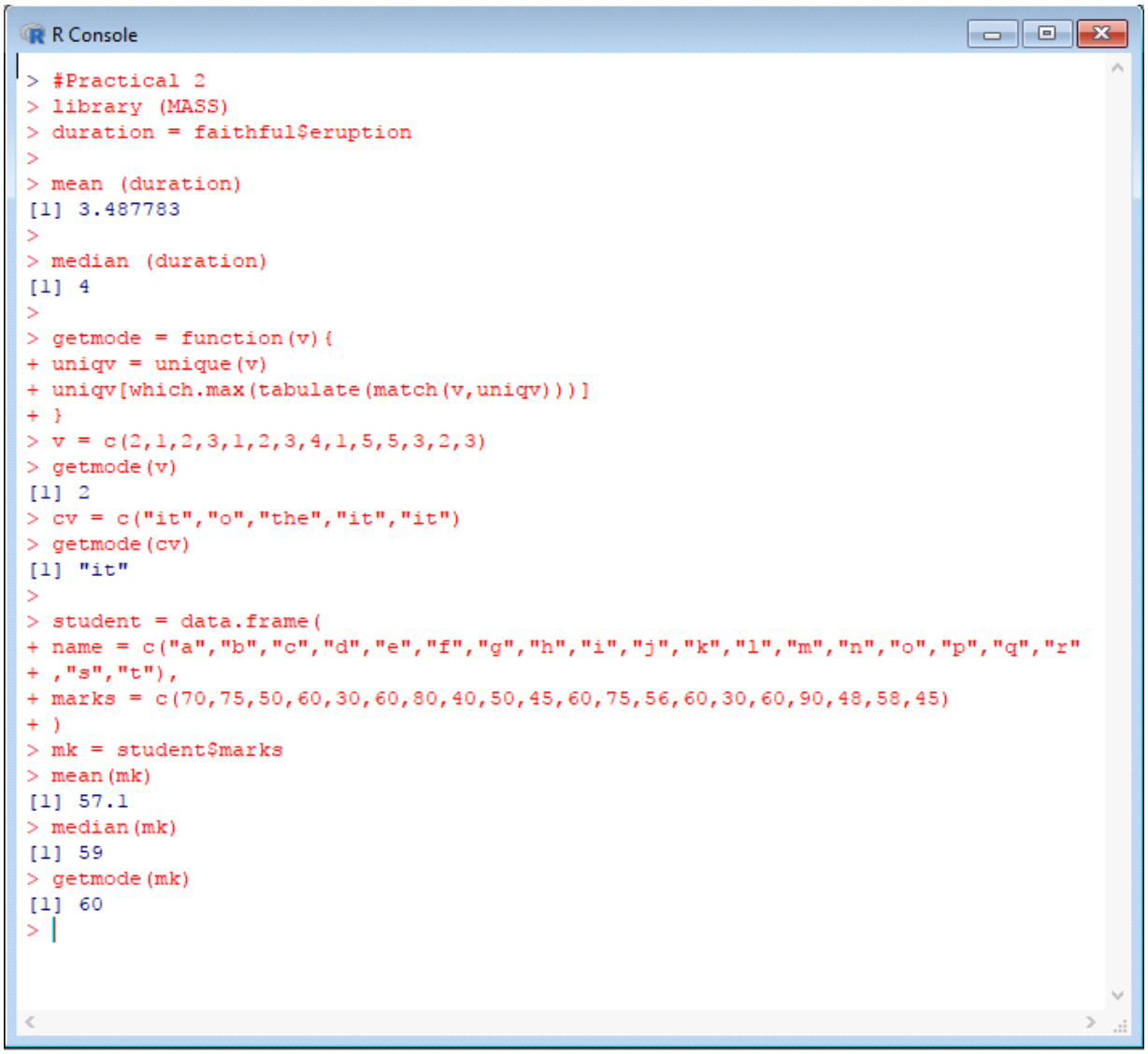
PRACTICAL 3

Aim:- Measures of central tendency

```
library (MASS)
duration = faithful$eruption
mean (duration)
median (duration)
getmode = function(v){
        uniqv = unique(v)
        uniqv[which.max(tabulate(match(v,uniqv)))]
}
v = c(2,1,2,3,1,2,3,4,1,5,5,3,2,3)
getmode(v)
cv = c("it","o","the","it","it")
getmode(cv)

student = data.frame(
name = c("a","b","c","d","e","f","g","h","i","j","k","l","m","n","o","p","q","r"
,"s","t"),
marks = c(70,75,50,60,30,60,80,40,50,45,60,75,56,60,30,60,90,48,58,45)
)
mk = student$marks
mean(mk)
median(mk)
getmode(mk)
```

