NAME- VARANASI KASYAP
REGISTRATION NUMBER-20BCE7315
Course Title: Applied Statistics
Instructor's name: Dr. Ankur
Course Code: MAT1011

Q1)
a) Enter the data {2,5,3,7,1,9,6} directly and store it in a variable x.

```
> x <- c(2,5,3,7,1,9,6)

> x

[1] 2 5 3 7 1 9 6
```

b) Find the number of elements in x.

```
> length(x)

[1] 7
```

c)Find the last element of x.

```
> x[length(x)]

[1] 6
```

d) Find the minimum and maximum elements of x.

```
> max(x)

[1] 9

> min(x)

[1] 1
```

Q2)
a)Enter the data {1, 2…,19,20} in a variable x.

```
> x <- 1:20

> x

 [1]  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
```

b)Find the 3rd to 5th element in the data list.

```
> 5(x, 3))

[1] 8
```

c)Find the 3rd element in the data list.

```
>  x[c(3)]

[1] 3
```

d)Find the 2nd,5th,6th,12th element in the list.

```
> x[c(2,5,6,12)]

[1]  2  5  6 12
```

e) Print the data as {20,19,…,2,1} without entering the data.

```
> numbers <- 20:1

> numbers

 [1] 20 19 18 17 16 15 14 13 12 11 10  9  8  7  6  5  4  3  2  1
```

Q3)

a)Enter data as 1,2,3. . . ,10

```
> x <- 1:10

> x

 [1]  1  2  3  4  5  6  7  8  9 10
```

b)Find sum of the numbers.

```
> sum(x)

[1] 55
```

c)Find Mean, median.

```
> sum(x)/length(x)

[1] 5.5  (or)

> mean(x)

[1] 5.5
```

```
> median(x)
[1] 5.5
```

d)Find sum of squares of these values.

```
> sum(x^2)
[1] 385
```

e) Find the value of1nn∑i=1|xi−¯x|, this is known as mean deviation about mean (M D¯x).

```
> sum(abs(x - mean(x)))/length(x)
[1] 2.5
```

f)Check whetherM D¯xis less than or equal to standard deviation.

```
> sd(x)
[1] 3.02765
> a=sum(abs(x - mean(x)))/length(x)
> a
[1] 2.5
> a==sd(x)
[1] FALSE
> a>x
 [1]  TRUE  TRUE
>
```

NAME- VARANASI KASYAP

REGISTRATION NUMBER-20BCE7315

Course Title: Applied Statistics

Instructor's name: Dr. Ankur

Course Code: MAT1011

Q.2) Create a data file as follows and store as

Code-

```
> price=c(52.00,54.75,57.50,57.50,59.75,62.50,64.75,67.25,67.50,69.75,70.00,75.50,77.50,78.00,81.25,82.50,86.25,87.50,88.00,92.00)

> F.A=c(1225,1230,1200,1000,1420,1450,1380,1510,1400,1550,1720,1700,1660,1800,1830,1790,2010,2000,2100,2240)

> Rooms=c(3,3,3,2,4,3,4,4,5,6,6,5,6,7,6,6,6,6,8,7)

> Age=c(6.2,7.5,4.2,4.8,1.9,5.2,6.5,9.2,0.0,5.7,7.3,4.5,6.8,0.7,5.6,2.3,6.7,3.4,5.6,3.4)

> C.H=c("YES","NO","NO","NO","YES","YES","NO","NO","NO","NO","YES","NO","YES","YES","YES","NO","YES","NO","YES","YES")

> df=data.frame(price,F.A,Rooms,Age,C.H)

> df

  price  F.A Rooms Age C.H
1  52.00 1225     3 6.2 YES
2  54.75 1230     3 7.5  NO
3  57.50 1200     3 4.2  NO
4  57.50 1000     2 4.8  NO
```

5	59.75	1420	4	1.9	YES
6	62.50	1450	3	5.2	YES
7	64.75	1380	4	6.5	NO
8	67.25	1510	4	9.2	NO
9	67.50	1400	5	0.0	NO
10	69.75	1550	6	5.7	NO
11	70.00	1720	6	7.3	YES
12	75.50	1700	5	4.5	NO
13	77.50	1660	6	6.8	YES
14	78.00	1800	7	0.7	YES
15	81.25	1830	6	5.6	YES
16	82.50	1790	6	2.3	NO
17	86.25	2010	6	6.7	YES
18	87.50	2000	6	3.4	NO
19	88.00	2100	8	5.6	YES
20	92.00	2240	7	3.4	YES

(a) How many rows are there in this table? How many columns are there?

```
> nrow(df)
[1] 20
> ncol(df)
[1] 5
```

(b) How to find the number of rows and number of columns by a single command?

nrow and ncol return the number of rows or columns present in x. NCOL and NROW do the same treating a vector as 1-column matrix.

```
nrow(x)
ncol(x)
NCOL(x)
NROW(x)
```

(c) What are the variables in the data file?

```
> ls()
[1] "Age"   "C.H"   "df"    "F.A"   "price" "Rooms"
```

(d) If the _le is very large, naturally we can not simply type 'a', because it will cover the entire screen and we won't be able to understand anything. So how to see the top or bottom few lines in
this file?

By using this command , we can see the first 8 lines from the top and last 8 lines from the bottom.

```
> df[1:8,]
  price  F.A Rooms Age C.H
1 52.00 1225     3 6.2 YES
2 54.75 1230     3 7.5  NO
3 57.50 1200     3 4.2  NO
4 57.50 1000     2 4.8  NO
5 59.75 1420     4 1.9 YES
6 62.50 1450     3 5.2 YES
7 64.75 1380     4 6.5  NO
8 67.25 1510     4 9.2  NO
> df[12:20,]
   price  F.A Rooms Age C.H
12 75.50 1700     5 4.5  NO
13 77.50 1660     6 6.8 YES
14 78.00 1800     7 0.7 YES
15 81.25 1830     6 5.6 YES
16 82.50 1790     6 2.3  NO
17 86.25 2010     6 6.7 YES
18 87.50 2000     6 3.4  NO
19 88.00 2100     8 5.6 YES
20 92.00 2240     7 3.4 YES
```

(e) If the number of columns is too large, again we may face the same problem. So how to see the first 5 rows and first three columns?

```
> df[1:5,1:3]
  price  F.A Rooms
1 52.00 1225     3
2 54.75 1230     3
3 57.50 1200     3
4 57.50 1000     2
5 59.75 1420     4
```

(f) How to get 1st, 3rd, 6th, and 10th row and 2nd, 4th, and 5th columns?

```
> df[c(1,3,6,10),c(2,4,5)]
    F.A Age C.H
1  1225 6.2 YES
3  1200 4.2  NO
6  1450 5.2 YES
10 1550 5.7  NO
```

(g) How to get values in a specific row or a column?

```
> df[14,]
   price  F.A Rooms Age C.H
14    78 1800     7 0.7 YES
> df[,3]
 [1] 3 3 3 2 4 3 4 4 5 6 6 5 6 7 6 6 6 6 8 7
```

Q.3) Calculate simple statistical measures using the values in the data file.

(a) Find means, medians, standard deviations of Price, Floor Area, Rooms, and Age.

```
mean(df[,1])

[1] 71.5875

> median(df[,1])

[1] 69.875

> sd(df[,1])

[1] 12.21094

> mean(df[,2])

[1] 1610.75

> median(df[,2])

[1] 1605

> sd(df[,2])

[1] 331.9649

> mean(df[,3])
```

```
[1] 5
> median(df[,3])
[1] 5.5
> sd(df[,3])
[1] 1.65434
> mean(df[,4])
[1] 4.875
> median(df[,4])
[1] 5.4
> sd(df[,4])
[1] 2.366182
```

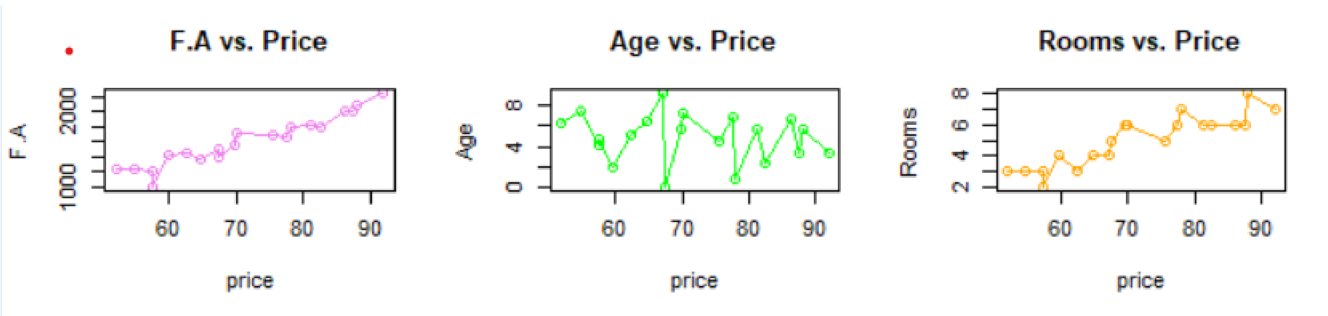
(b) How many houses have central heating and how many don't have?

```
> K=subset(df,df$C.H=='YES')
> nrow(K)
[1] 10
> J=subset(df,df$C.H=='NO')
> nrow(J)
[1] 10
```

(c) Plot Price vs. Floor, Price vs. Age, and Price vs. Rooms, in separate graphs.

```
> par(mfrow=c(4,3))
> plot(price,F.A,type = "o",col="violet",main="F.A vs. Price")
> plot(price,Age,type="o",col="green",main="Age vs. Price")
```

> plot(price,Rooms,type="o",col="orange",main="Rooms vs. Price")

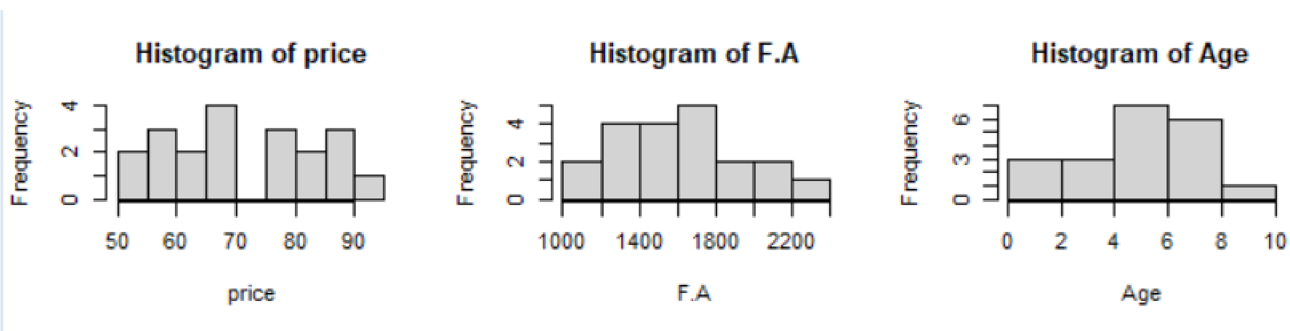


(d) Draw histograms of Prices, Floor Area, and Age.

> hist(price)

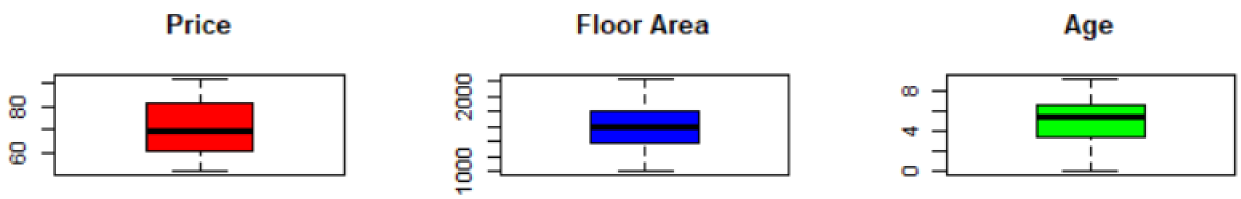
> hist(F.A)

> hist(Age)



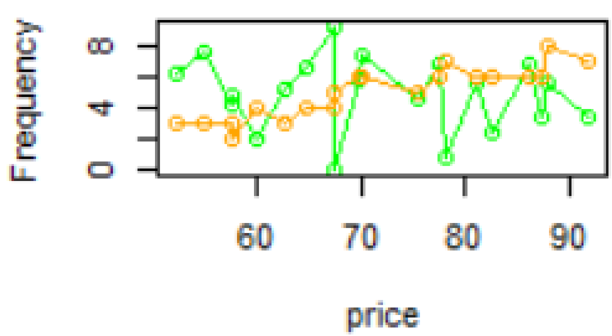
(e) Draw box plots of Price, Floor Area, and Age.

```
> boxplot(price,col=c("red"),main="Price")

> boxplot(F.A,col=c("blue"),main="Floor Area")

> boxplot(Age,col=c("green"),main="Age")
```



(f) Draw all the graphs in (c), (d), and (e) in the same graph paper.

```
> plot(price,Age,type = "o",col="green",ylab="Frequency")

> lines(price,F.A,type="o",col="violet")

> lines(price,Rooms,type="o",col="orange")
```



NAME- VARANASI KASYAP

REGISTRATION NUMBER-20BCE7315

Course Title: Applied Statistics

Instructor's name: Dr. Ankur

Course Code: MAT1011

Q.1)

(a)Create a data list (4,4,4,4,3,3,3,5,5,5) using 'rep' function.

```
> rep(c(4, 3, 5), times = c(4,3,3))
 [1] 4 4 4 4 3 3 3 5 5 5
```

(b)Create a list (4,6,3,4,6,3, . . . ,4,6,3) where there 10 occurrences of 4,6, and 3 in the given order.

```
> rep(c(4, 6, 3), times = 10)
 [1] 4 6 3 4 6 3 4 6 3 4 6 3 4 6 3 4 6 3 4 6 3 4 6 3 4 6 3 4 6 3
```

(c)Create a list (3,1,5,3,2,3,4,5,7,7,7,7,7,7,6,5,4,3,2,1,34,21,54) using one line command.

```
> a <- c(3,1,5,3,2:5);b <- rep(7, 6);c <- c(6:1,34,21,54);d <- c(a,b,c);d
 [1]  3  1  5  3  2  3  4  5  7  7  7  7  7  7  6  5  4  3  2  1 34 21 54
```

(d)First create a list (2,1,3,4). Then append this list at the end with another list (5,7,12,6,−8).Check whether the number of elements in the augmented list is 11.

```
> x<-c(2,1,3,4)
> D= append(x,c(5,7,12,6,(-8)))
> D
[1]  2  1  3  4  5  7 12  6 -8
```

No, the number of elements in the augmented list are 9 because we added five elements to the existing four elements. So there are nine elements.

