Programing language R (object oriented language)

Introduction to R programming language : (Sheet: 01)

Assignment no: 1

Mathematical operation:

Question	Answer
pi	> pi=[1] 3.141593
x<-2;y<-3;x^y;	> x<-2; > y<-3; > x^y;=[1] 8
sqrt(x)	> sqrt(x)=[1] 1.414214
abs(x)	> abs(x)=[1] 2
factorial(x)	> factorial(x)=[1] 2
log(x,base=2)	> log(x,base=2)=[1] 1
log10(x)	> log10(x)=[1] 0.30103
log2(x)	> log2(x)=[1] 1
exp(2)	> exp(2)=[1] 7.389056
cos(x)	> cos(x)=[1] -0.4161468
sin(x)	> sin(x)=[1] 0.9092974
tan(x);	> tan(x)=[1] -2.18504
x<-1;	> x<-1;
acos(x)	> acos(x) = [1] 0
asin(x)	> asin(x)=[1] 1.570796
atan(x)	> atan(x)=[1] 0.7853982
2*pi/3-sqrt(4)	> 2*pi/3-sqrt(4) =[1] 0.0943951
abs(12-17*2/3-9)	> abs(12-17*2/3-9) = [1] 8.333333
exp(log(2))	>exp(log(2))=2
10^log10(2)	10^log10(2)= 2
asin(1/sqrt(2))*180/pi	asin(1/sqrt(2))*180/pi= 45

(1)Creating vector:

Question	Answer
a.(1,2,3……..19,20)	> x<-c(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.(20,19……..2,1)	rev(x)=[1] 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 or > x<-c(20:1) > x=[1] 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
c.(1,2…….19,20,20,19……2,1)	x<-c(1:20,20:1); >x; =[1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 20 19 18 17 16 [26] 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
d. (4,6,3) and assign it to data……	> x<- c(4,6,3);x;=[1] 4 6 3
e. (4,6,3,4,6,3, , , , , , 4,6,3) where there are 10 occurance of 4,6,3….	> data<- c(4,6,3);data;=[1] 4 6 3

	> rep(data,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 4 6 3 4 6 3 4 6 3 4 6 3 4 6 3
f. (4,6,3,4,6,3,,,,,,,,,,,,,,,4,6,3,4) where there are 11 occurance of 4, 10 occurance of 6 and 10 occurance of 3..	> data<-c (4, 6, 3) > rep (data,l=31)=[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 4 6 3 4 6 3 4 6 3 4 6 3 4 6 3 4
g. (4,4,4,,,,,,,,,,6,6,6,,,,,,,,,,,,3,3,3,3) where there are 10 occurance of 4, 20 occurance of 6 and 30 occurance of 3……	z<-rep(data,times=c(10,20,30));z; =[1] 4 4 4 4 4 4 4 4 4 4 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 3 3 3 3 3 3 3 3 [39] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3

h. Create a vector of the values of $e^x \cos(x)$ at x= (3,3.1,3.2,……,5.9,6)	>x<- seq(3,6 by=0.1);
	>y<- exp(x)*cos(x);
	>z<- data.frame(x,y);z;
	`x y 11 4.0 -35.687732 21 5.0 42.099201`
	`1 3.0 -19.884531 12 4.1 -34.685042 22 5.1 61.996630`
	`2 3.1 -22.178753 13 4.2 -32.693695 23 5.2 84.929067`
	`3 3.2 -24.490697 14 4.3 -29.538816 24 5.3 111.061586`
	`4 3.3 -26.773182 15 4.4 -25.032529 25 5.4 140.525075`
	`5 3.4 -28.969238 16 4.5 -18.975233 26 5.5 173.405776`
	`6 3.5 -31.011186 16 4.5 -18.975233 27 5.6 209.733494`
	`7 3.6 -32.819775 17 4.6 -11.157417 28 5.7 249.468441`
	`8 3.7 -34.303360 18 4.7 -1.362099 29 5.8 292.486707`
	`9 3.8 -35.357194 19 4.8 10.632038 30 5.9 338.564378`
	`10 3.9 -35.862834 20 4.9 25.046705 31 6.0 387.360340`

Question no 2: Create the following vector:

Question	Answer
a) (a)$(0.1^3 0.2^1, 0.1^6 0.2^4, .., 0.1^{36} 0.2^{34})$	> x<-(.1^seq(3,36,by=3)*.2^seq(1,34,by=3)) > x; [1] 2.000000e-04 1.600000e-09 1.280000e-14 1.024000e-19 8.192000e-25 [6] 6.553600e-30 5.242880e-35 4.194304e-40 3.355443e-45 2.684355e-50 [11] 2.147484e-55 1.717987e-60
(b) $\left(2, \frac{2^2}{2}, \frac{2^3}{3}, \ldots, \frac{2^{25}}{25}\right)$	> x<-c(2^(seq(1,25,by=1))/(1:25)) > x; [1] 2.000000e+00 2.000000e+00 2.666667e+00 4.000000e+00 6.400000e+00 [6] 1.066667e+01 1.828571e+01 3.200000e+01 5.688889e+01 1.024000e+02 [11] 1.861818e+02 3.413333e+02 6.301538e+02 1.170286e+03 2.184533e+03 [16] 4.096000e+03 7.710118e+03 1.456356e+04 2.759411e+04 5.242880e+04 [21] 9.986438e+04 1.906502e+05 3.647221e+05 6.990507e+05 1.342177e+06