Output:-



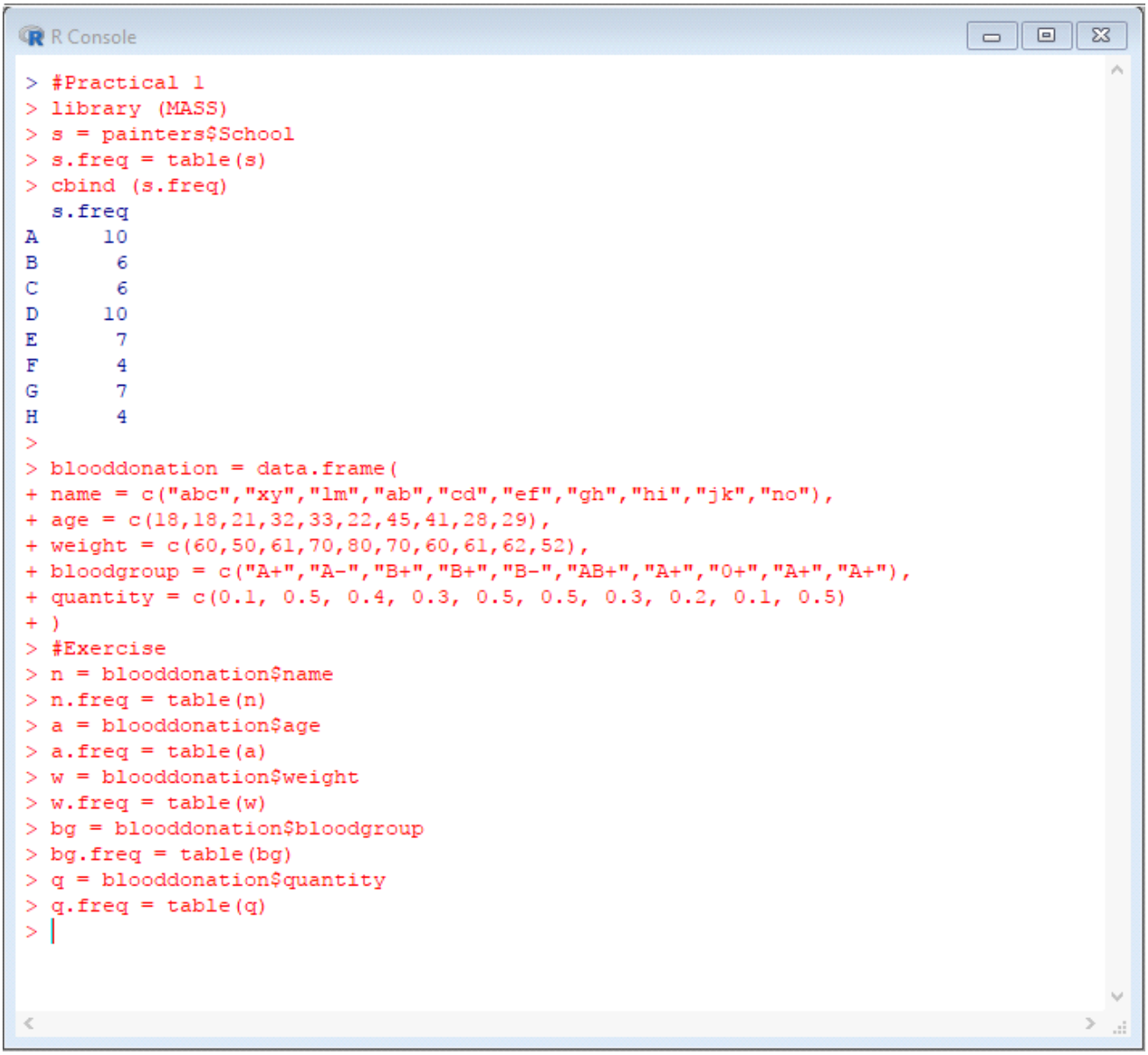
PRACTICAL 4

Aim:- Frequency distribution and data presentation

```
library (MASS)
s = painters$School
s.freq = table(s)
cbind (s.freq)

blooddonation = data.frame(
name = c("abc","xy","lm","ab","cd","ef","gh","hi","jk","no"),
age = c(18,18,21,32,33,22,45,41,28,29),
weight = c(60,50,61,70,80,70,60,61,62,52),
bloodgroup = c("A+","A-","B+","B+","B-","AB+","A+","0+","A+","A+"),
quantity = c(0.1, 0.5, 0.4, 0.3, 0.5, 0.5, 0.3, 0.2, 0.1, 0.5)
)
#Exercise
n = blooddonation$name
n.freq = table(n)
a = blooddonation$age
a.freq = table(a)
w = blooddonation$weight
w.freq = table(w)
bg = blooddonation$bloodgroup
bg.freq = table(bg)
q = blooddonation$quantity
q.freq = table(q)
```

Output:-



PRACTICAL 5.a

Aim:- Frequency distribution using cut(), table()