Q.2)

(a)Print all numbers starting with 3 and ending with 7 with an increment of 0 : 0.5. Storethese numbers in x.

```
> X=seq(3,7, by=0.5)
> X
[1] 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 7.0
```

(b)Print all even numbers between 2 and 14 (both inclusive).

```
> even<-seq(2,14,2)
> even
[1]  2  4  6  8 10 12 14
```

(c)Type 2∗xand see what you get. Each element ofxis multiplied by 2.

```
> 2*x
[1]  6  7  8  9 10 11 12 13 14
```

Q.3) Collect atleast 75 students list and analyse the data by using descriptive statistics and interpret the results.

```
> getwd()
[1] "C:/Users/HP/OneDrive/Documents"
> setwd("C:/Users/HP/OneDrive/Documents")
> A=read.csv("list.csv")
> A
```

	Stdent.Name	Registration.number	Maths	Physics	Chemistry	Java	English
1	A	20001234	10	11	12	10	12
2	B	20001235	11	9	12	12	9
3	C	20001236	12	8	9	13	11
4	D	20001237	13	11	11	15	15
5	E	20001238	14	13	15	11	14
6	F	20001239	12	14	14	13	11
7	G	20001240	9	11	11	15	13
8	H	20001241	8	10	13	13	10
9	I	20001242	14	15	10	14	11

10	J	20001243	15	12	11	12	15
11	K	20001244	13	11	15	9	11
12	L	20001245	15	13	11	8	12
13	M	20001246	11	11	12	14	15
14	N	20001247	14	9	15	15	13
15	O	20001248	13	8	13	13	14
16	P	20001249	11	11	14	15	15
17	Q	20001250	13	13	15	11	6
18	R	20001251	14	14	6	14	8
19	S	20001252	12	11	8	13	13
20	T	20001253	9	10	13	11	11
21	U	20001254	11	15	11	13	13
22	V	20001255	15	12	13	14	14
23	W	20001256	14	11	14	12	12
24	X	20001257	11	13	12	9	9
25	Y	20001258	13	14	9	11	11
26	Z	20001259	10	14	11	15	15
27	AA	20001260	11	11	15	14	14
28	BB	20001261	10	10	14	11	11
29	CC	20001262	15	15	11	13	15
30	DD	20001263	12	12	13	10	14
31	EE	20001264	11	12	10	11	12
32	FF	20001265	9	9	11	10	9
33	GG	20001266	8	11	10	15	8
34	HH	20001267	11	15	15	13	14
35	II	20001268	13	14	12	14	15
36	JJ	20001269	14	11	15	14	13
37	KK	20001270	11	13	12	13	15
38	LL	20001271	10	10	11	14	11

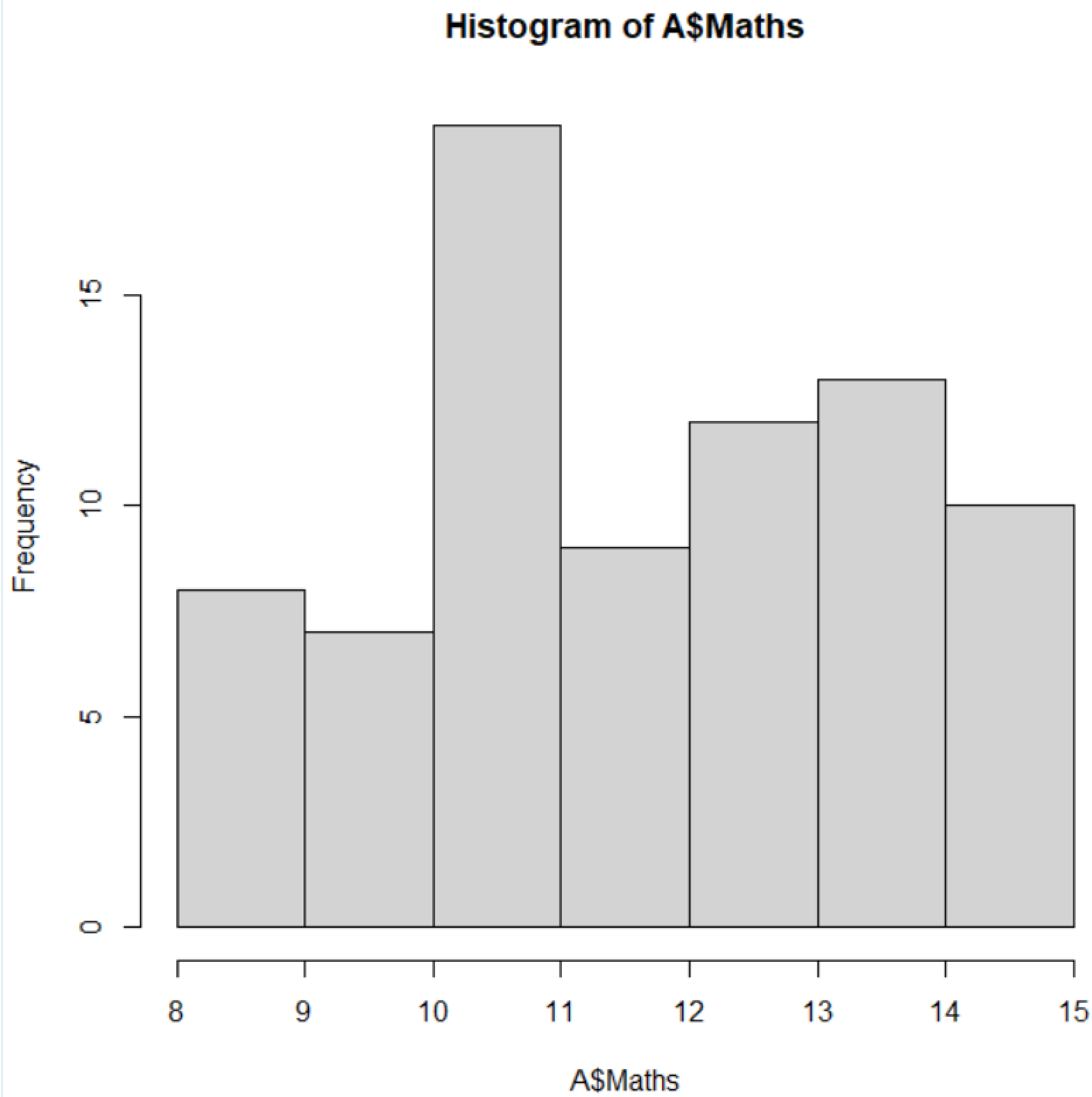
39	MM	20001272	15	11	13	12	14
40	NN	20001273	12	15	14	9	13
41	OO	20001274	11	11	14	11	11
42	PP	20001275	13	12	13	15	13
43	QQ	20001276	14	15	14	14	14
44	RR	20001277	14	13	12	11	12
45	SS	20001278	11	14	9	13	9
46	TT	20001279	10	15	11	10	11
47	UU	20001280	15	6	15	11	15
48	VV	20001281	12	8	14	15	14
49	WW	20001282	11	13	11	11	11
50	XX	20001283	9	11	13	12	13
51	YY	20001284	8	13	10	15	10
52	ZZ	20001285	11	14	11	13	11
53	AAA	20001286	13	12	11	15	10
54	BBB	20001287	14	9	15	11	15
55	CCC	20001288	11	11	11	12	13
56	DDD	20001289	10	15	12	9	15
57	EEE	20001290	15	14	15	11	11
58	FFF	20001291	12	11	13	15	12
59	GGG	20001292	11	13	14	14	11
60	HHH	20001293	13	10	15	11	13
61	III	20001294	14	11	6	13	12
62	JJJ	20001295	14	10	8	10	10
63	KKK	20001296	13	15	13	11	9
64	LLL	20001297	14	12	11	15	8
65	MMM	20001298	12	9	13	11	14
66	NNN	20001299	9	11	14	12	11
67	OOO	20001300	11	12	12	15	13

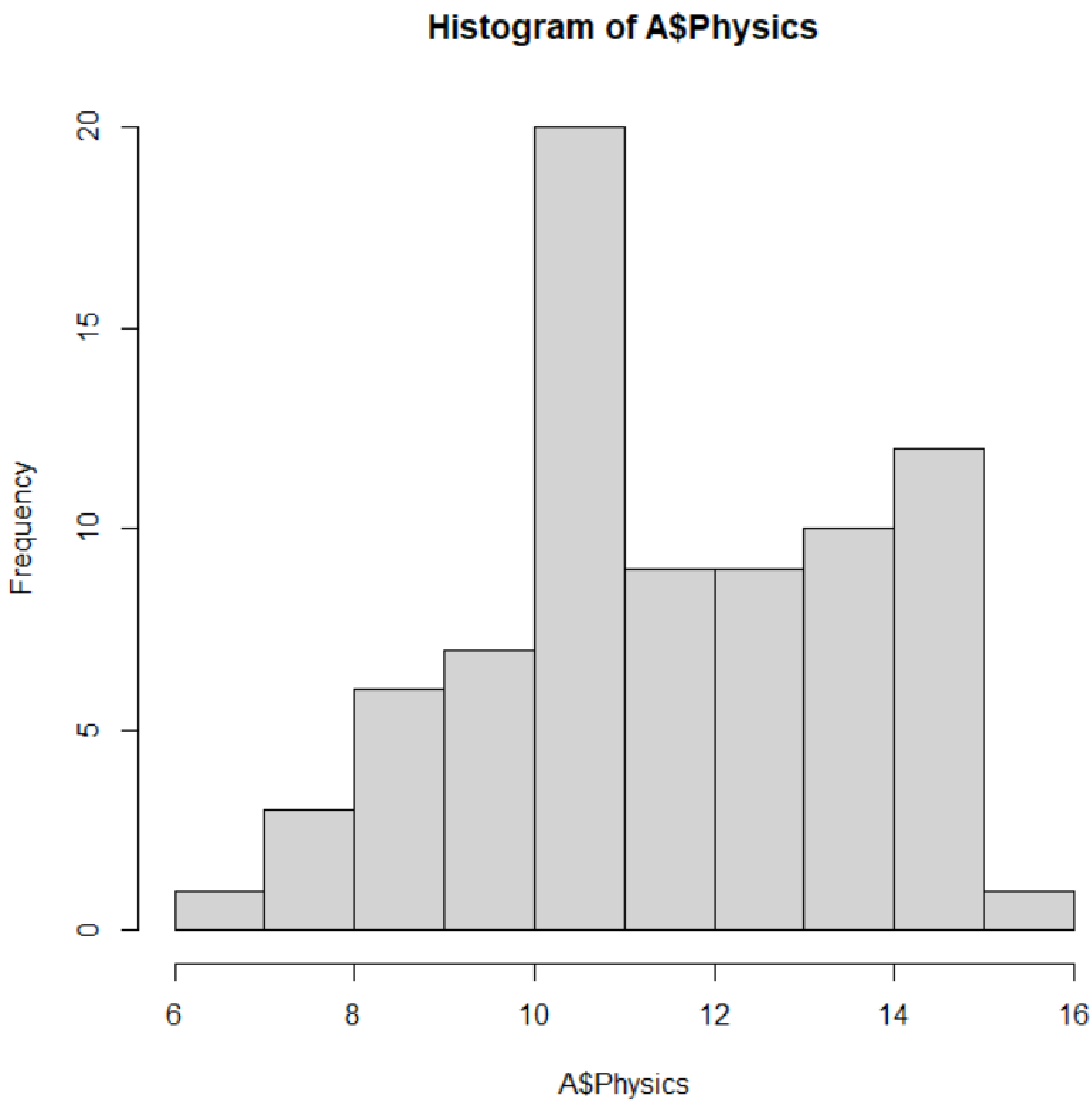
68	PPP	20001301	15	14	9	13	10
69	QQQ	20001302	14	15	11	14	11
70	RRR	20001303	11	16	15	15	15
71	SSS	20001304	13	15	14	6	11
72	TTT	20001305	10	12	11	8	12
73	UUU	20001306	11	11	13	13	15
74	VVV	20001307	15	10	10	11	11
75	WWW	20001308	11	9	11	13	13
76	XXX	20001309	12	11	10	14	14
77	YYY	20001310	15	15	15	12	12
78	ZZZ	20001311	13	14	12	9	9

```
> mean(A[,3])
[1] 12.10256
> median(A[,3])
[1] 12
> sd(A[,3])
[1] 1.951288
> mean(A[,4])
[1] 11.98718
> median(A[,4])
[1] 12
> sd(A[,4])
[1] 2.159206
> mean(A[,5])
[1] 12.14103
> median(A[,5])
[1] 12
> sd(A[,5])
[1] 2.142485
```

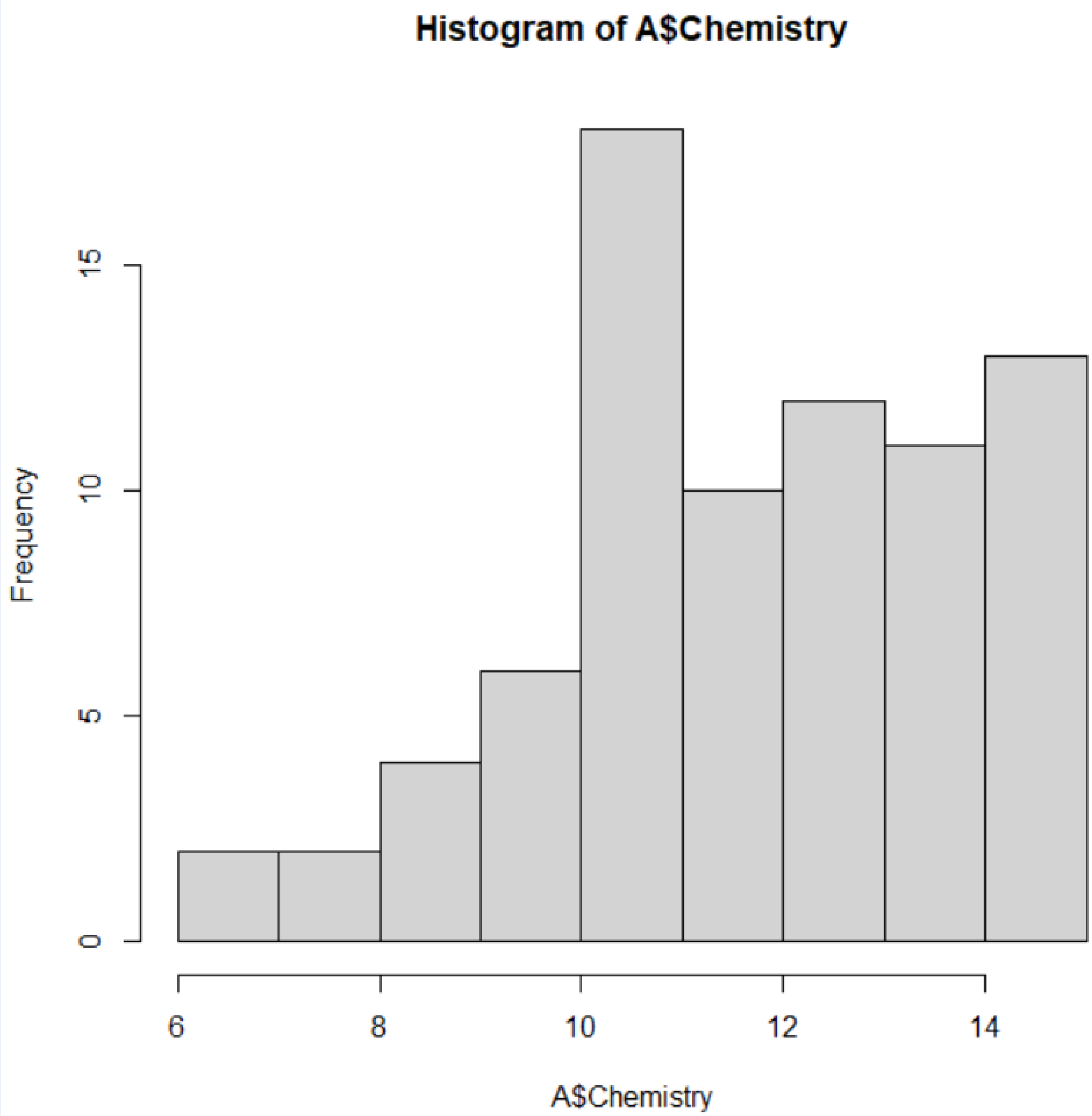
```
> mean(A[,6])
[1] 12.33333
> median(A[,6])
[1] 13
> sd(A[,6])
[1] 2.068105
> mean(A[,7])
[1] 12.11538
> median(A[,7])
[1] 12
> sd(A[,7])
[1] 2.12579
```

```
> hist(A$Math)
> hist(A$Physics)
> hist(A$Chemistry)
> hist(A$Java)
> hist(A$English)
```