Question no 3: Create the following vector:

Question	Answer
(a) $\sum_{i=10}^{100}(i^3 + 4i^2)$	> i<-seq(10,100) > y<-sum(i^3+4*i^2) > y; [1] 26852735
(b) $\sum_{i=1}^{25}\left(\frac{2^i}{i} + \frac{3^i}{i^2}\right)$	> i<-(1:25) > x<-sum(2^i/i+3^i/(i*i)) > x; [1] 2129170437

Question no 4: Create the following vector:

Question	Answer

(a) Create a vector from 0 to 1 in increments of 0.1	`> m<-seq(0,1,by=.1)`
	`> m;`
	`[1] 0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0`

(b) Print out the first 3 element of m and only third element, replace the third element.	`> x<-(1:3)`
	`> m[x]`
	`[1] 0.0 0.1 0.2`
	`> m[3]`
	`[1] 0.2`
	`> m[3]=4`

(c) Create a vector consisting of 3 names	`> m<-c("bob","bill","sue")`
	`> m;`
	`[1] "bob" "bill" "sue"`

(d) Create vector x,y and include in matrix	`> z<-matrix(c(x,y),2,3)`
	`> z;`
	`[,1] [,2] [,3]`
	`[1,] 1 3 4`
	`[2,] 2 3 5`

Assignment no:2

Question : Trees data set provides the measurement of the girth height and volume of timber in 31 felled black cherry trees. Give the answer of following question.

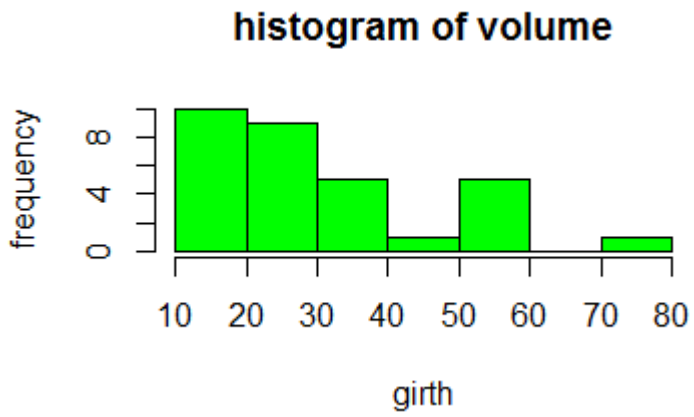
Question	Answer

1) Read the trees data set.	`> trees`
	`Girth Height Volume`
	`1 8.3 70 10.3`
	`2 8.6 65 10.3`
	`3 8.8 63 10.2`
	`4 10.5 72 16.4`
	`5 10.7 81 18.8`
	`6 10.8 83 19.7`
	`7 11.0 66 15.6`
	`8 11.0 75 18.2`
	`9 11.1 80 22.6`
	`10 11.2 75 19.9`
	`11 11.3 79 24.2`
	`12 11.4 76 21.0`
	`13 11.4 76 21.4`
	`14 11.7 69 21.3`
	`15 12.0 75 19.1`
	`16 12.9 74 22.2`
	`17 12.9 85 33.8`
	`18 13.3 86 27.4`
	`19 13.7 71 25.7`
	`20 13.8 64 24.9`
	`21 14.0 78 34.5`
	`22 14.2 80 31.7`
	`23 14.5 74 36.3`
	`24 16.0 72 38.3`
	`25 16.3 77 42.6`
	`26 17.3 81 55.4`
	`27 17.5 82 55.7`
	`28 17.9 80 58.3`
	`29 18.0 80 51.5`
	`30 18.0 80 51.0`
	`31 20.6 87 77.0`

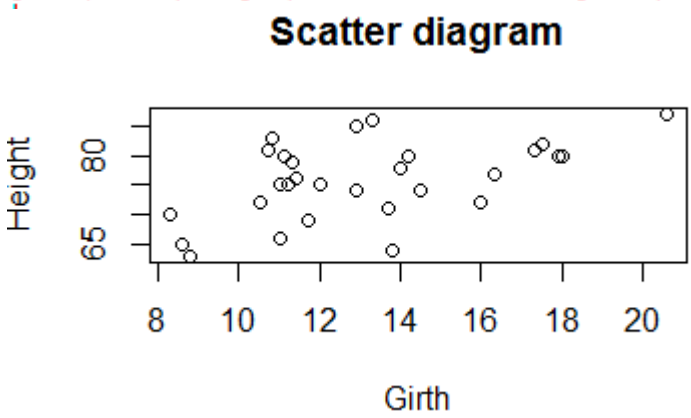
2) Just read the variable name girth.	`> trees$Girth`
	` [1]  8.3  8.6  8.8 10.5 10.7 10.8 11.0 11.0 11.1 11.2 11.3 11.4 11.4 11.7 12.0`
	`[16] 12.9 12.9 13.3 13.7 13.8 14.0 14.2 14.5 16.0 16.3 17.3 17.5 17.9 18.0 18.0`
	`[31] 20.6`