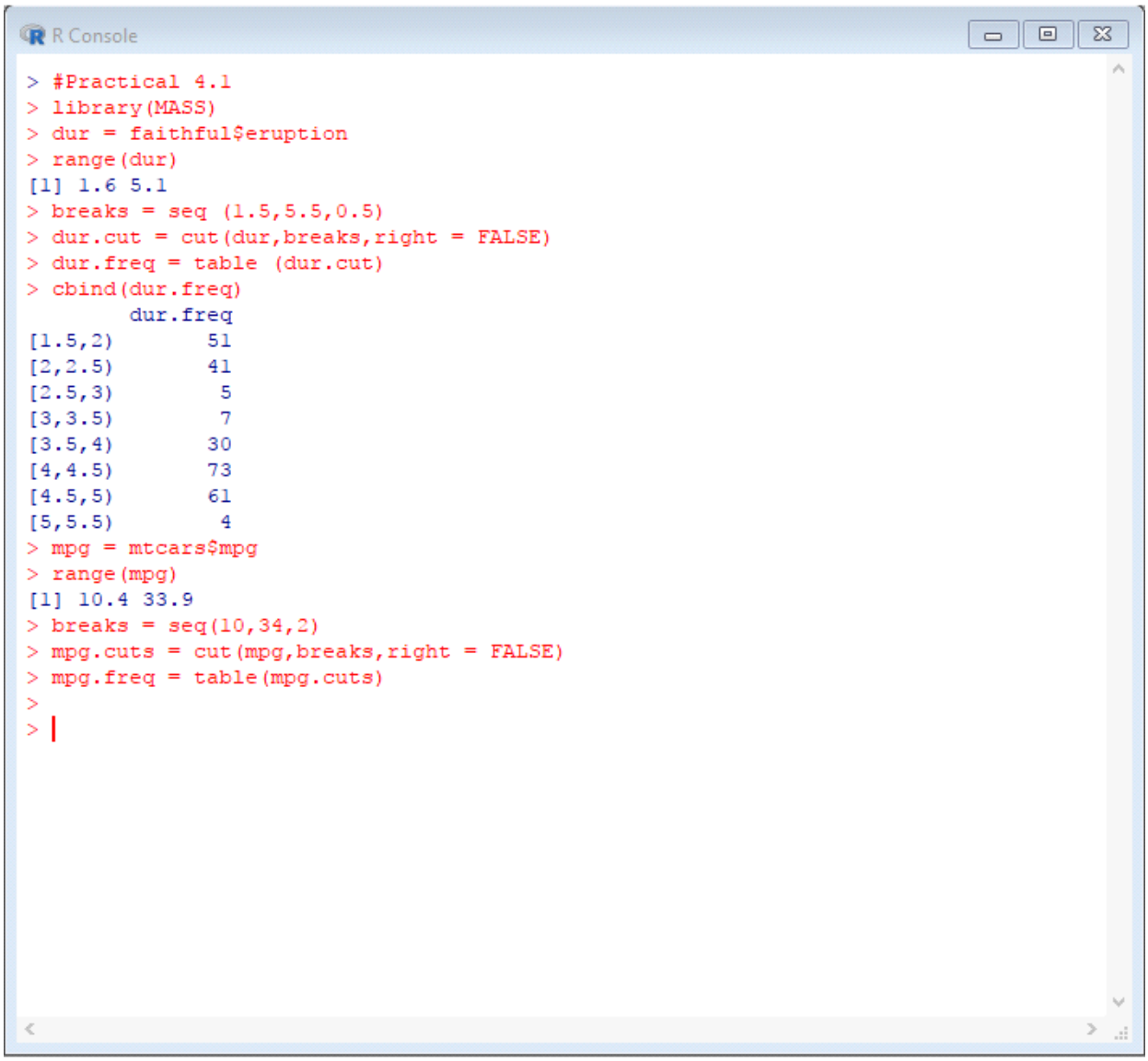
#Practical 5.a

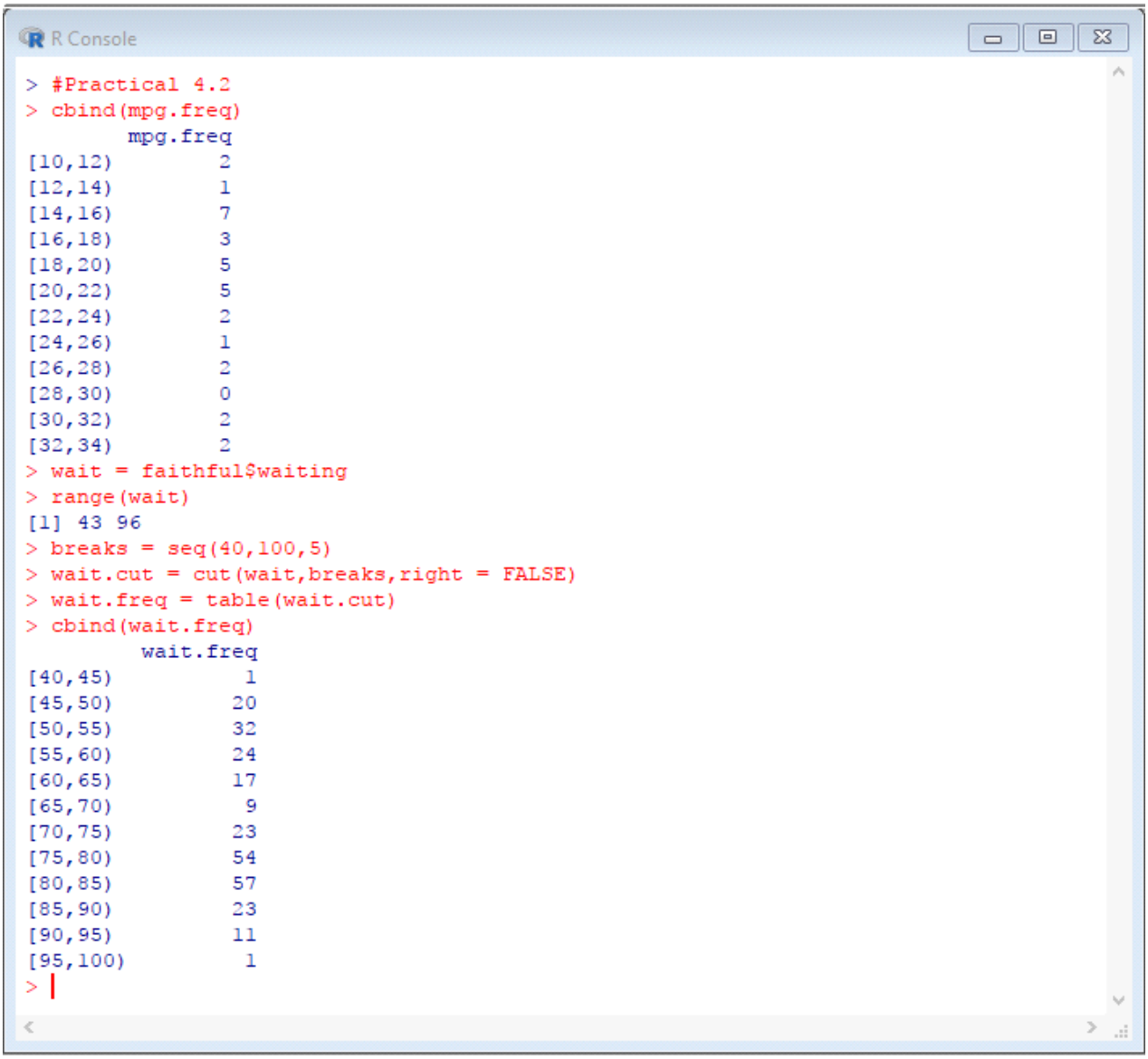
```
library(MASS)
dur = faithful$eruption
range(dur)
breaks = seq (1.5,5.5,0.5)
dur.cut = cut(dur,breaks,right = FALSE)
dur.freq = table (dur.cut)
cbind(dur.freq)
mpg = mtcars$mpg
range(mpg)
breaks = seq(10,34,2)
mpg.cuts = cut(mpg,breaks,right = FALSE)
mpg.freq = table(mpg.cuts)
```

#Practical 5.b

```
cbind(mpg.freq)
 wait = faithful$
waiting range(wait)
breaks = seq(40,100,5)
wait.cut = cut(wait,breaks,right = FALSE)
wait.freq = table(wait.cut)
cbind(wait.freq)
```