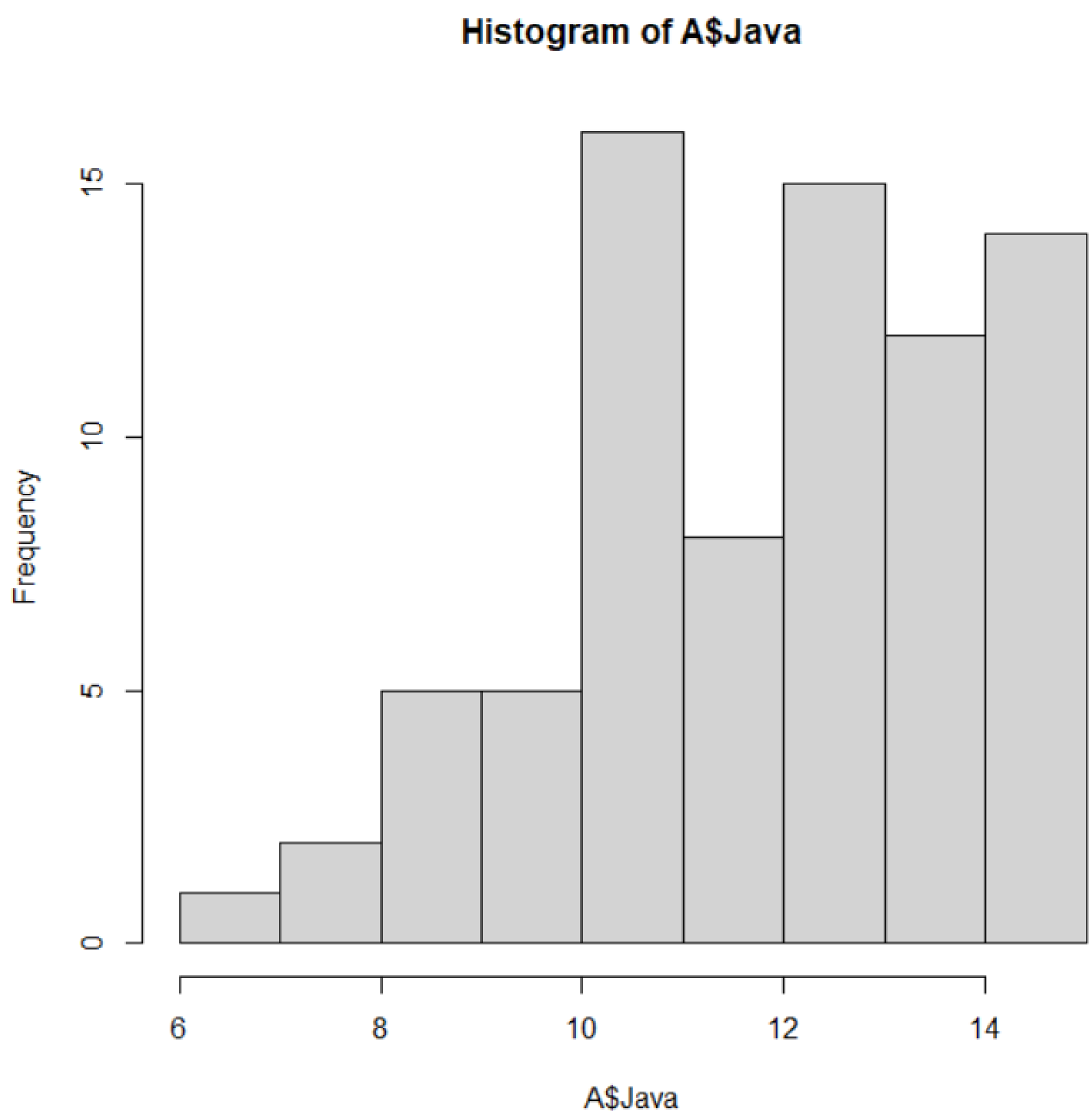

Histogram of A$Maths
Frequency
0
5
10
15
8
9
10
11
12
13
14
15
A$Maths

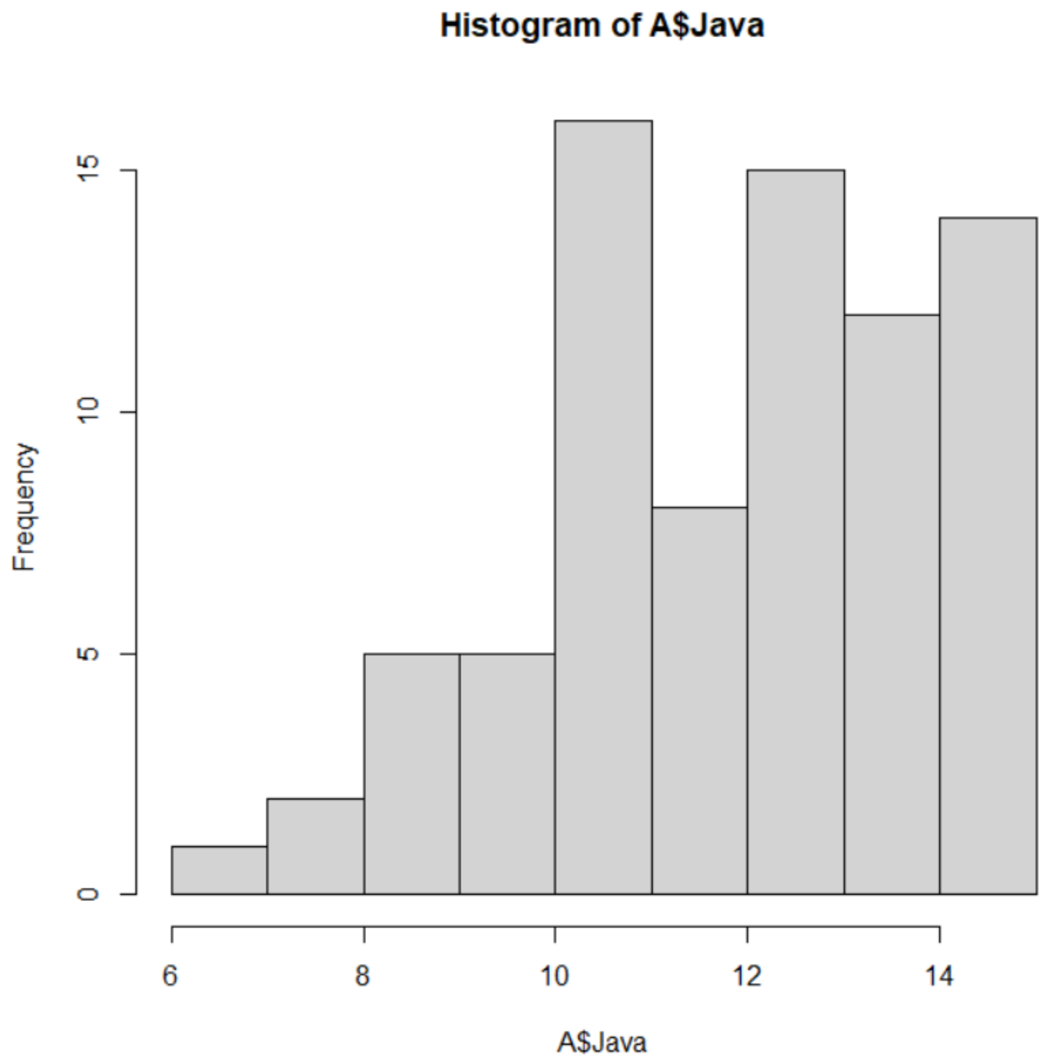


Histogram of A$Physics
Frequency
0
5
10
15
20
6
8
10
12
14
16
A$Physics

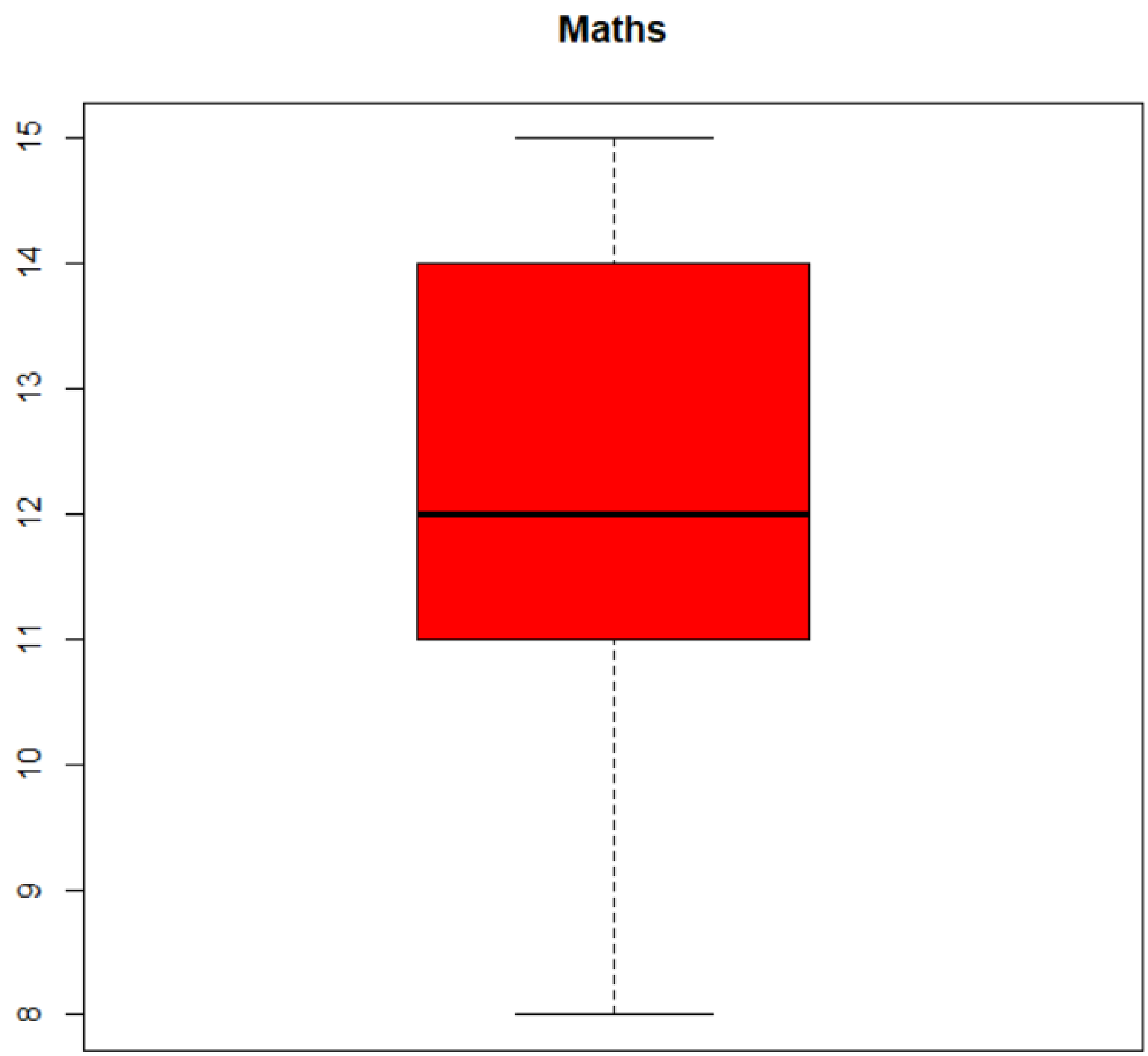


Histogram of A$Chemistry
Frequency
0
5
10
15
6
8
10
12
14
A$Chemistry

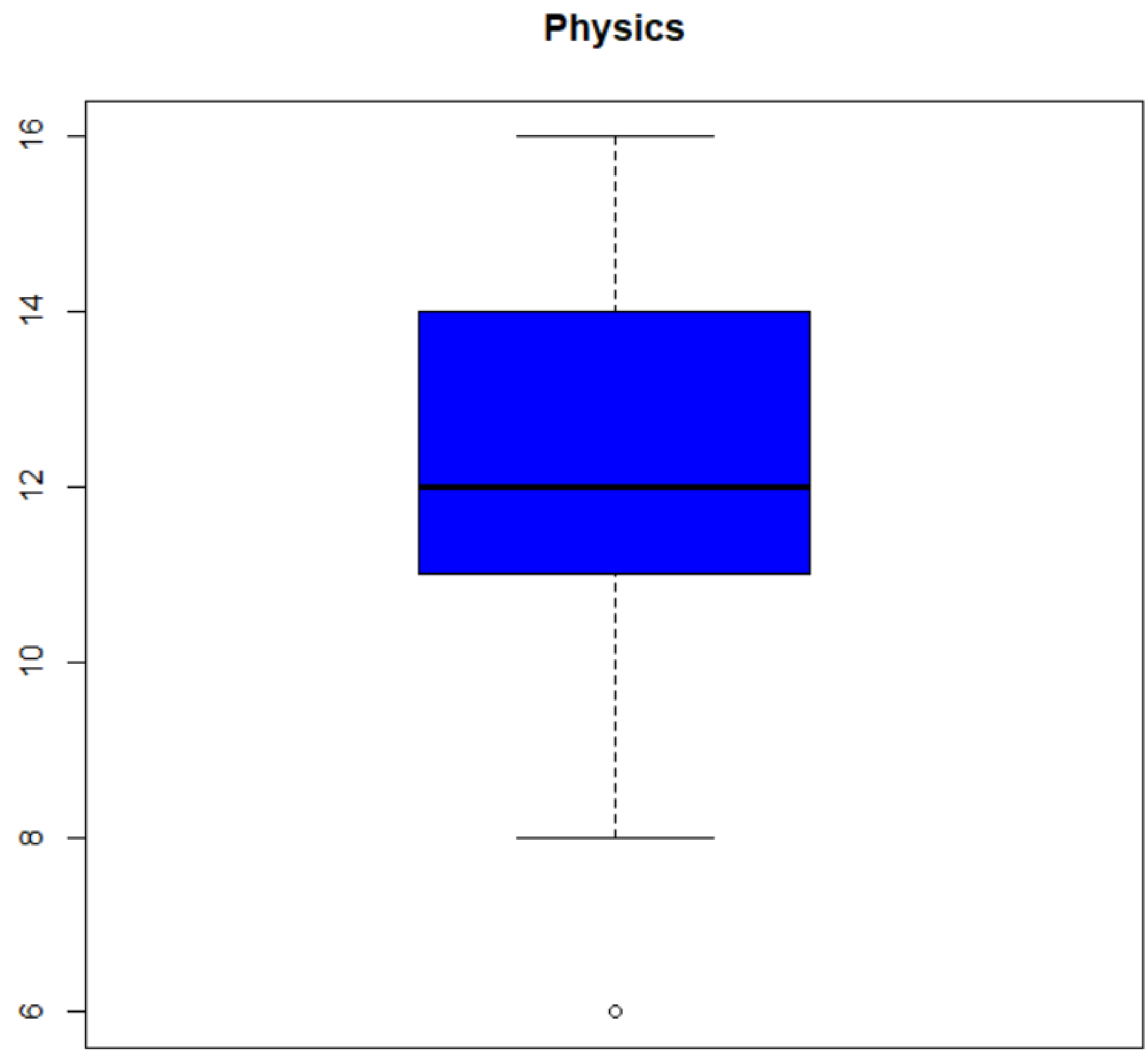