3) Attached trees data set	`> attach(trees)` `The following object(s) are masked from 'trees (position 3)':`

4) Remove tress data set from work space	`> dev.off()`
	`null device`
	`1`

5) Simple histogram of volume	`> x<-(trees$Volume)`
	`> x`
	` [1] 10.3 10.3 10.2 16.4 18.8 19.7 15.6 18.2 22.6 19.9 24.2 21.0 21.4 21.3 19.1`
	`[16] 22.2 33.8 27.4 25.7 24.9 34.5 31.7 36.3 38.3 42.6 55.4 55.7 58.3 51.5 51.0`
	`[31] 77.0`
	`> hist(x,xlab="girth",ylab="frequency",main="histogram of girth",col="green")`
	 

6) Box plot of girth	`> x<-(trees$Girth)`
	`> boxplot(x,main="girth",col="black")`
	 
	Comment: There is no outlier in girth variable.

7) Scatter plot of girth and height	`plot(Girth,Height,main="Scatter diagram",col="black")` 

8)Make five number summary of girth variable	`summary(Girth)`
	`Min. 1st Qu. Median Mean 3rd Qu. Max.`
	`8.30 11.05 12.90 13.25 15.25 20.60`

9)Find variance and length ,standard deviation of girth variable	`> var(Girth)`
	`[1] 9.847914`
	`> sd(Girth)`
	`[1] 3.138139`
	`> length(Girth)`
	`[1] 31`
	`> length(Height)`
	`[1] 31`

10)Calculate correlation coefficient.	`> cor(trees$Girth,trees$Height);` `[1] 0.5192801`

<u>Introduction to MATRICES:</u> (Sheet: 02)

Assignment no :03 (sheet-2)

Question	Answer
$A = \begin{matrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{matrix}$	> A<-matrix(c(1,5,-2,1,2,-1,3,6,-3),3,3); > A;= [,1] [,2] [,3] [1,] 1 1 3 [2,] 5 2 6 [3,] -2 -1 -3
a .Check that A^3=0;	> result<-A%*%A%*%A; result; = [,1] [,2] [,3] [1,] 0 0 0 [2,] 0 0 0 [3,] 0 0 0

b .Replace the third column of A by the sum of the second and third columns.	`> A[,3]<-A[,2]+A[,3];`
	`> A;`
	`[,1] [,2] [,3]`
	`[1,] 1 1 4`
	`[2,] 5 2 8`
	`[3,] -2 -1 -4`

c. replace the 2nd row of A by the subtraction of 2nd and third of A.	`> A[2,]<-abs(A[2,]-A[3,]);`
	`> A;`
	`[,1] [,2] [,3]`
	`[1,] 1 1 4`
	`[2,] 7 3 12`
	`[3,] -2 -1 -4`

##create the following matrix B with 15 row.	`> B<-matrix(c(rep(10,15),rep(-10,15),rep(10,15)),15,3);`
	`> B;`
	`[,1] [,2] [,3]`
	`[1,] 10 -10 10`
	`[2,] 10 -10 10`
	`[3,] 10 -10 10`
	`[4,] 10 -10 10`
	`[5,] 10 -10 10`
	`[6,] 10 -10 10`
	`[7,] 10 -10 10`
	`[8,] 10 -10 10`
	`[9,] 10 -10 10`
	`[10,] 10 -10 10`
	`[11,] 10 -10 10`
	`[12,] 10 -10 10`
	`[13,] 10 -10 10`
	`[14,] 10 -10 10`
	`[15,] 10 -10 10`

## calculate the 3*3 matrix $B'B$	`> result<-t(B)%*%B;result;` 	
	`[,1] [,2] [,3]`
	`[1,] 1500 -1500 1500`
	`[2,] -1500 1500 -1500`
	`[3,] 1500 -1500 1500`

## show system date	`> date()` `[1] "Tue Nov 01 20:41:20 2016"`

Question : 01

A relationship between a childs height and their head circumstance …….