Output:-



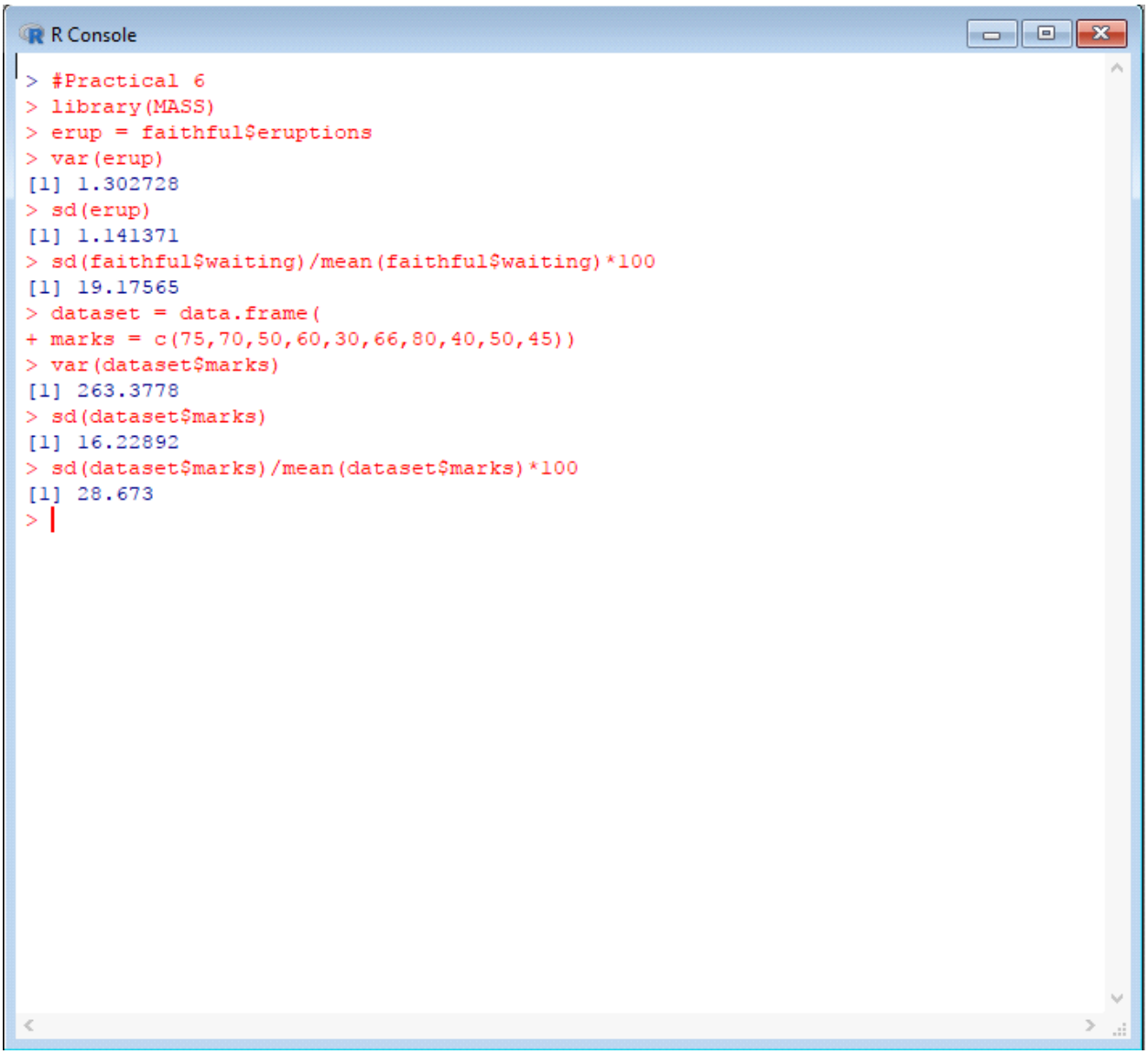

R Console
> #Practical 4.2
> cbind(mpg.freq)
mpg.freq
[10,12) 2
[12,14) 1
[14,16) 7
[16,18) 3
[18,20) 5
[20,22) 5
[22,24) 2
[24,26) 1
[26,28) 2
[28,30) 0
[30,32) 2
[32,34) 2
> wait = faithful$waiting
> range(wait)
[1] 43 96
> breaks = seq(40,100,5)
> wait.cut = cut(wait,breaks,right = FALSE)
> wait.freq = table(wait.cut)
> cbind(wait.freq)
wait.freq
[40,45) 1
[45,50) 20
[50,55) 32
[55,60) 24
[60,65) 17
[65,70) 9
[70,75) 23
[75,80) 54
[80,85) 57
[85,90) 23
[90,95) 11
[95,100) 1
>

PRACTICAL 6

Aim:- Summary Statistics (measures of central tendency, dispersion)

```
library(MASS)
erup = faithful$eruptions
var(erup)
sd(erup)
sd(faithful$waiting)/mean(faithful$waiting)*100
dataset = data.frame(
marks = c(75,70,50,60,30,66,80,40,50,45))
var(dataset$marks)
sd(dataset$marks)
sd(dataset$marks)/mean(dataset$marks)*100
```