Histogram of A$Java
Frequency
0
5
10
15
6
8
10
12
14
A$Java

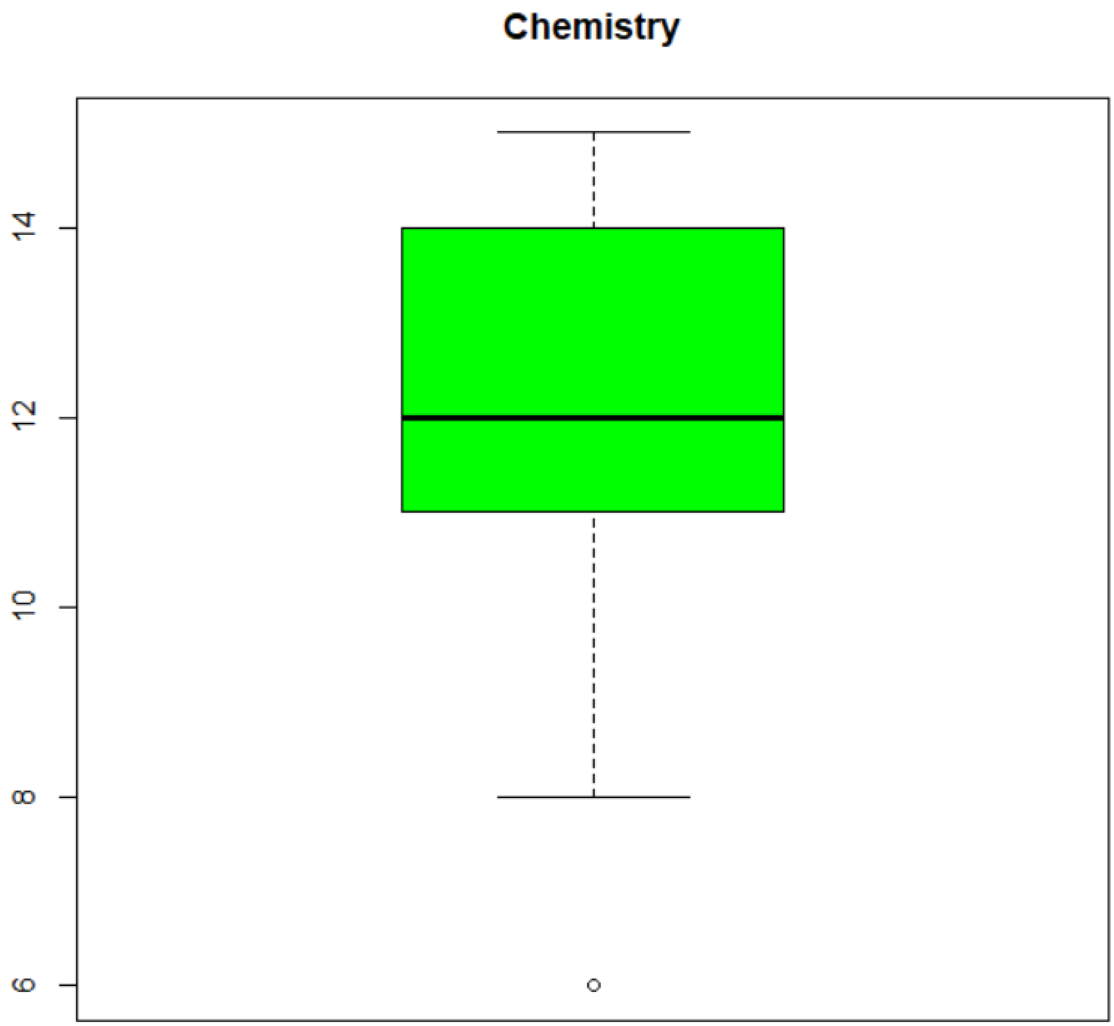


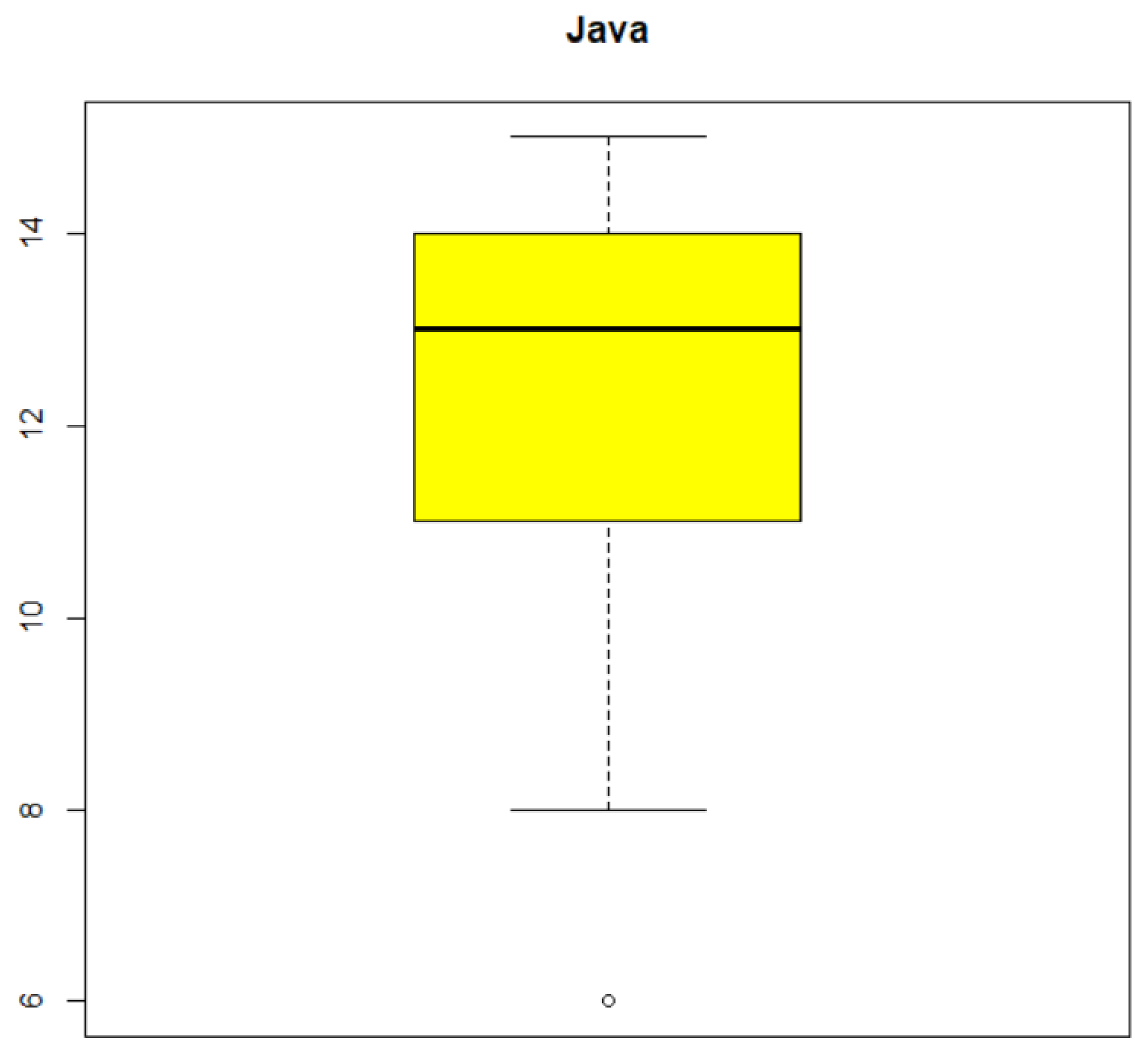

```
> boxplot(A$Maths,col=c("red"),main="Maths")
> boxplot(A$Physics,col=c("blue"),main="Physics")
> boxplot(A$Chemistry,col=c("green"),main="Chemistry")
> boxplot(A$Java,col=c("yellow"),main="Java")
> boxplot(A$English,col=c("orange"),main="English")
```