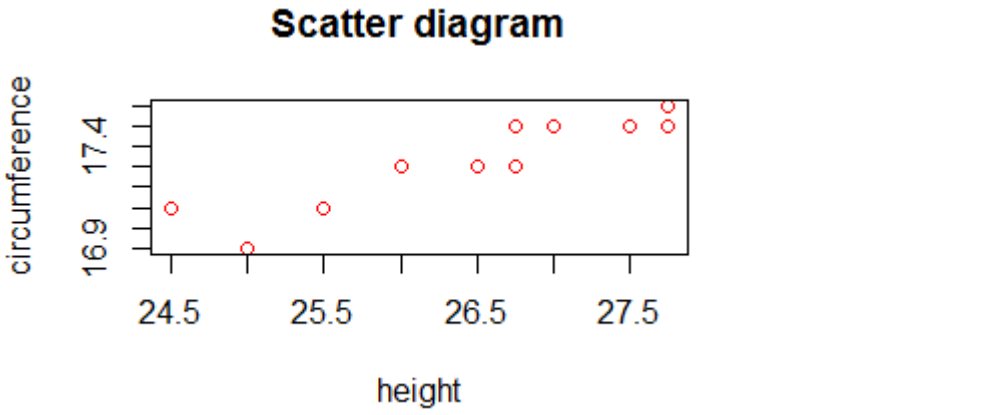
Height	27.75	24.5	25.5	26	25	27.75	26.5	27	26.75	26.75	27.5
Circum-ference	17.5	17.1	17.1	17.3	16.9	17.6	17.3	17.5	17.3	17.5	17.5

Answer:

Question	Answer

a.Read this data in R	`> z<-read.table("E:\\Head.txt",header=T)`
	`> z`
	`height circumference`
	`1 27.75 17.5`
	`2 24.50 17.1`
	`3 25.50 17.1`
	`4 26.00 17.3`
	`5 25.00 16.9`
	`6 27.75 17.6`
	`7 26.50 17.3`
	`8 27.00 17.5`
	`9 26.75 17.3`
	`10 26.75 17.5`
	`11 27.50 17.5`

	Or,
	`> height<-c(27.75,24.5,25.5,26,25,27.75,26.5,27,26.75,26.75,27.5);height;`
	`[1] 27.75 24.50 25.50 26.00 25.00 27.75 26.50 27.00 26.75 26.75 27.50`
	`> circumference<-c(17.5,17.1,17.1,17.3,16.9,17.6,17.3,17.5,17.3,17.5,17.5);circumference;`
	`[1] 17.5 17.1 17.1 17.3 16.9 17.6 17.3 17.5 17.3 17.5 17.5`
	
	variable<-data.frame(height,circumference);
	attach(variable);variable;
	plot(variable$height,variable$circumference);

b. make a scatter plot circumference against height.	`> attach(z)`
	`The following object(s) are masked from 'z (position 3)':`
	
	`circumference, height`
	`> plot(height,circumference,main="Scatter diagram",col="red")`
	 
	Comment: There appearance to be a positive linear relationship between two variables.

c. Find co-variance of height and circumference and comment.	`> cov(height,circumference);`
	`[1] 0.2188636`
	`> cov(z)`
	`height circumference`
	`height 1.1977273 0.21886364`
	`circumference 0.2188636 0.04818182`

d. find co-relation between height and circumference.	`> cor(height,circumference);`
	`[1] 0.9110727`
	`> cor(z)`
	`height circumference`
	`height 1.0000000 0.9110727`
	`circumference 0.9110727 1.0000000`

e. Fit a simple linear regression of height and circumference.	`> model<-lm(circumference~height)`
	`> model`
	
	`Call:`
	`lm(formula = circumference ~ height)`
	
	`Coefficients:`
	`(Intercept)  height`
	`12.4932  0.1827`
	
	`> plot(height,circumference,main="Scatter",col="red");`
	`> abline(model)`
	`> summary(model)`
	
	`Call:`
	`lm(formula = circumference ~ height)`
	
	`Residuals:`
	`Min  1Q  Median  3Q  Max`
	`-0.16148 -0.05842 -0.01831  0.06442  0.12989`
	
	`Coefficients:`
	`Estimate Std. Error t value Pr(>|t|)`
	`(Intercept) 12.49317  0.72968  17.12 3.56e-08 ***`
	`height  0.18273  0.02756  6.63 9.59e-05 ***`
	`---`
	`Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1`
	
	`Residual standard error: 0.09538 on 9 degrees of freedom`
	`Multiple R-squared: 0.8301,  Adjusted R-squared: 0.8112`
	`F-statistic: 43.96 on 1 and 9 DF,  p-value: 9.59e-05`

<u>Introduction to probability distribution:</u> (Sheet : 03)

Assignment no: 04

Probability mass function for the binomial distribution.

<u>Question</u>	<u>Answer</u>
suppose that a fair dice is rolled 10 times. What is the probability of throwing exactly two sixes?	`> dbinom(2,10,1/6)` `[1] 0.29071`

<u>Probability density function for normal distribution :</u>

<u>Question</u>	<u>Answer</u>

Fin the value of the pdf at x=2.5 for a normal distribution with the mean of 5 and a standard deviation of 2.	`> dnorm(2.5,5,2);` `[1] 0.09132454`

Finding probability :

Question	Answer
1. For the standard normal distribution find the probability that a randomly selected value will be less than or equal to 2.5.	`> pnorm(2.5,0,1)` `[1] 0.9937903`

2. For normal distribution where mean=5, SD=2,
a. When X=6
b. When x>6
	a.
	`> pnorm(6,5,2);`
	`[1] 0.6914625`
	`> 1-pnorm(6,5,2);`
	`[1] 0.3085375`
	b.