Output:-

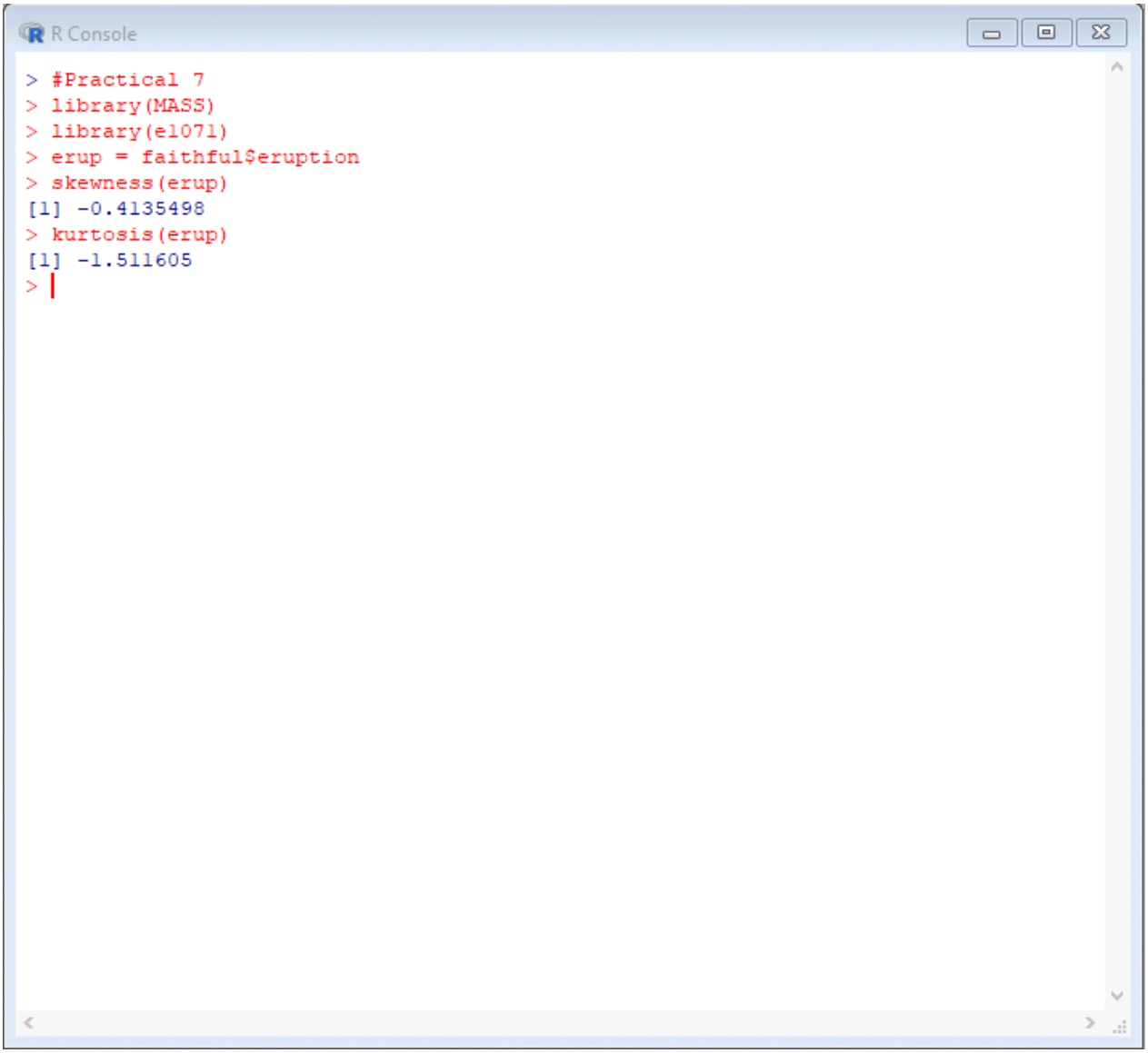


PRACTICAL 7

Aim:- Measures of skewness and kurtosis

```
library(MASS)
library(e1071)
erup = faithful$eruption
skewness(erup)
kurtosis(erup)
```