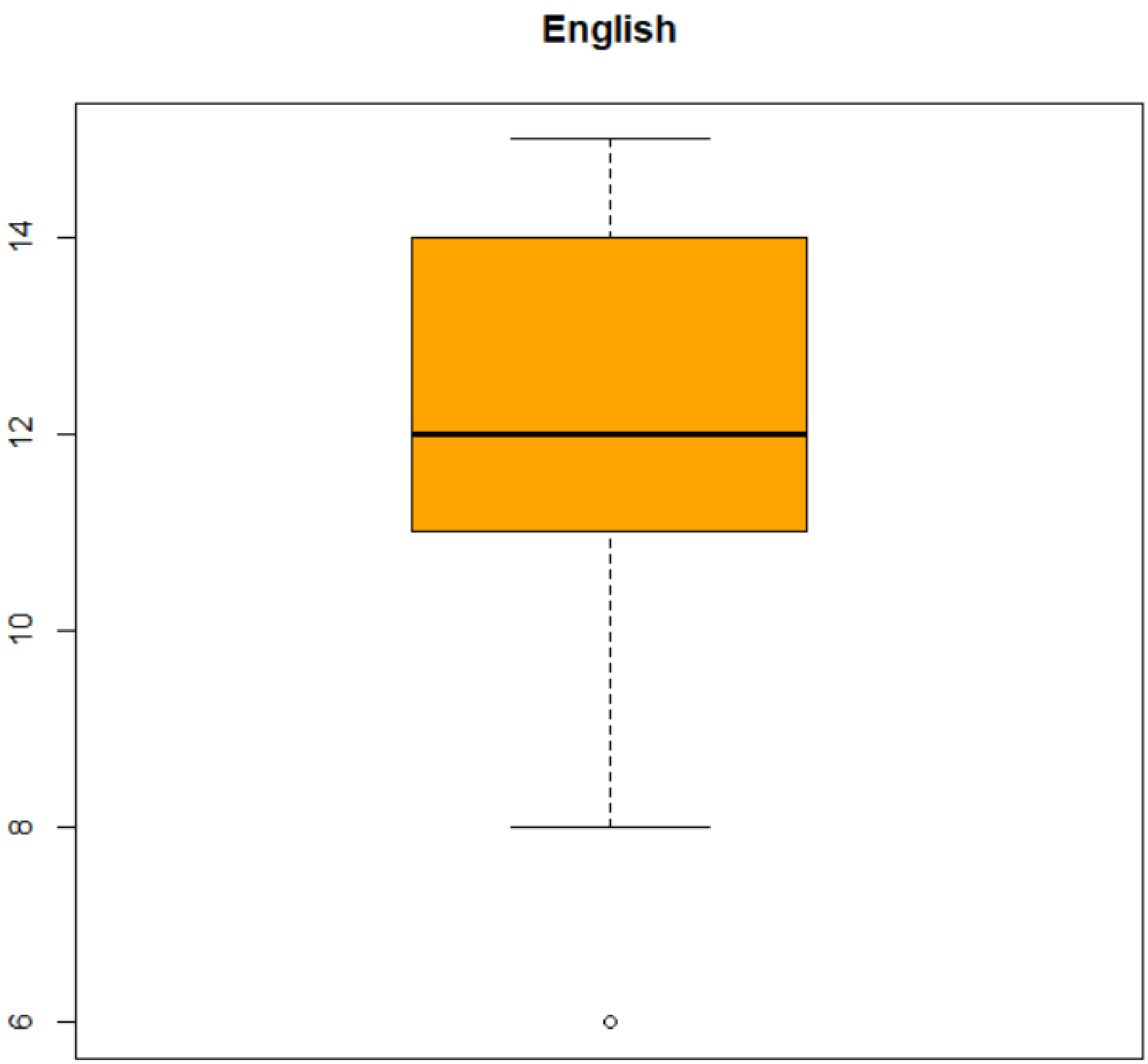
Maths
15
14
13
12
11
10
9
8


Physics
16
14
12
10
8
6

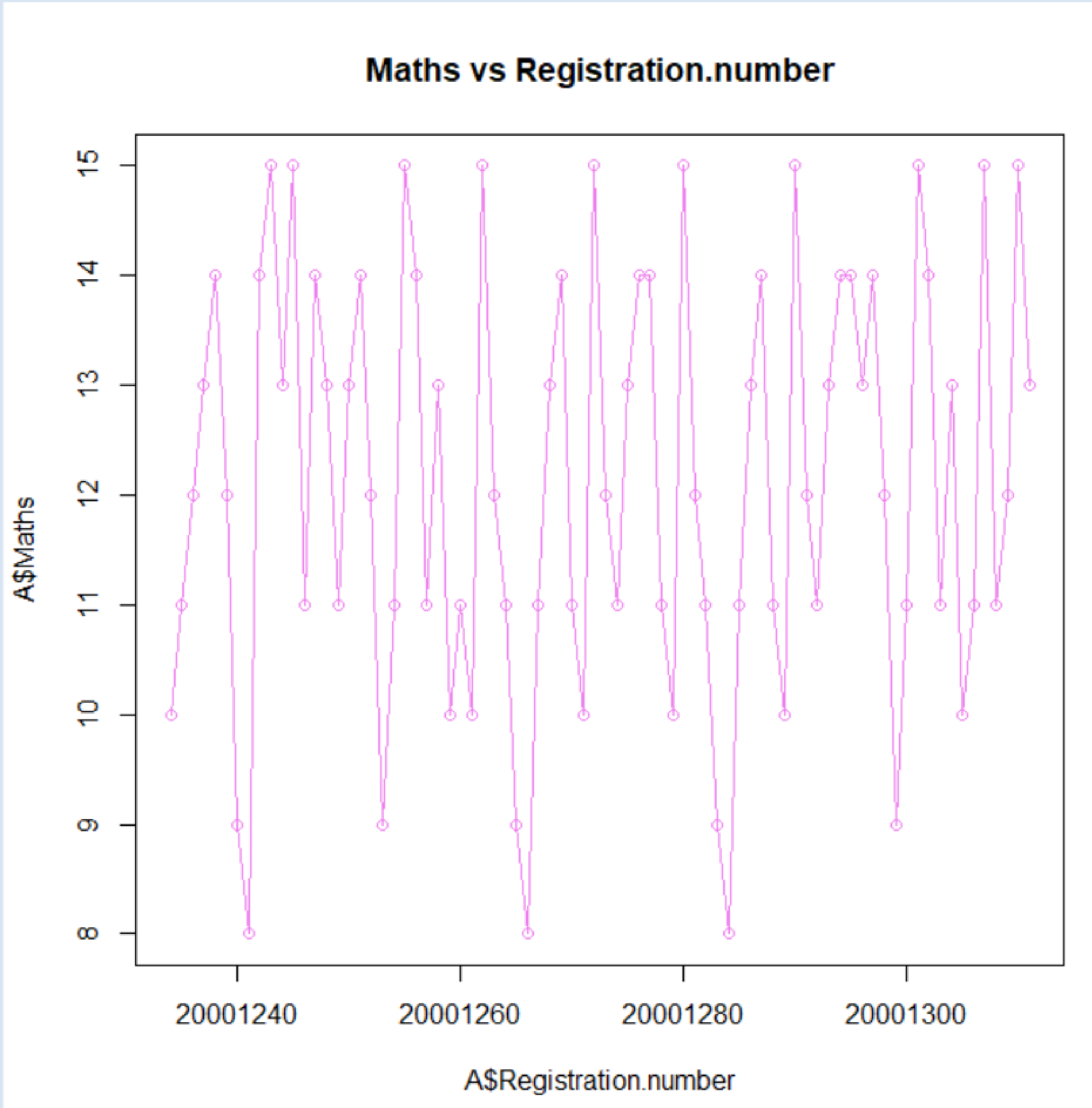

Chemistry
14
12
10
8
6

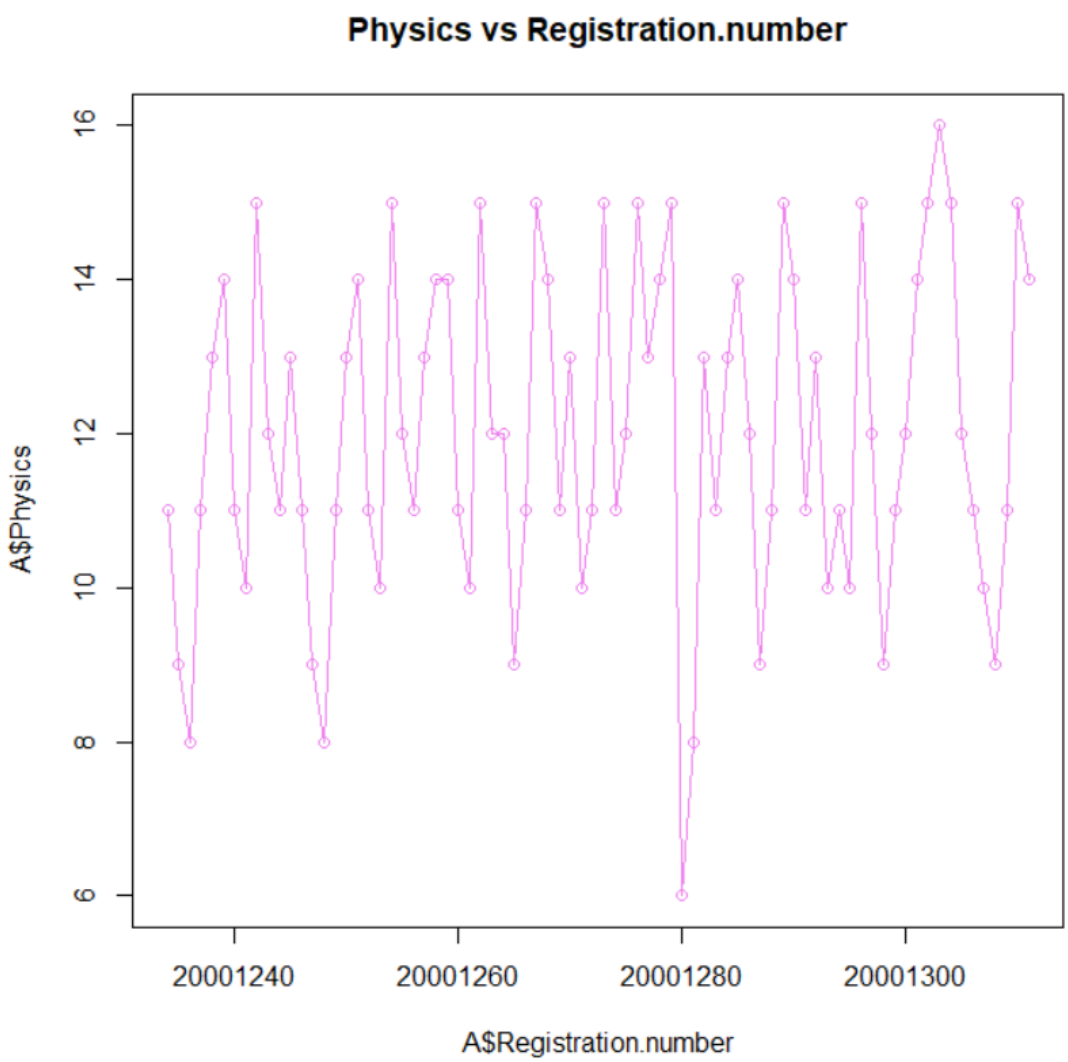


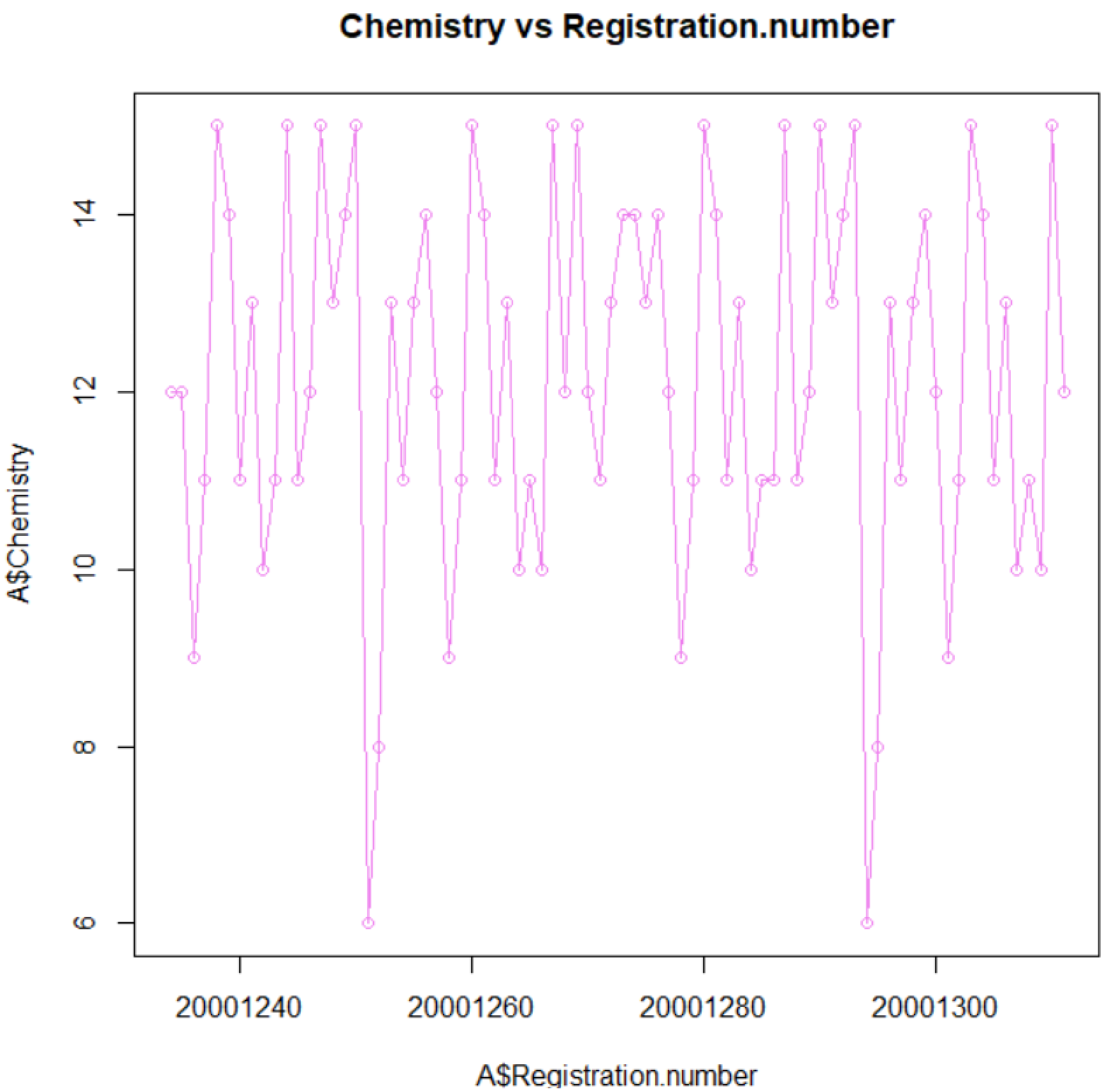
Java
14
12
10
8
6

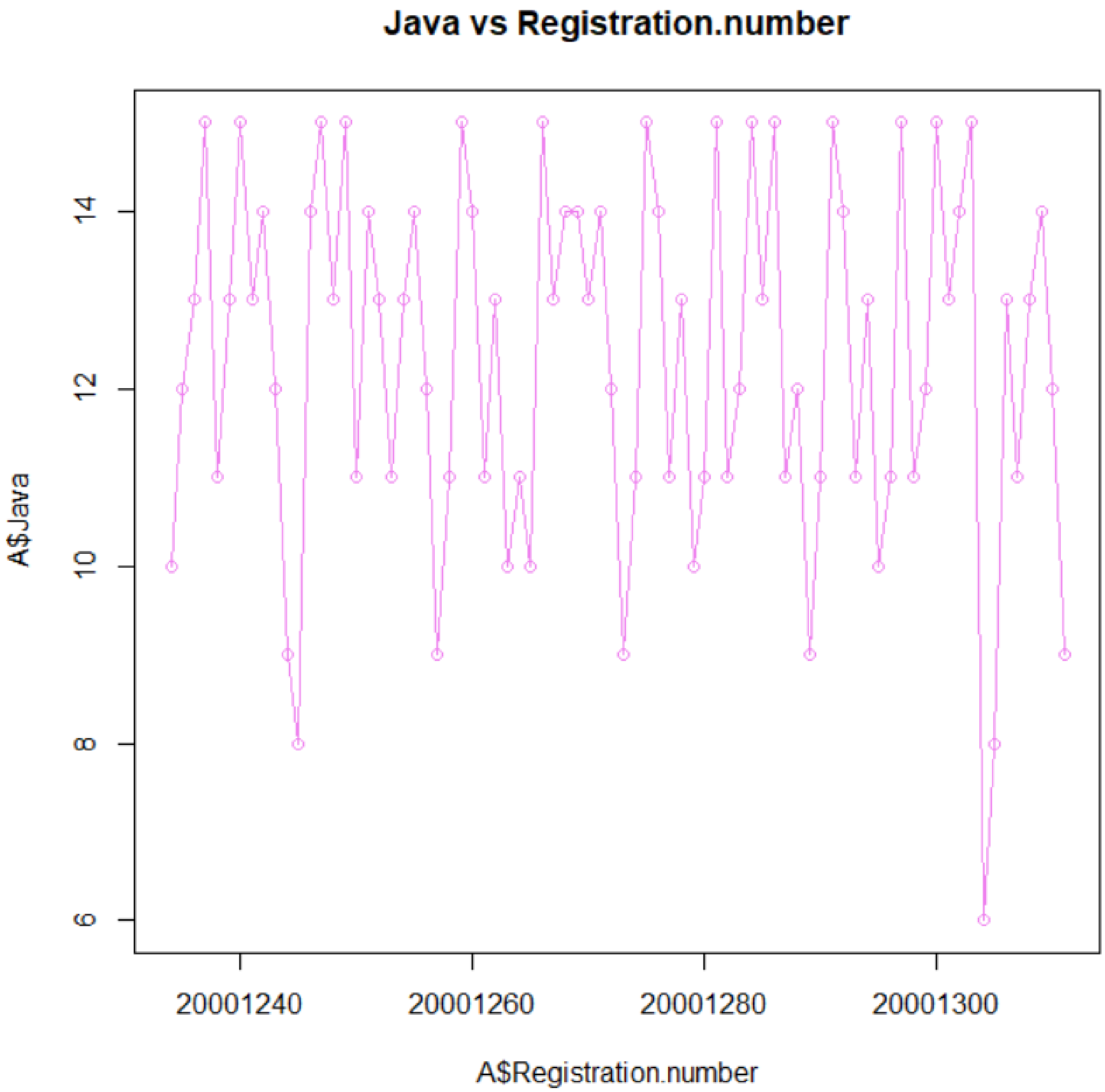



```
> plot(A$Registration.number,A$Maths,type = "o",col="violet",main="Maths vs Registration.number")

> plot(A$Registration.number,A$Physics,type = "o",col="violet",main="Physics vs Registration.number")

> plot(A$Registration.number,A$Chemistry,type = "o",col="violet",main="Chemistry vs Registration.number")

> plot(A$Registration.number,A$Java,type = "o",col="violet",main="Java vs Registration.number")

> plot(A$Registration.number,A$English,type = "o",col="violet",main="English vs Registration.number")
```

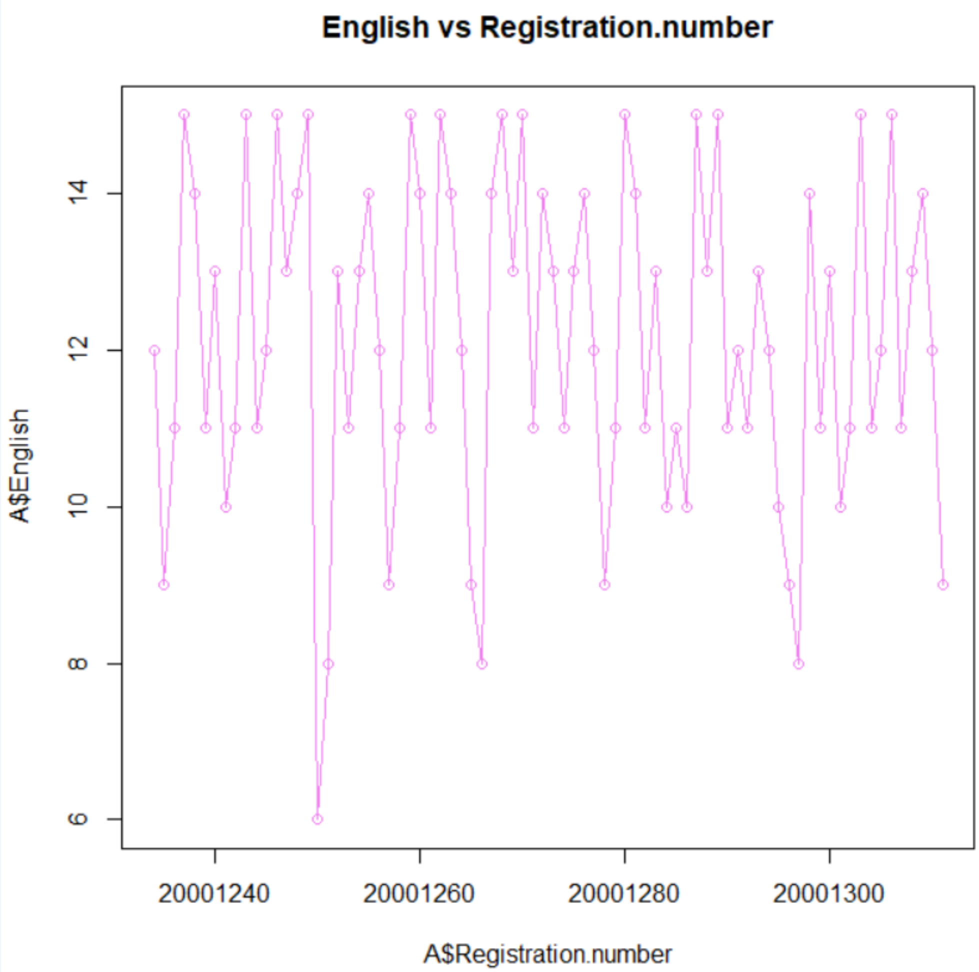

Maths vs Registration.number
A$Maths
8
9
10
11
12
13
14
15
20001240
20001260
20001280
20001300
A$Registration.number


Physics vs Registration.number
A$Physics
6
8
10
12
14
16
20001240
20001260
20001280
20001300
A$Registration.number


Chemistry vs Registration.number
A$Chemistry
6
8
10
12
14
20001240
20001260
20001280
20001300
A$Registration.number