3. The number of lobster ordered in a restaurant on a given day is known to follow a poisson distribution with a mean of 20. What is the probability that exactly 18 lobsters will be ordered tomorrow ?	`> dpois(18,20)` `[1] 0.08439355`
Finding quartiles	
1. For a standard normal distribution find the value below which 95% of values fall?	`> qnorm(.95);` `[1] 1.644854`

2. Suppose a variables known to be normal distributed with a mean of 5 and standard deviation of 2.
a. Find the value below which 95% of the population falls?
b. Value above which 95% of the population falls?
	Answer:
	`> qnorm(.95,5,2);`
	`[1] 8.289707`
	a.
	`> qnorm(.05,5,2);`
	`[1] 1.710293`
	b.

3. A manufacturer of special type of one size glove wants to design the glove to fit at least 99% OF THE POPULATION. HAND span is known to be normally distributed with a mean of 195 millimeters and standard deviation of 17 millimeters.
 	
## *what range of the hand spans must the glove accommodate?* 	
	`> x<-qnorm(.995,195,17);x;`
	`[1] 238.7891`
	`> y<-qnorm(.005,195,17);y;`
	`[1] 151.2109`
	`> z<-x-y;`
	`> z;`
	`[1] 87.5782`

One sample t test	

*An outbreak of salmonella related illness was attributed to ice-cream produced at a certain factory scientist measured the level of salmonella in 9 ice-cream: (**.593, .142, .329, .691, .231, .793, .519, .392, .418)**
 	
## is there any evidence that the mean level of salmonella in the ice-cream is greater than 0.3MPN/g? 	
	x<-c(.593,.142,.329,.691,.231,.793,.519,.392,.418);
	x;
	t.test(x,alternative=c("greater"),mu=.3);
	Answer:
	
	One Sample t-test
	
	data: x
	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
	
	From output: x'=0.45, t static=2.2, degree of freedom=8, p-value=.029 , here p<level of significance (.05)
	So, we may reject null hypothesis.(m1<=.3)
	And accept m1>.30

*Six subjects were drug treatment group and additional 6 subjects a place. there reaction time was measured … ***Control group: (91,87,99,77,88,91)*** ***Treatment group: (101,110,103,93,99,104)***	x<-c(91,87,99,77,88,91) x y<-c(101,110,103,93,99,104) y t.test(x,y,alternative=c("less"),paired=FALSE,var.equal=TRUE) Two Sample t-test data: x and y t = -3.4456, df = 10, p-value = 0.003136 alternative hypothesis: true difference in means is less than 0 95 percent confidence interval: -Inf -6.082744 sample estimates: mean of x mean of y 88.83333 101.66667 mean of x and y is 88.83333 101.66667 t = -3.4456, df = 10, p-value = 0.003136<level of significance (.05) So, we may reject null hypothesis m1=m2 and accept m1<m2 .
A study was performed to test whether cars get better millage an premium gas than regular gas. Each of 10 cars was first filled with regular gas or premium gas decided by a toss and the millage for that tank was recorded again for the same cars using the other kind of gasoline use paired **t test** to determine whether cars get significantly better millage with premium gas. **Regular gas:** **(16,20,21,22,23,22,27,25,27,28)**	a<-c(16,20,21,22,23,22,27,25,27,28) a b<-c(19,22,24,24,25,25,26,26,28,32) b t.test(b,a,alternative=c("greater"),paired=TRUE)

Premium gas: (19,22,24,24,25,25,26,26,28,32)	Paired t-test data: b and a t = 4.4721, df = 9, p-value = 0.0007749 alternative hypothesis: true difference in means is greater than (95 percent confidence interval: 1.180207 Inf sample estimates: mean of the differences 2 data: b and a t = 4.4721, df = 9, p-value = 0.0007749<level of significance(.05), So we may reject null(m2=m1) hypothesis and accept alter (m2>m1) hypothesis.