Output:-

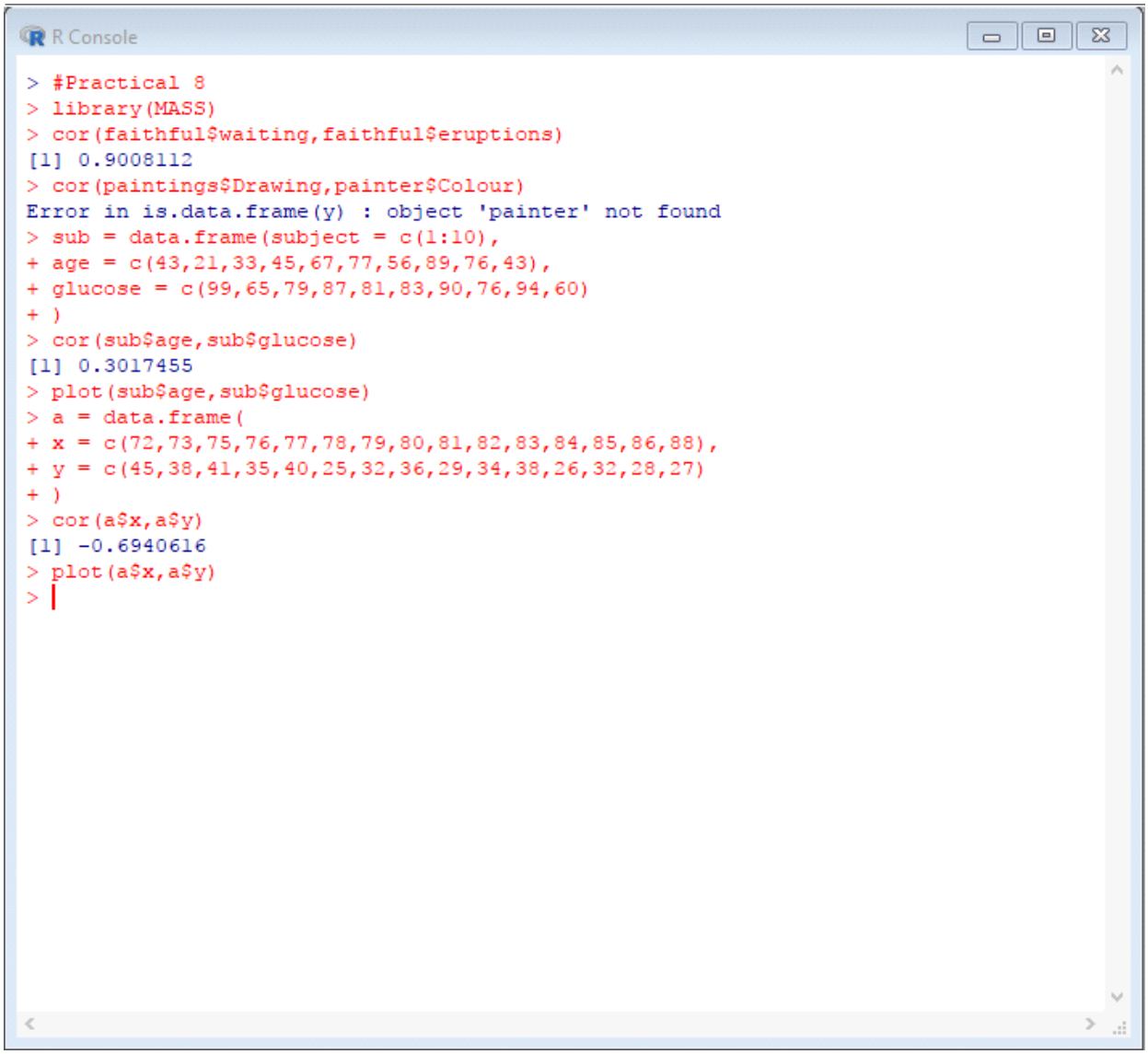


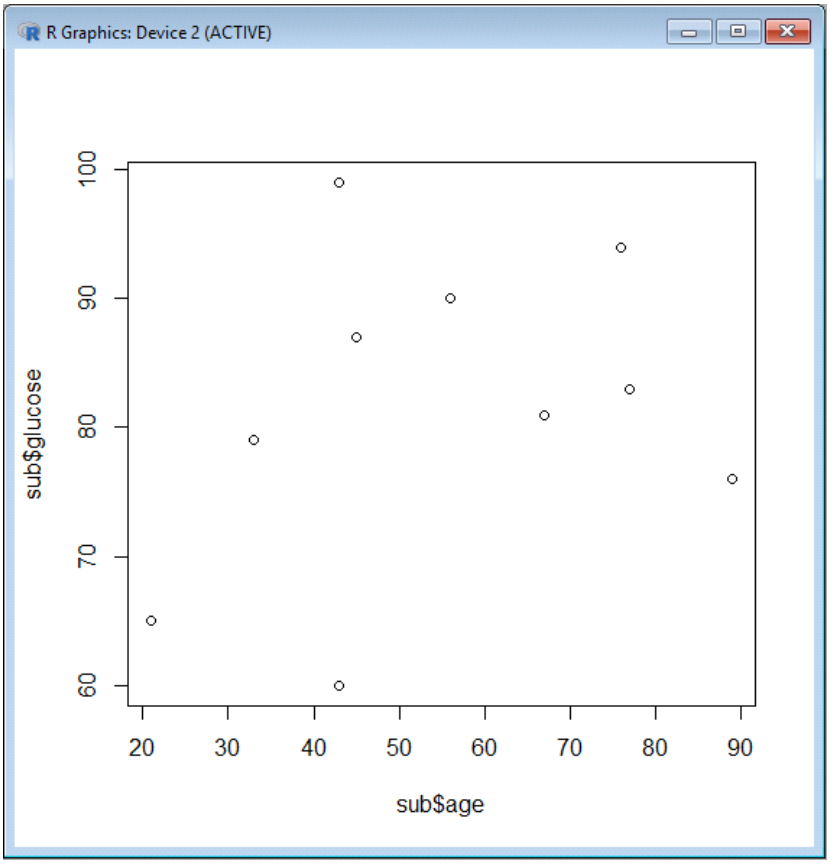
PRACTICAL NO:8

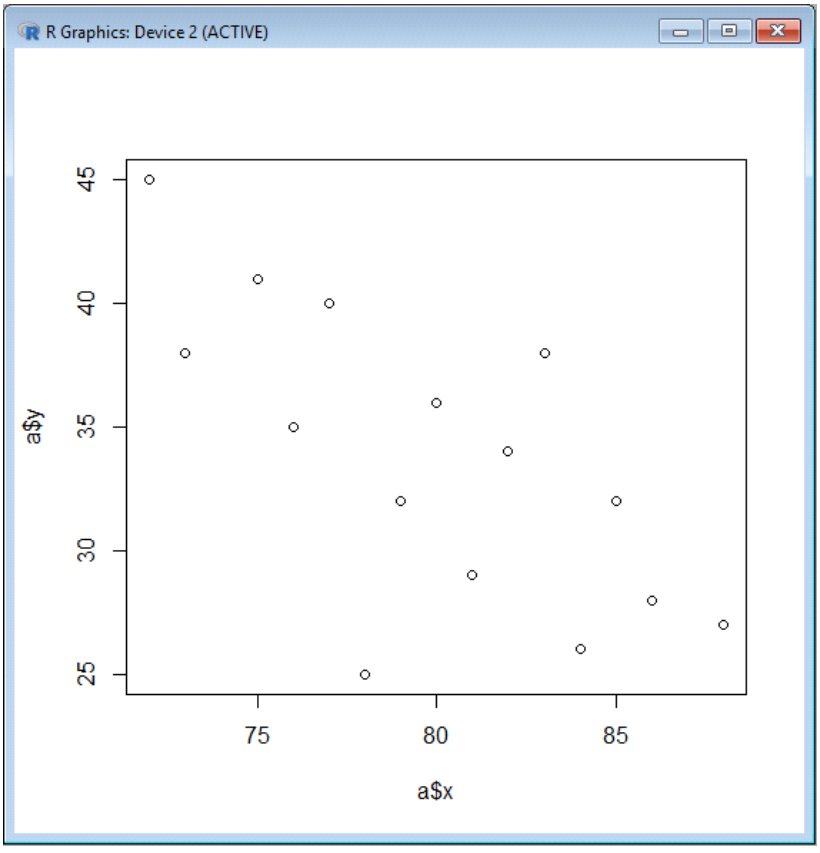
Aim:- Correlation and regression

```
library(MASS)
cor(faithful$waiting,faithful$eruptions)
cor(painters$Drawing,painters$Colour)
sub = data.frame(subject = c(1:10),
age = c(43,21,33,45,67,77,56,89,76,43),
glucose = c(99,65,79,87,81,83,90,76,94,60)
)
cor(sub$age,sub$glucose)
plot(sub$age,sub$glucose)
a = data.frame(
x = c(72,73,75,76,77,78,79,80,81,82,83,84,85,86,88),
y = c(45,38,41,35,40,25,32,36,29,34,38,26,32,28,27)
)
cor(a$x,a$y)
plot(a$x,a$y)
```

Output:-




R Graphics: Device 2 (ACTIVE)
sub$glucose
sub$age
60
70
80
90
100
20
30
40
50
60
70
80
90


R Graphics: Device 2 (ACTIVE)
a$y
a$x
25
30
35
40
45
75
80
85

Example 18: An urn contains 10 white, 6 red and 9 black balls. If 6 balls are drawn at random find the probability that

(i) Four of the balls drawn are white

(ii) Two is of each colour

(iii) None is red

(iv) At least one is white

Solution: R code is

```
>n=choose(25,6)
> m1=choose(10,4)*choose(15,2)
> p1=m/n
> m2=choose(10,2)*choose(9,2)*choose(6,2)
> P2=m2/n
> m3=choose(19,4)
> p3=m3/n
> m4=choose(15,6)
> p4=1-m4/n
> cat("Prob four of the balls drawn are white=",p1)
Prob four of the balls drawn are white= 0.1245059
> cat("Prob Two  is of each colour =",p2)
Prob Two  is of each colour = 0.3666667
> cat("Prob None is red=",p3)
Prob None is red= 0.02188594
> cat("Prob at least one is white 1=",p4)
Prob at least one is white 1= 0.9717391
```

```
[1] 10 11 12 13 14 15
> p1=m1/n;p2=m2/n;p3=m3/n;p4=m4/n
> cat("Prob that token drawn has a number less than 6 =", p1)
Prob that token drawn has a number less than 6 = 0.2