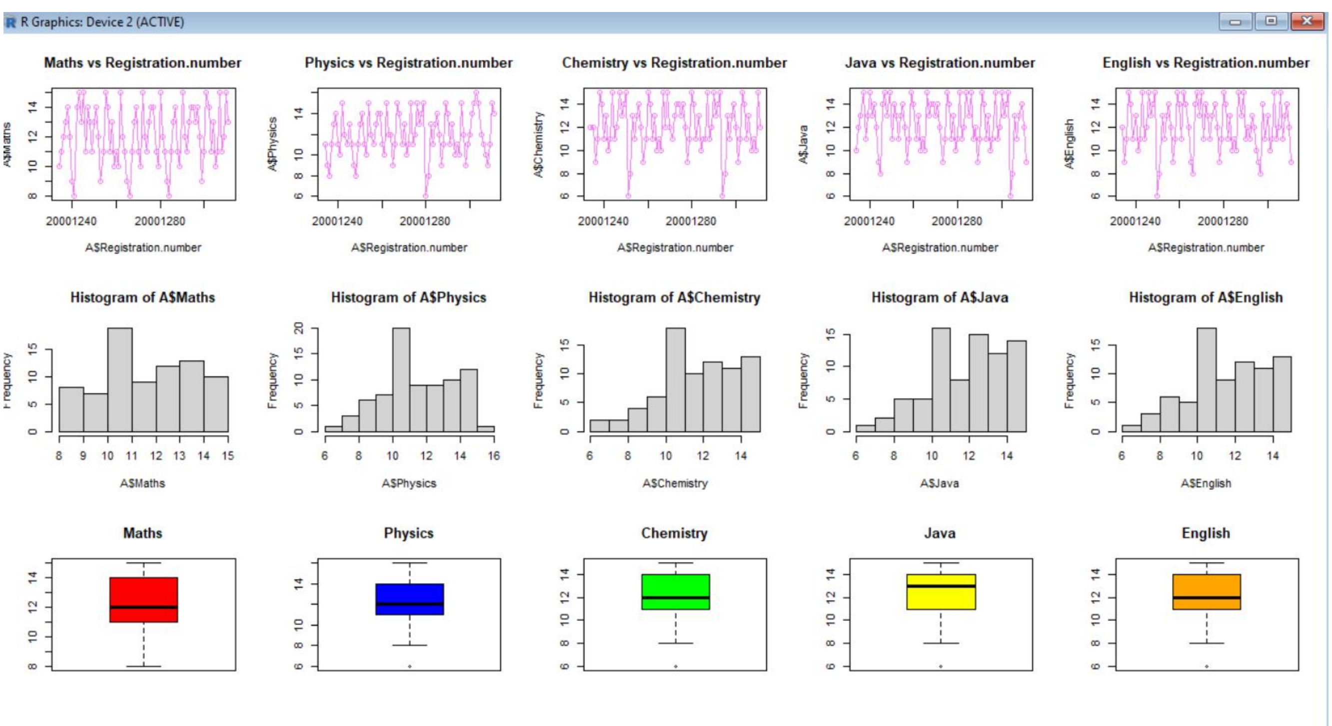


Java vs Registration.number
A$Java
A$Registration.number
6
8
10
12
14
20001240
20001260
20001280
20001300




```
> par(mfrow=c(3,5))
> plot(A$Registration.number,A$Maths,type = "o",col="violet",main="Maths vs Registration.number")
> plot(A$Registration.number,A$Physics,type = "o",col="violet",main="Physics vs Registration.number")
> plot(A$Registration.number,A$Chemistry,type = "o",col="violet",main="Chemistry vs Registration.number")
```

```
> plot(A$Registration.number,A$Java,type = "o",col="violet",main="Java vs
Registration.number")

> plot(A$Registration.number,A$English,type = "o",col="violet",main="English vs
Registration.number")

> hist(A$Maths)

> hist(A$Physics)

> hist(A$Chemistry)

> hist(A$Java)

> hist(A$English)

> boxplot(A$Maths,col=c("red"),main="Maths")

> boxplot(A$Physics,col=c("blue"),main="Physics")

> boxplot(A$Chemistry,col=c("green"),main="Chemistry")

> boxplot(A$Java,col=c("yellow"),main="Java")

> boxplot(A$English,col=c("orange"),main="English")
```



NAME- VARANASI KASYAP

REGISTRATION NUMBER-20BCE7315

Course Title: Applied Statistics

Instructor's name: Dr. Ankur

Course Code: MAT1011

Problem Set-III

Q.1) Matrices and arrays:

(a)Matrices and arrays are represented as vectors with dimensions: Create one matrix with 1 to 12 numbers with 3×4 order.

```
> x<-rbind(c(1:4),c(5:8),c(9:12))
> x
     [,1] [,2] [,3] [,4]
[1,]    1    2    3    4
[2,]    5    6    7    8
[3,]    9   10   11   12
```

(b)Create same matrix with matrix function.

```
> x<-matrix(1:12,nrow=3,ncol=4,byrow = TRUE)
> x
     [,1] [,2] [,3] [,4]
[1,]    1    2    3    4
[2,]    5    6    7    8
[3,]    9   10   11   12
```

(c)Give name of rows of this matrix with A,B,C.

```
> rownames(x)<-c("A","B","C")
> x
  [,1] [,2] [,3] [,4]
A    1    2    3    4
B    5    6    7    8
C    9   10   11   12
```

(d)Transpose of the matrix.

```
> t(x)
     A B  C
[1,] 1 5  9
[2,] 2 6 10
[3,] 3 7 11
```

```
[4,] 4 8 12
```

(e)Use functionscbindandrbindseparately to create different matrices.

```
> cbind(x,c(99,98,97))
  [,1] [,2] [,3] [,4] [,5]
A    1    2    3    4   99
B    5    6    7    8   98
c    9   10   11   12   97
> rbind(x,c(99,98,97,96))
  [,1] [,2] [,3] [,4]
A    1    2    3    4
B    5    6    7    8
c    9   10   11   12
    99   98   97   96
```

(f)Use arbitrary numbers to create matrix.

```
> y<-sample(1:100,12)
> z<-matrix(y,nrow=3,ncol=4,byrow=TRUE)
> z
     [,1] [,2] [,3] [,4]
[1,]   66   26   17    5
[2,]   74   37   34   49
[3,]   15   40   61   45
```

(g)Verify matrix multiplication.

```
> x*z
  [,1] [,2] [,3] [,4]
A   66   52   51   20
B  370  222  238  392
```

c 135 400 671 540

Q.2) Random sampling

(a)In R you can simulate these situations with the sample function. Pick five numbers at random from the set 1 : 40.

```
> x<-1:40
> x<-sample(x,5)
> x
[1] 13 24 29 28 32
```

(b)Notice that the default behaviour of sample is sampling without replacement. That is the samples will not contain the same number twice, and obviously can not be bigger than the length of the vector to be sampled. If you want sampling with replacement, then you need to add the argument replace=T RU E. Sampling with replacement is suitable for modelling coin tosses or throws of a die. So, for instance, simulate 10-coin tosses.

```
> x<-c("HEAD","TAIL")
> x<-sample(x,10,replace=TRUE)
> x
 [1] "HEAD" "TAIL" "TAIL" "HEAD" "HEAD" "TAIL" "TAIL" "HEAD" "TAIL"
"TAIL"
```

(c)In fair coin-tossing, the probability of heads should equal the probability of tails, but the idea of a random event is not restricted to symmetric cases. It could be equally well applied to other cases, such as the successful outcome of a surgical procedure. Hopefully there would be a better than 50% chance of this. Simulate data with non-equal probabilities for the outcomes (say, a 90%chance of success) by using the prob argument to sample.

```
> x<-c("HEAD","TAIL")
> x<-sample(x,10,replace=TRUE,prob<-c(10,30))
```

```
> x
 [1] "TAIL" "TAIL" "HEAD" "TAIL" "TAIL" "TAIL" "TAIL" "TAIL" "HEAD"
"TAIL"
```

(d)The choose function can be used to calculate the following expression.(405)=40!5!35!.

```
> choose(40,5)
[1] 658008
```

(e)Find 5!

```
> factorial(5)
[1] 120
```

NAME- VARANASI KASYAP

REGISTRATION NUMBER-20BCE7315

Course Title: Applied Statistics

Instructor's name: Dr. Ankur

Course Code: MAT1011

Problem Set-IV

Q.1) Five terminals on an on-line computer system are attached to a communication line to the central computer system. The probability that any terminal is ready to transmit is 0.95. Let X denote the number of ready terminals.

(a)Find the probability of getting exactly 3 ready terminals.

```
n=5
p=0.95
 dbinom(3,n,p)
[1] 0.02143438
```

(b)Find all the probabilities.

```
n=5
 p=0.95
 for(x in 0:n)
```

```
 {
v=dbinom(x,n,p)
 print(v)
 }
[1] 3.125e-07 // for 0
[1] 2.96875e-05 // for 1
[1] 0.001128125
[1] 0.02143438
[1] 0.2036266
[1] 0.7737809
```

Q.2) It is known that 20% of integrated circuit chips on a production line are defective. To maintained monitor the quality of the chips, a sample of twenty chips is selected at regular intervals forinspection. Let X denote the number of defectives found in the sample. Find the probability of different number of defectives found in the sample?

```
n=20
p=0.2
for(x in 0:n)
{
 v=dbinom(x,n,p)
 print(v)
}
 [1] 0.01152922
 [1] 0.05764608
 [1] 0.1369094
 [1] 0.2053641
 [1] 0.2181994
 [1] 0.1745595
 [1] 0.1090997
 [1] 0.05454985
```

```
[1] 0.02216088
[1] 0.007386959
[1] 0.002031414
[1] 0.0004616849
[1] 8.656592e-05
[1] 1.331783e-05
[1] 1.664729e-06
[1] 1.664729e-07
[1] 1.30057e-08
[1] 7.65041e-10
[1] 3.187671e-11
[1] 8.388608e-13
[1] 1.048576e-14
```

Q.3) It is known that 1% of bits transmitted through a digital transmission are received in error.One hundred bits are transmitted each day. Find the probability of different number of bits foundin error each day.?

```
n=100
 p=0.01
for(x in 0:n)
{
 v=dbinom(x,n,p)
 print(v)
 }
```

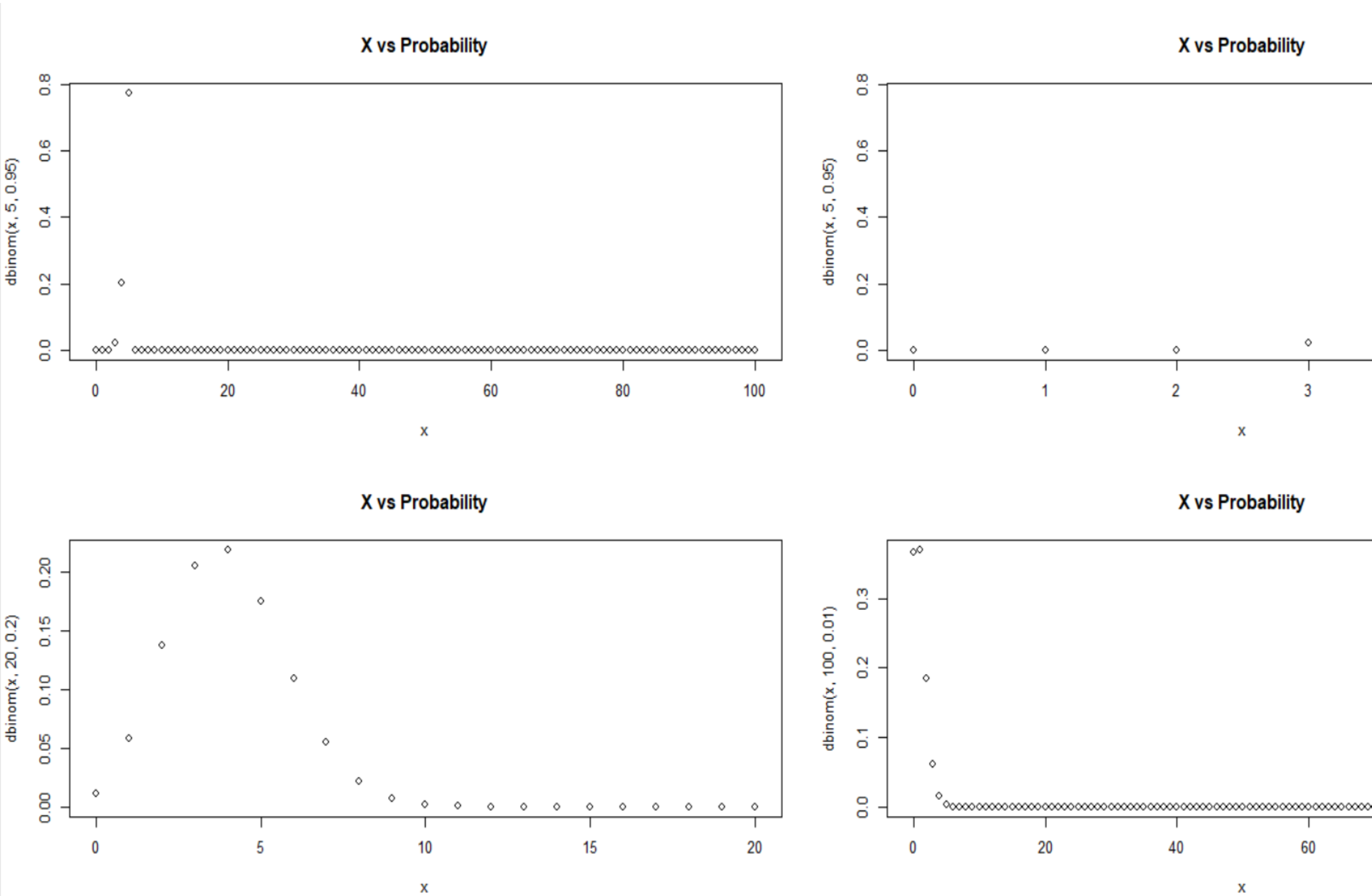
```
[1] 0.3660323
[1] 0.3697296
[1] 0.1848648
[1] 0.06099917
[1] 0.01494171
[1] 0.002897787
[1] 0.0004634508
```

```
[1] 6.286346e-05
[1] 7.381694e-06
[1] 7.621951e-07
[1] 7.006036e-08
[1] 5.790112e-09
[1] 4.33771e-10
[1] 2.965956e-11
[1] 1.861747e-12
[1] 1.078184e-13
[1] 5.785707e-15
[1] 2.887697e-16
[1] 1.344999e-17
[1] 5.863367e-19
[1] 2.39865e-20
[1] 9.230014e-22
[1] 3.347893e-23
[1] 1.146841e-24
[1] 3.716614e-26
[1] 1.141263e-27
```

[1] 3.325359e-29
[1] 9.206008e-31
[1] 2.424382e-32
[1] 6.079954e-34
[1] 1.453457e-35
[1] 3.315151e-37
[1] 7.220499e-39
[1] 1.502889e-40
[1] 2.991491e-42
[1] 5.698078e-44
[1] 1.039212e-45
[1] 1.815713e-47
[1] 3.040667e-49
[1] 4.882708e-51
[1] 7.521343e-53
[1] 1.111802e-54
[1] 1.577594e-56
[1] 2.149411e-58
[1] 2.81259e-60
[1] 3.535467e-62
[1] 4.269888e-64
[1] 4.955382e-66
[1] 5.526836e-68
[1] 5.924459e-70
[1] 6.103988e-72
[1] 6.044749e-74
[1] 5.753549e-76
[1] 5.263395e-78
[1] 4.627377e-80
[1] 3.909263e-82
[1] 3.173103e-84
[1] 2.474154e-86
[1] 1.852815e-88
[1] 1.332276e-90
[1] 9.195842e-93
[1] 6.09097e-95
[1] 3.870118e-97
[1] 2.357936e-99
[1] 1.376951e-101
[1] 7.703225e-104

```
[1] 4.126306e-106
[1] 2.115098e-108
[1] 1.036813e-110
[1] 4.856976e-113
[1] 2.172673e-115
[1] 9.273039e-118
[1] 3.772701e-120
[1] 1.46168e-122
[1] 5.387028e-125
[1] 1.886367e-127
[1] 6.267832e-130
[1] 1.973343e-132
[1] 5.877609e-135
[1] 1.653336e-137
[1] 4.383845e-140
[1] 1.093365e-142
[1] 2.558996e-145
[1] 5.605686e-148
[1] 1.145943e-150
[1] 2.178859e-153
[1] 3.838722e-156
[1] 6.239651e-159
[1] 9.310773e-162
[1] 1.268066e-164
[1] 1.565513e-167
[1] 1.737722e-170
[1] 1.717116e-173
[1] 1.492009e-176
[1] 1.122294e-179
[1] 7.159768e-183
[1] 3.766713e-186
[1] 1.568973e-189
[1] 4.851495e-193
[1] 9.9e-197
[1] 1e-200
```