Assignment no:05 (sheet no :4)

Construction of matrix:

Question	Answer

1 2 3
If X=4 5 6
7 8 9
a) Find transpose of X.
b) Calculate X' and find XX'.
c) Find diagonal of X.
d) Find determinant of X
e) Find inverse of X
f) Find dimension of X
	a).
	`> x<-matrix(1:9,3);`
	`> x;`
	`[,1] [,2] [,3]`
	`[1,] 1 4 7`
	`[2,] 2 5 8`
	`[3,] 3 6 9`
	b)
	`> t(x);`
	`[,1] [,2] [,3]`
	`[1,] 1 2 3`
	`[2,] 4 5 6`
	`[3,] 7 8 9`
	`> z<-(x%*%y);`
	`> z;`
	`[,1] [,2] [,3]`
	`[1,] 14 32 50`
	`[2,] 32 77 122`
	`[3,] 50 122 194`
	c)
	`> diag(x);`
	`[1] 1 5 9`
	d)
	`> det(y);`
	`[1] 0`
	e)
	`> solve(x)`
	`[,1] [,2] [,3]`
	`[1,] -0.6666667 -1.333333 1`
	`[2,] -0.6666667 3.666667 -2`
	`[3,] 1.0000000 -2.000000 1`
	f)
	`> dim(x);`
	`[1] 3 3`

Construct this:
Y= 1.5 0.0 0.00 0.0
0.0 2.1 0.00 0.0
0.0 0.0 0.08 0.0
0.0 0.0 0.00 4.1
	`> y<-diag(c(1.5,2.1,.08,4.1));`
	`> y;`
	`[,1] [,2] [,3] [,4]`
	`[1,] 1.5 0.0 0.00 0.0`
	`[2,] 0.0 2.1 0.00 0.0`
	`[3,] 0.0 0.0 0.08 0.0`
	`[4,] 0.0 0.0 0.00 4.1`

Create a matrix **Y** given below:

11,12,13……20
Y= 32,34,36……50
35,37,39……53

Form Y :

a) Create Y1=1st row.
b) Create Y2=2nd row.
c) Create Y3=3rd row.

```
> y<-matrix(c(11:20,seq(32,50,2),seq(35,53,2)),3,10,byrow=T);
> y;
     [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
[1,]   11   12   13   14   15   16   17   18   19    20
[2,]   32   34   36   38   40   42   44   46   48    50
[3,]   35   37   39   41   43   45   47   49   51    53
```

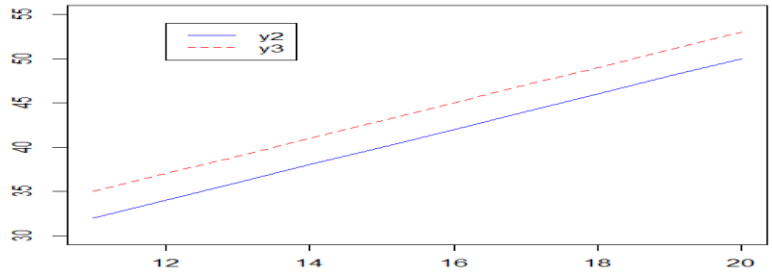
a.
```
> y1<-y[1,]
> y1;
 [1] 11 12 13 14 15 16 17 18 19 20
```

b.
```
> y2<-y[2,]
> y2;
 [1] 32 34 36 38 40 42 44 46 48 50
```

c.
```
> y3<-y[3,]
> y3;
 [1] 35 37 39 41 43 45 47 49 51 53
```

Plot (y2 vs y1) AND (y3 vs Y1)

```
> plot(y1,y2,type="l",lty=1,ylim=c(12,55),xlab="y1",ylab="y2 and y3",col="red")
> lines(y1,y3,type="l",lty=2,col="blue")
> legend(12,54,c("y2","y3"),lty=c(1,2),col=c("red","blue"))
```



Question:

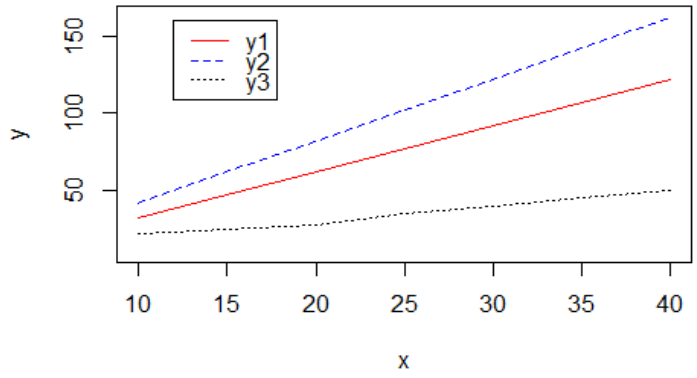
X1	10	15	20	25	30	35	40
Y1	32	47	62	77	92	107	122
Y2	42	62	82	102	122	142	162
Y3	22	25	28	35	40	45	50

Question	Answer

Plot X1 vs Y1 and X1 vs Y2 and X1 vs Y3

```
> x1<-c(seq(10,40,by=5)); x1;
[1] 10 15 20 25 30 35 40
> y1<-c(seq(32,122,by=15));y1;
[1]  32  47  62  77  92 107 122
> y2<-c(seq(42,162,by=20));y2;
[1]  42  62  82 102 122 142 162
> y3<-c(seq(22,28,3),seq(35,50,5));y3;
[1] 22 25 28 35 40 45 50

> plot(x1,y1,xlab="x",ylab="y",ylim=c(10,163),type="l",lty=1,col="red")
> lines(x1,y2,type="l",lty=2,col="blue")
> lines(x1,y3,type="l",lty=3,col="black")
> legend(12,160,c("y1","y2","y3"),lty=c(1,2,3),col=c("red","blue","black"))
```