> cat("Prob that token drawn has a number greater than 20 =", p2)
Prob that token drawn has a number greater than 20 = 0.2
> cat("Prob that token drawn has a number multiple of 5 =", p3)
Prob that token drawn has a number multiple of 5 = 0.2
> cat("Prob that token drawn has a number lying between 10 and 15,both inclusive =", p4)
Prob that token drawn has a number lying between 10 and 15 , both inclusive = 0.24
```

row for a photograph.

watches Hindi movies given that he watches English movies=",p4)

Example 20: A stockiest has 40 items in a lot. Out of which 30 are non-defective and 10 are defective. A customer selects 6 items from the lot. What is the probability that out of these six items (i) all items are non-defective (ii) four are non defective and two is defective.

Solution : R code is

```
> n= choose (40,6) ;m1= choose (30,6);m2= choose(30,4)* choose(10,2)
> p1=m1/n;p2=m2/n
> cat("Prob all selected are non-defective =",p1)
Prob all selected are non-defective = 0.1546942
> cat("Prob among selected 4 are non-defective and 2 defective =",p2)
Prob among selected 4 are non-defective and 2 defective = 0.3212879.
```

Example 21: A box contains 25 tokens numbered 1,2,3…….25. A token is drawn from a box and a number on it is observed. Obtain the probability that token drawn has a number

(i) less than 6

(ii) greater than 20

(iii) multiple of 5

(iv) lying between 10 and 15 , both inclusive.

Write sample space and events.

Solution:

R code is

```
> s=seq(1,25);n=length(s)
> s
 [1]  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
> e1=seq(1,5);m1= length(e1)
> e1
[1] 1 2 3 4 5
> e2=seq(21,25);m2= length(e2)
> e2
[1] 21 22 23 24 25
> e3=seq(5,25,5);m3= length(e3)
> e3
[1]  5 10 15 20 25
> e4=seq(10,15);m4= length(e4)
> e4
```