Q.4) Plot all of the above problems in a single window for random variable and respective Probability distribution.



par(mfrow=c(2,2)) plot(x,dbinom(x,5,0.95),main="X vs Probability") x=0:5

plot(x,dbinom(x,5,0.95),main="X vs Probability") x=0:20

plot(x,dbinom(x,20,0.2),main="X vs Probability") x=0:100

plot(x,dbinom(x,100,0.01),main="X vs Probability")

Q.5) For Q.No. 1 Find P (X = 3) and P(X>3). For Q. No. 2 Find P (X = 4) and P(X>4). Find all the cumulative probabilities and round to 4 decimal places.

```
n=5
p=0.95

dbinom(3,n,p)

1-pbinom(3,n,p)

n=20
p=0.2

dbinom(4,n,p)

1-pbinom(4,n,p)
```

```
> dbinom(3,n,p)
> [1] 0.02143438
> 1-pbinom(3,n,p)
```

```
>    [1] 0.977407
>    dbinom(4,n,p
)
>    [1]
0.2181994
>    1-pbinom(4,n,p)
>    [1] 0.3703517
```

Q.6) The probability that a patient recover from a rare blood disease is 0.4. If 15 people are known to have contracted this disease, what is the probability that

(a)at least 10 survive

(b)from 3 to 8 survive

(c)exactly 5 survive?

```
n
=
1
5
p
=
0
.
4
```

```
sum((1-pbinom(10,n,p)),dbinom(10,n,p))
```

```
sum((pbinom(8,n,p)-
pbinom(3,n,p)),dbinom(3,n,p)) dbinom(5,n,p)
```

```
> sum((1-pbinom(10,n,p)),dbinom(10,n,p))
> [1] 0.0338333
> sum((pbinom(8,n,p)-pbinom(3,n,p)),dbinom(3,n,p))
> [1] 0.8778386
> dbinom(5,
n,p)
> [1]
0.1859378
```

NAME- VARANASI KASYAP

REGISTRATION NUMBER-20BCE7315

Course Title: Applied Statistics

Instructor's name: Dr. Ankur

Course Code: MAT1011

Q.1) During a laboratory experiment, the average number of radioactive particles passing through a counter in 1 millisecond is 4. What is the probability that 6 particles enter the counter in a given millisecond?

dpois(x = 6, lambda = 4)

[1] 0.1041956

Q.2) In a certain industrial facility, accidents occur infrequently. It is known that the probability of an accident on any given day is 0.005 and accidents are independent of each other.
(a) What is the probability that in any given period of 400 days there will be an accident on one day?

dpois(x = 1, lambda = 2)

[1] 0.2706706

(b) What is the probability that there are at most three days with an accident?

ppois(q = 3, lambda = 2, lower.tail = TRUE)

[1] 0.8571235

Q.3)In manufacturing process where glass products are made , defectsor bubbles occur ,occasionally rendering the piece undesirable for marketing. It is known that, on average, 1 in every 1000 of these items produced has one or more bubbles. What is the probability that a random sample of 8000 will yield fewer than 7 items possessing bubbles?

ppois(q = 6, lambda = 8, lower.tail = TRUE)

[1] 0.3133743

Q.4) On average, 3 traffic accidents per month occur at a certain intersection. What is the probability that in any given month at this intersection
(a) exactly 5 accidents will occur?

dpois(x = 5, lambda = 3)

[1] 0.1008188

(b) fewer than 3 accidents will occur?

ppois(q = 2, lambda = 3, lower.tail = TRUE)

[1] 0.4231901

(c) at least 2 accidents will occur?

1 - ppois(q = 1, lambda = 3, lower.tail = TRUE)

[1] 0.8008517

Q.5) The potential buyer of a particular engine requires (among other things) that the engine start successfully 10 consecutive times. Suppose the probability of a successful start is 0.990. Let us assume that the outcomes of attempted starts are independent.
(a) What is the probability that the engine is accepted after only 10 starts?

dpois(x=0,lambda=0.01*10)

[1] 0.9048374

(b) What is the probability that 12 attempted starts are made during the acceptance process?

dpois(x=12,lambda= 10*0.99)/10

[1] 0.009284745

Q.6) The acceptance scheme for purchasing lots containing a large number of batteries is to test no more than 75 randomly selected batteries and to reject a lot if a single battery fails. Suppose the probability of a failure is 0.001.
(a) What is the probability that a lot is accepted?

dpois(0,75*0.001)

[1] 0.9277435

(b) What is the probability that a lot is rejected on the 20th test? (c) What is the probability that it is rejected in 10 or fewer trials?

(0.001) * ((0.999)^19)

[1] 0.00098117

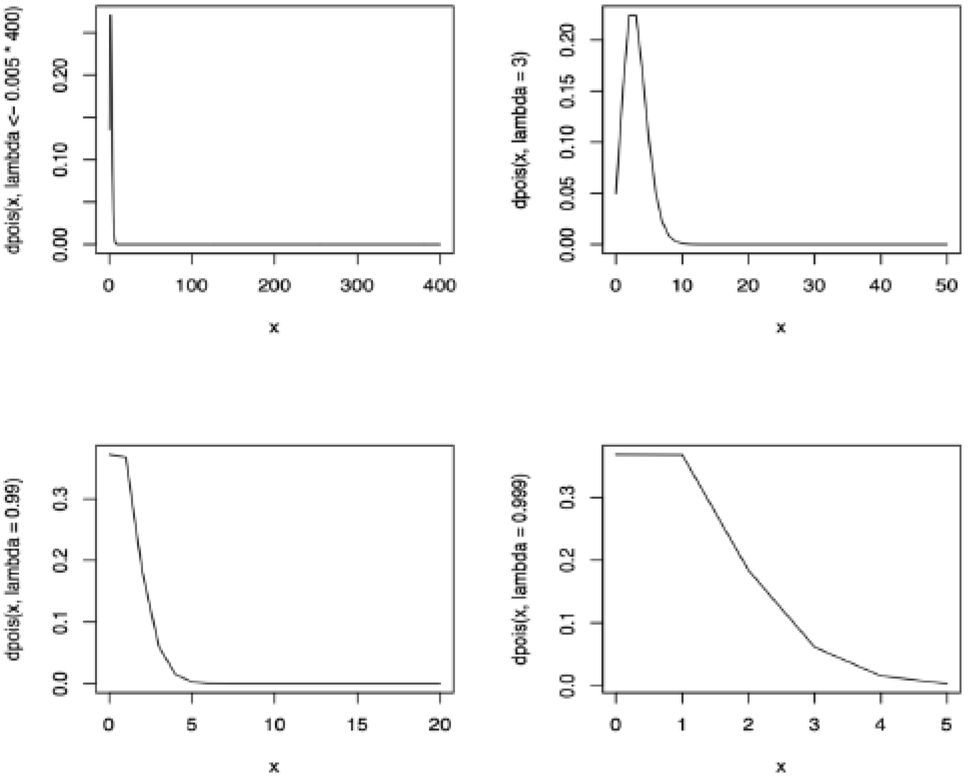
(c) What is the probability that it is rejected in 10 or fewer trials?

1 - 0.999^10

[1] 0.00995512

Q.7) Plot the graph for Q. No. 2, 4, 5 and 6 for Random Variable against Probability Distribution function.

```
par(mfrow = c(2,2))
x <- 0 : 400
plot(x,dpois(x,lambda <- 0.005*400) , type = "l")
x <- 0 : 50
plot(x,dpois(x,lambda = 3 ) , type = "l")
x <- 0 : 20
plot(x,dpois(x,lambda = 0.99) , type = "l")
x<- 0 : 5
plot(x,dpois(x,lambda = 0.999) , type = "l")
```



dpois(x, lambda <- 0.005 * 400)
0.00
0.10
0.20
0
100
200
300
400
x
dpois(x, lambda = 3)
0.00
0.05
0.10
0.15
0.20
0
10
20
30
40
50
x
dpois(x, lambda = 0.99)
0.0
0.1
0.2
0.3
0
5
10
15
20
x
dpois(x, lambda = 0.999)
0.0
0.1
0.2
0.3
0
1
2
3
4
5
x

NAME- VARANASI KASYAP

REGISTRATION NUMBER-20BCE7315

Course Title: Applied Statistics

Instructor's name: Dr. Ankur

Course Code: MAT1011

Problem Set-VI

Q.1) IQ is a normal distribution of mean of 100 and standard deviation of 15

m=100

s=15

(a)What percentage of people have an IQ<125?

Pnorm(125,m,s)*100

[1]95.22096

(b)What percentage of people have an IQ>110?

(1-pnorm(110,m,s))*100

[1]25.24925

(c)What percentage of people have 110< IQ <125?

(pnorm(125,m,s)-pnorm(110,m,s))*100

[1]20.47022

(d)Find 25% for standard normal distribution.

qnorm(0.25,0,1)

[1]-0.6744898

(e)Find 25% normal distribution with mean and standard deviation 2& 3.

qnorm(0.25,2,3)

[1]-0. 02346925

(f)What IQ separates the lower 25% from the others?

qnorm(0.25,100,15,lower.tail=TRUE)

[1]89.88265

(g)What IQ separates the top 25% from the others?

qnorm(0.25,100,15,lower.tail=FAIL)

[1]110.1173

(h)Find 25 percentile for mean 100 and SD 15

qnorm(0.25,100,15)

[1]89.88265

Q.2) Generate the 20 random number for a normal distribution with mean 572 and SD is 51.Calculate mean and SD of data set.Q.

Y<- morm(20,572,51)

Y

mean(y)

sd(y)

>y

637.2627 662.8077 538.3794 561.2866 689.4733 577.1023 542.7777

575.2301 554.0807 631.4813 496.3323 655.3692 498.9150 528.5452

599.2436 514.8669 496.0762 636.8012 590.6751 583.3823

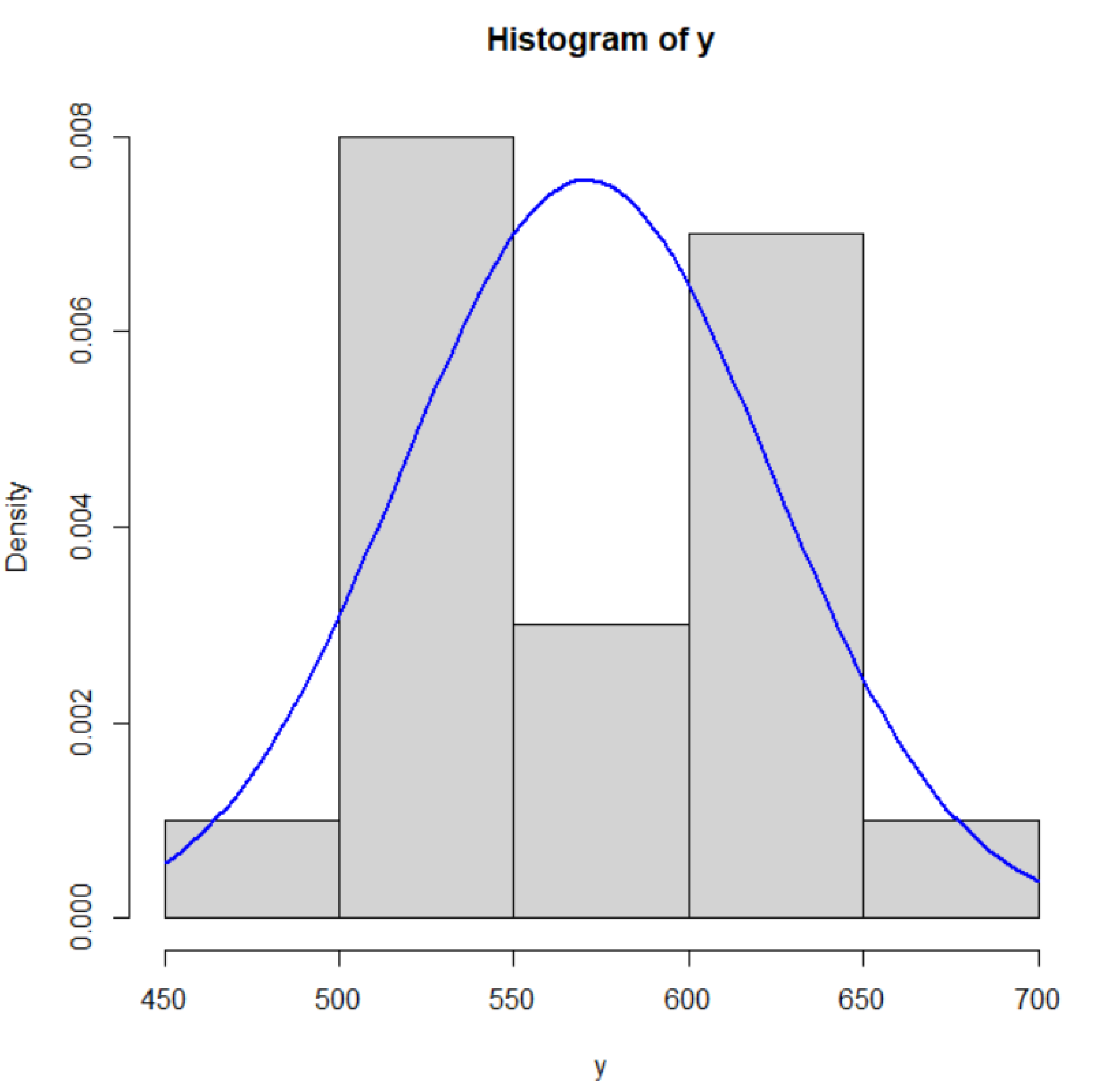
>mean(y)

[1]578.5044

```
>sd(y)
[1]58.8521
```

Q.3) Make appropriate histogram of data in above question and visually assume if normal densitycurve & histogram density estimates are similar.

```
> z<-dnorm(y,mean(y),sd(y))
> hist(y,freq=FALSE)
> curve(dnorm(x,mean(y),sd(y)),add=T,col="blue",lwd=2)
```



Name – Kasyap Varanasi

Roll Number – 20BCE7315

Assignment-8

Q.1) Test the hypothesis that the mean systolic blood pressure in a certain population equals 140mmHg. The standard deviation has a known value of 20 and a data set of 55 patients is available.BBb

Ans > No<-seq(1:55)

> Status <- c(rep(0,25),rep(1,30))

> M<-
c(120,115,94,118,111,102,102,131,105,107,115,139,115,113,114,105,115,134,109,109,93,118,109,106,125,150,142,119,127,141,149,144,142,149,161,143,140,148,149,141,146,159,152,135,134,161,130,125,141,148,153,145,137,147,175)

> BP<-data.frame(No,Status,M)

> BP

> MU=140

> XB=mean(BP$M)

> Sigma=20

> N=55

> ##Z-Value

> Z=(XB-MU)/(Sigma/sqrt(N))

> Z

[1] -3.660905

> ##P-Value

> P=2*pnorm(-abs(Z))

> P

[1] 0.0002513257

> if(P<0.5) {

+ print("The Null Hypothesis is Rejected")

+ } else{

+ print("The Null Hypothesis is Aceepted")

+ }

[1] "The Null Hypothesis is Rejected"

Q.2) A coin is tossed 100 times and turns up head 43 times Test the claim that this is a fair coin.Use 5% level of significance to test the claim.