Question:

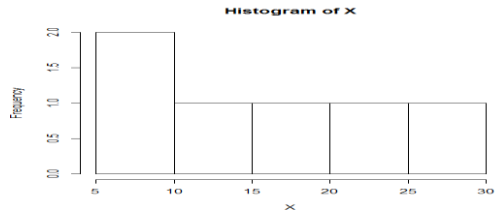
X =	5	10	15	20	25	30

1. Draw a histogram of X.
2. Select randomly 50 observations from the normal distribution. And draw a histogram.

Answer: (1)

```
> x<-c(seq(5,30,5));x;
[1]  5 10 15 20 25 30
> hist(x,xlab="x",ylab="frequency",main="histogram of x",col="white")
```

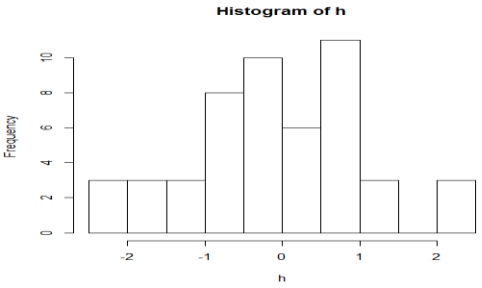
histogram:



(2)

```
> h<-rnorm(50,0,1);
> hist(h,xlab="x",ylab="frequency",main="histogram of x",col="white")
```

```
> rnorm(50,0,1);
 [1] -0.11130789 -0.40542787 -0.42460583  1.25244153  0.70227519 -1.47094950
 [7]  0.91229153  0.48633228 -0.38609907  0.26436717 -1.36762175 -1.88798894
[13] -0.62005678  0.66074346 -0.23921817 -1.46837583  1.01331938  0.34734130
[19] -0.88752247 -0.84400800  0.06816585  0.48370105 -1.10597984  1.42603190
[25] -0.32120160  0.73018597  2.87167526  0.13299202 -2.01877921  1.74750024
[31] -1.29720924 -2.47083089 -0.69523862 -0.84988128  1.42624112 -0.63671031
[37] -0.23320606  0.34525283 -1.41698916  0.45780691  0.54642141  1.60333468
[43] -0.12729823  0.79862332 -0.23757612 -0.34597320  1.13990007  0.52575021
[49] -0.45566704  0.37870277
```



Sampling distribution:

Question: Draw a random sample of five number from the set of 1:40;
# without replacement. 	
## with replacement. 	
	`> sample(1:40,5);`
	`[1] 13 8 37 7 38`
	`> sample(1:40,5,replace=T);`
	`[1] 2 9 18 18 22`

Question : simulate 10 coin tosses .where head (H) and tail(T) have chance .	`> sample(c("H","T"),10,replace=T);` `[1] "H" "H" "H" "H" "T" "T" "T" "H" "T" "H"`

Question: simulate 90% chance of success and 10% chance of failure of a random experiment.	`> sample(c("Success","Failure"),10,replace=T,prob=c(0.9,0.1));`
	`[1] "Success" "Success" "Success" "Success" "Success" "Success" "Success"`
	`[8] "Success" "Failure" "Success"`

Question : how many ways you choose give numbers out of 40.	`> choose(40,5);` `[1] 658008`

Probability distribution:

1. X~binomial (20,0.3)

Question	Answer
a. P(x=3)	`> dbinom(3,20,0.3);` `[1] 0.07160367`
b. P(x<=3)	`> pbinom(3,20,0.3);` `[1] 0.1070868`
c. P(3<x<7)	`> pbinom(6,20,0.3)-pbinom(3,20,0.3);` `[1] 0.500923`

d. Draw a random sample (n=100) for binomial (20,0.3)	`> rbinom(100,20,0.3);`
	`[1] 4 4 3 5 4 8 3 7 5 6 5 5 2 6 4 8 3 8 3 7 9 6 5 3 7`
	`[26] 6 6 9 8 10 5 4 6 6 9 5 9 5 2 3 8 8 2 6 4 5 8 7 10 8`
	`[51] 7 3 4 6 9 6 7 7 6 6 5 8 7 5 7 3 5 4 6 5 9 4 6 5 3`
	`[76] 8 7 5 8 4 5 8 7 5 6 11 9 6 8 5 5 7 2 7 11 10 8 7 5 9`

2. X~poisson (ʹY=7.2)

Question	Answer

a. P(x=3)	`> dpois(3,7.2);` `[1] 0.04644361`
b. P(x<=2)	`> ppois(2,7.2);` `[1] 0.02547351`
c. P(X>7)	`> 1-ppois(7,7.2);` `[1] 0.4310588`
d. P(x>=9)	`> 1-ppois(8,7.2);` `[1] 0.2973317`
e. P(3<x<9)	`> ppois(8,7.2)-ppois(3,7.2);` `[1] 0.6307511`
f. P(3<X<=9)	`> ppois(9,7.2)-ppois(3,7.2);` `[1] 0.7377328`
g. P(3<=x<=9)	`> ppois(9,7.2)-ppois(2,7.2);` `[1] 0.7841764`