Ans > a=100

> b=43

> c=b/a

> d=0.5

> e=1-d

> #Z – value

> f=(c-d)/(sqrt((d*e)/a))

> f

[1] -1.4

> ## P - value

> g=2*pnorm(-abs(f),lower.tail=FALSE)

> g

[1] 1.838487

> if(f<g){

+ print("This Coin is Fair")

+ } else{

+ print("This Coin is Not Fair")

+ }

[1] "This Coin is Fair"

Q.3) A manufacturer of sports equipment has developed a new synthetic fishing line that the company claims has a mean breaking strength of 8 kilograms with a standard deviation of 0.5kilogram. Test the hypothesis that μ=8 kilograms against the alternative that μ is not equal to 8kilograms if a random sample of 50 lines is tested and found to have a mean breaking strength of7.8 kilograms. Use a 0.01 level of significance.

```
Ans > a=7.8
> b=8
> c=0.5
> d=50
> ##Z-value
> e=(a-b)/(c/sqrt(d))
> e
[1] -2.828427
> ##P-Value
> f=2*pnorm(-abs(Z))
> f
[1] 0.0002513257
> if(e<f){
+ print("The Null Hypothesis is Rejected")
+ }else{
+ print("The Null Hypothesis is Accepted")
+ }
[1] "The Null Hypothesis is Rejected"
```

NAME- VARANASI KASYAP

REGISTRATION NUMBER-20BCE7315

Course Title: Applied Statistics

Instructor’s name: Dr. Ankur

Course Code: MAT1011

Problem Set-VIII

Question 1

An outbreak of salmonella-related illness was attributed to ice produced at certain factory. Scientists measured the level of salmonella in 9 randomly sampled batches ice cream. The levels (in MPN/g) were:0.5930.1420.3290.6910.2310.7930.5190.3920.418Is there evidence that the mean level of salmonella in ice cream greater than 0.3 MPN/g.

Code-

```
> A=c(0.593,0.142,0.329,0.691,0.231,0.793,0.519,0.392,0.418)

> t.test(A,alternative="greater",mu=0.3)
```

```
        One Sample t-test

data:  A
t = 2.2051, df = 8, p-value = 0.02927
alternative hypothesis: true mean is greater than 0.3
95 percent confidence interval:
 0.3245133       Inf
sample estimates:
mean of x
0.4564444
```

Question 2

Suppose that 10 volunteers have taken an intelligence test; here are the results obtained. The average score of the entire population is 75 in the entire test. Is there any significant difference (with a significance level of 95%) between the sample and population means, assuming that the variance of the population is not known?

Scores: 65, 78, 88, 55, 48, 95, 66, 57, 79, 81.

Code-

```
> a = c(65, 78, 88, 55, 48, 95, 66, 57, 79, 81)

> t.test (a, mu=75)


        One Sample t-test


data:  a

t = -0.78303, df = 9, p-value = 0.4537

alternative hypothesis: true mean is not equal to 75

95 percent confidence interval:

 60.22187 82.17813

sample estimates:

mean of x

     71.2


> qt(0.975, 9)

[1] 2.262157
```

Question 3

Comparing two independent sample means, taken from two population with unknown variance.The following data shows the heights of the individuals of two different countries with unknown population variances. Is there any significant difference between the average heights of the two groups?

A: 175, 168, 168, 190, 156, 181, 182, 175, 174, 179

B: 185, 169, 173, 173, 188, 186, 175, 174, 179, 180

Code-

```
> a = c(175, 168, 168, 190, 156, 181, 182, 175, 174, 179)

> b = c(185, 169, 173, 173, 188, 186, 175, 174, 179, 180)

> t.test(a,b, var.equal=TRUE, paired=FALSE)


        Two Sample t-test


data:  a and b

t = -0.94737, df = 18, p-value = 0.356

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

 -10.93994   4.13994

sample estimates:

mean of x mean of y

    174.8     178.2


> qt(0.975, 18)

[1] 2.100922
```

NAME- VARANASI KASYAP

REGISTRATION NUMBER-20BCE7315

Course Title: Applied Statistics

Instructor's name: Dr. Ankur

Course Code: MAT1011

PROBLEM SET- IX

Q.1) It is important that scientific researchers in the area of forest products be able to studycorrelation among the anatomy and mechanical properties of trees. For the study QuantitativeAnatomical Characteristics of Plantation Grown Loblolly Pine (Pinus Taeda L.) and Cottonwood(Populus deltoides Bart. Ex Marsh.) and Their Relationships to Mechanical Properties, conductedby the Department of Forestry and Forest Products at Virginia Tech, 29 loblolly pines wererandomly selected for investigation. Table shows the resulting data on the specific gravity ingrams/cm3 and the modulus of rupture in kilopascals (kPa). Compute and interpret the samplecorrelation coefficient.

```
> xi<-c(0.414 ,0.383 ,0.399 ,0.402 ,0.442 ,0.422 ,0.466 ,0.500 ,0.514 ,0.530 ,0.569,0.558
+  ,0.577 ,0.572 ,0.548 ,0.581, 0.557, 0.550, 0.531, 0.550, 0.556, 0.523, 0.602,0.569
+  ,0.544, 0.557, 0.530, 0.547, 0.585)
> xi
```

```
 [1] 0.414 0.383 0.399 0.402 0.442 0.422 0.466 0.500 0.514 0.530 0.569 0.558
[13] 0.577 0.572 0.548 0.581 0.557 0.550 0.531 0.550 0.556 0.523 0.602 0.569
[25] 0.544 0.557 0.530 0.547 0.585
>  yi<-c(29186,29266, 26215, 30162, 38867, 37831, 44576, 46097, 59698, 67705,6608 ,78486
+ ,89869, 77369, 67095, 85156, 69571, 84160, 73466, 78610, 67657, 74017,87291,86836
+ ,82540,81699,82096, 75657, 80490)
> yi
 [1] 29186 29266 26215 30162 38867 37831 44576 46097 59698 67705  6608 78486
[13] 89869 77369 67095 85156 69571 84160 73466 78610 67657 74017 87291 86836
[25] 82540 81699 82096 75657 80490
> xbar<- mean(xi)
>  xbar
[1] 0.519931
>  yi<- B$Modulus.of.Rupture
Error: object 'B' not found
>  yi
 [1] 29186 29266 26215 30162 38867 37831 44576 46097 59698 67705  6608 78486
[13] 89869 77369 67095 85156 69571 84160 73466 78610 67657 74017 87291 86836
[25] 82540 81699 82096 75657 80490
>  ybar<- mean(yi)
>  ybar
[1] 63388.83
>  r<-sum((xi-xbar)*(yi-ybar))
>  xsq<-sum((xi-xbar)^2)
```

```
> ysq<-sum((yi-ybar)^2)
> corp<- r/(sqrt(sum((xi-xbar))^2)*(sum((yi-ybar))^2))
> cor1<- r/(sqrt(xsq*ysq))
> cor1
[1] 0.7624076
> cor(xi,yi)
[1] 0.7624076
```

Q.2) Compute and interpret the correlation coefficient for the following grades of 6 studentsselected at random:

```
> ai<-c(70, 92, 80, 74, 65,83)
> ai
[1] 70 92 80 74 65 83
> abar<- mean(ai)
> bi<-c( 74, 84, 63, 87, 78, 90)
> bi
[1] 74 84 63 87 78 90
> bbar<- mean(bi)
> bbar
[1] 79.33333
> c<-sum((ai-abar)*(bi-bbar))
> asq<-sum((ai-abar)^2)
> bsq<-sum((bi-bbar)^2)
> cor<- c/(sqrt(sum((ai-abar))^2)*(sum((bi-bbar))^2))
```

```
>  cor2<- c/(sqrt(asq*bsq))
>  cor2
[1] 0.2396639
>  cor(ai,bi)
[1] 0.2396639
```

Q.3) Assume that x and y are random variables with a bivariate normal distribution.Calculater.

```
> di<-c(17.3,19.3,19.5,19.7,22.9,23.1,26.4,26.8,27.6,28.1,28.2,28.7,29.0,29.6,29.9
+ ,29.9,30.3,31.3,36.0,39.5,40.4,44.3,44.6,50.4,55.9)
> di
 [1] 17.3 19.3 19.5 19.7 22.9 23.1 26.4 26.8 27.6 28.1 28.2 28.7 29.0 29.6 29.9
[16] 29.9 30.3 31.3 36.0 39.5 40.4 44.3 44.6 50.4 55.9
>  dbar<- mean(di)
>  dbar
[1] 31.148
> ei<-c(71.7,48.3,88.3,75.0,91.7,100.0,73.3,65.0,75.0,88.3,68.3,96.7,
+ 76.7,78.3,60.0,71.7,85.0,85.0,88.3,100.0,100.0,100.0,91.7,100.0,71.7)
> ei
 [1]  71.7  48.3  88.3  75.0  91.7 100.0  73.3  65.0  75.0  88.3  68.3  96.7
[13]  76.7  78.3  60.0  71.7  85.0  85.0  88.3 100.0 100.0 100.0  91.7 100.0
[25]  71.7
> ebar<- mean(ei)
> ebar
```

```
[1] 82
> f<-sum((di-dbar)*(ei-ebar))
> dsq<-sum((di-dbar)^2)
> esq<-sum((ei-ebar)^2)
> cor<- f/(sqrt(sum((di-dbar))^2)*(sum((ei-ebar))^2))
> cor3<- f/(sqrt(dsq*esq))
> cor3
[1] 0.3916965
> cor(di,ei)
[1] 0.3916965
```

Que: In a certain type of metal test specimen, the normal stress on a specimen is known to be functionally related to the shear resistance. The following is a set of coded experimental data on the two variables

Normal Stress, x	Shear Resistance, y
26.8	26.5
25.4	27.3
28.9	24.2
23.6	27.1
27.7	23.6
23.9	25.9
24.7	26.3
28.1	22.5
26.9	21.7
27.4	21.4
22.6	25.8
25.6	24.9

• Estimate the shear resistance for a normal stress of 24.5.

• Plot the data; does it appear that a simple linear regression will be a suitable model?

Ans:

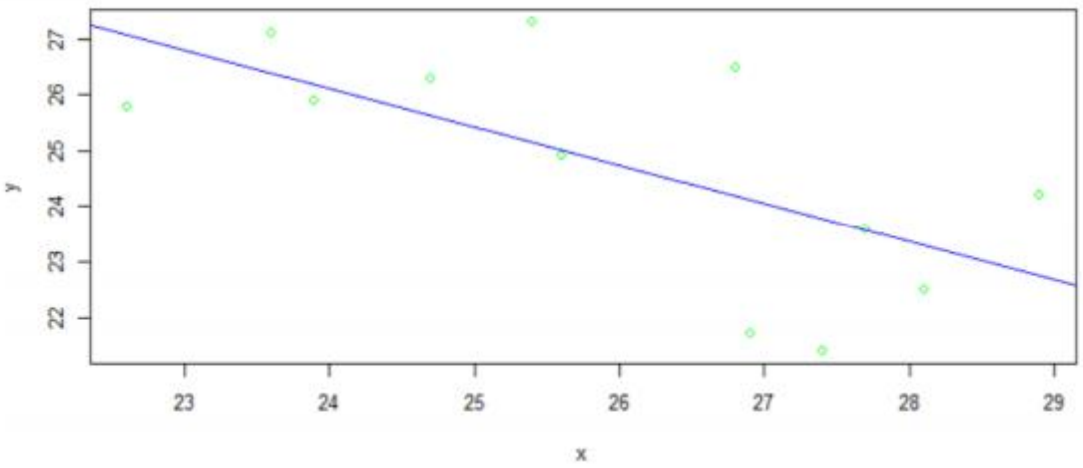
CODE:

```
x=c(26.8,25.4,28.9,23.6,27.7,23.9,24.7,28.1,26.9,27.4,22.6,25.6)

y=c(26.5,27.3,24.2,27.1,23.6,25.9,26.3,22.5,21.7,21.4,25.8,24.9)

lm(y~x)

a<-sum(x*y)-((sum(x)*sum(y))/length(x))

b<-sum(x*x)-((sum(x)*sum(x))/length(x))

c<-sum(y*y)-((sum(y)*sum(y))/length(x))

slope<-a/b
```

```
const<-(sum(y)-(slope*sum(x)))/length(x)
slope
const
plot(x,y,col="green")
abline(a=const,b=slope,col="red")
abline(lm(y~x),col="blue")
yval<-slope*24.5+const
yval
```

OUTPUT:

```
Call:
lm(formula = y ~ x)
Coefficients:
(Intercept)          x
    42.5818    -0.6861
slope
[1] -0.6860771
> const
[1] 42.5818
yval
[1] 25.77291
```



2.) A study was made by a retail merchant to determine the relation between weekly advertising expenditures and sales.

(a) Plot a scatter diagram.

(b) Find the equation of the regression line to predict weekly sales from advertising expenditures.

Advertising Costs ($)	Sales ($)
40	385
20	400
25	395
20	365
30	475
50	440
40	490
20	420
50	560
40	525
25	480
50	510

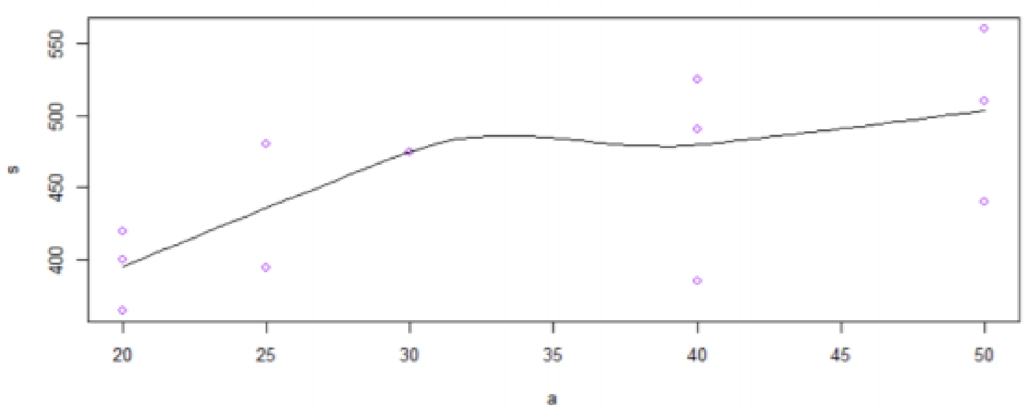
Ans:

CODE:

```
a<- c(40,20,25,20,30,50,40,20,50,40,25,50)
s<-c(385,400,395,365,475,440,490,420,560,525,480,510)
x<-sum(a)
y<-sum(s)
sxs<-sum(a*s)
xs<-sum(a^2)
l<-length(a)
ls<-length(s)
lls<-length(a*s)
z<-(l*sxs-x*y)/(l*xs-x^2)
z
d<-(y-z*x)/l
d
scatter.smooth(a,s,col="purple")
```

OUTPUT:

```
z
[1] 3.220812
> d<-(y-z*x)/l
> d
[1] 343.7056
```



Real life problem-

QUESTION:

The mean height of 61 males from the name state was 68.2 inches with an estimated standard deviation of 2.5 inches, while 61 males from another state had a mean height of 67.5 inches estimated standard deviation of 2.8 inches. The heights are normally distributed. Test the hypothesis if there is any significant difference.

```
> x1bar=68.2
> n1=61
> x2bar=67.5
> n2=61
> sigma1=2.5
> sigma2=2.8
> level=0.05
> z=(x1bar-x2bar)/sqrt(((sigma1)^2/n1)+((sigma2)^2/n2))
> z
[1] 1.45649
> pvalue=2*pnorm(z,lower.tail=TRUE)
> pvalue
[1] 1.854743
> if(pvalue<=level)
+ {
+ print("NULL HYPOTHESIS H0 is rejected")
+ }else
+ {
+ print("NULL HYPOTHESIS H0 is accepted")
+ }
[1] "NULL HYPOTHESIS H0 is accepted"
```