1. Let x~N(0,1) find P(X>3).

Answer:
```
> 1-pnorm(3,0,1);
[1] 0.001349898
```

2. X follows normal distribution with mean 35 and standard deviation 6.

** find P(X>42)

Answer:
```
> 1-pnorm(42,35,6);
[1] 0.1216725
```

3. X~N(20,4)

a. Find f(22)	`> dnorm(22,20,2);` `[1] 0.1209854`
b. Find P(16<X<24)	`> pnorm(24,20,2)-pnorm(16,20,2);` `[1] 0.9544997`
c. Find b such that P[X<b]=0.5	`> qnorm(.5,20,2);` `[1] 20`

d. Draw a random sample of size 30 and create vector.	`> rnorm(20,4,30);`
	`[1] 14.9677181 -2.1367220 -2.7929748 0.2254493 27.4877998 -19.8568804`
	`[7] 9.9052908 22.8252161 -10.7558166 42.0824110 24.2301196 -10.6006485`
	`[13] 53.1432210 -7.8589988 7.1917119 -36.4636384 6.9928317 51.0650659`
	`[19] 22.9343899 -33.4730797`

Writing my own function :

<table>
<tr><th>Question</th><th>Answer</th></tr>
<tr><td>1. F(X)=x^2+2x+3</td><td>

```
> m<-function(x)
+ {
+       x*x+2*x+3
+ }
> x<-1:100
> m(x)
  [1]    6   11   18   27   38   51   66   83  102  123  146  171
 [13]  198  227  258  291  326  363  402  443  486  531  578  627
 [25]  678  731  786  843  902  963 1026 1091 1158 1227 1298 1371
 [37] 1446 1523 1602 1683 1766 1851 1938 2027 2118 2211 2306 2403
 [49] 2502 2603 2706 2811 2918 3027 3138 3251 3366 3483 3602 3723
 [61] 3846 3971 4098 4227 4358 4491 4626 4763 4902 5043 5186 5331
 [73] 5478 5627 5778 5931 6086 6243 6402 6563 6726 6891 7058 7227
 [85] 7398 7571 7746 7923 8102 8283 8466 8651 8838 9027 9218 9411
 [97] 9606 9803 10002 10203
```

</td></tr>
<tr><td>2. Write and an R function that will take a vector (x,say) as input and calculate .
a. $\sum x^3$/n-1
b. $\sum(Xi - x)^3$/n-1</td><td>a.

```
k<-function (x)
{
sum(x^3)/(length(x)-1)
}
x<-1:10
k(x)
Output:
 > k(x)
 [1] 336.1111
```

b.

```
h<-function (x)
{
sum((x-mean(x))^3)/(length(x)-1)
}
x<-1:10
h(x)
Output:
> h(x)
 [1] 0
```

</td></tr>
<tr><td>3. write a R function that will take a vector X and a constant C as argument.
• a.Find $\sum x^2$
• b.If $\sum x^2$>C ,return 1 or 0.</td><td>a.

```
> f<-function(x)
+ {   sum(x^2)
+ }
>   x<-1:10
>   f(x)
 [1] 385
```

b.

```
m<-function(x,c)
{
d<-sum(x^2)
v<-0
if(d>c){v<-1}
v
}
m(x,1000000000)

Output:
 [1] 0
```

</td></tr>
</table>

4. Suppose you have a p*q dimensional matrix write a R function that will find the column such as vector.

```
> x<-matrix(1:100,nrow=10)
> abc<- function(x)
+ {
+ csum<- c(rep(0,ncol(x)))
+ for(i in 1:ncol(x))
+ {
+ csum[i]<-sum(x[,i])
+ }
+ csum
+ }
> abc(x)
 [1]  55 155 255 355 455 555 655 755 855 955
```

Some other topic:

Legal Operator: Suppose U=(1,2,3,4,5,6,7,8,9,10).

1. Calculate sum of U when (u>4).
2. Replace the value of U by 10 when U<4 and calculate the sum of U.
3. Replace U<4 or U>8 by 10 and calculate sum of U.
4. Replace not equal 2 by 6 and calculate sum of U.

Solution 1:
Code:

```
u<-1:10
sum(u[u>4])
```

Output: 45

Solution 2:
Code:

```
u<-1:10
u[u<4]=10;u
sum(u)
```

Output: 79

Solution 3:
Code:

```
u<-1:10
u[u<=4 | u>=8]=0;u
sum(u)
```

Output: 18

Solution 4:
Code:

```
u<-1:10
u[u!=2]=6;u
sum(u)
```

Output: 56