SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. Generate the following matrix

1 6 21 26
2 7 22 27
3 8 23 28
4 9 24 29
5 10 25 30

(i)Find the mean of each row of the above matrix.
(ii)Compute the median
(iii)Compute the sum of first 2 columns.

2. Write a program for creating a pie-chart in R using the input vector(21,62,10,53). Provide labels for the chart as 'London', 'New York', 'Singapore', 'Mumbai'. Add a title to the chart as 'city pie-chart' and add a legend at the top right corner of the chart.

3. (i) Write a function in R programming to find a factorial of a given number.
(ii) Find sum of natural numbers up-to 10, without formula using loop statement.

4. For this exercise, use the (built-in) dataset Titanic.

a. Draw a Bar chart to show details of "Survived" on the Titanic based on passenger Class
b. Modify the above plot based on gender of people who survived
c. Draw histogram plot to show distribution of feature "Age"

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. Write R program to find the given number is Prime or not?

2. Write a program for creating a box plot using the vectors H=c(7,12,28,3,41) and M=c("mar", "apr", "may", "jun", "jul"). Add a title to the chart as "Revenue chart".

3. Load dataset named ChickWeight,
   (i).Order the data frame, in ascending order by feature name "weight" grouped by feature "diet" and Extract the last 6 records from order data frame.
   (ii).a Perform melting function based on "Chick", "Time", "Diet" features as ID variables
   b. Perform cast function to display the mean value of weight grouped by Diet
   c. Perform cast function to display the mode of weight grouped by Diet
4. Write a R program to create an array of two 3x3 matrices each with 3 rows and 3 columns from two given two vectors. Print the second row of the second matrix of the array and the element in the 3rd row and 3rd column of the 1st matrix.

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES CES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. Write R Program To Find The Given Number Is Positive Or Negative.
2. Write a R code for histogram charts for the below given age Attribute. Age : 9,13,21,8,36,22,12,41,31,33,19. Add a title to the chart as 'Age' and add color, border, x and y limits.
3. a. Melt 'airquality' data set which inbuild dataset in 'R' and display as a long – format data?

b. Melt air quality data and specify month and day to be "ID variables"?

c. Cast the molten 'airquality' data set.

d. Use cast function appropriately and compute the average of Ozone, Solar, Wind and temperature per month?

e. Create a boxplot for ozone reading of 'airquality' dataset. Add title, label and color.

4. Write a R program:
   - i) Sort a vector in ascending and descending order
   - ii) To find sum, mean and product of vectors
   - iii) To create a sequence of numbers from 20 to 50 and find the mean of numbers from 20 to 60 and sum of numbers from 51 to 91.
   - iv) To calculate the area of Rectangle.

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. A student recorded his/her scores on weekly R programming quizzes that were marked out of a possible 10 points. His/Herscores were as follows:

   8, 5, 8, 5, 7, 6, 7, 7, 5, 7, 5, 5, 6, 6, 9, 8, 9, 7, 9, 9, 6, 8, 6, 6, 7

   What is the mean, median and mode of his/her scores on the weekly R programming quizzes?
2. The merchant sale a products such as tee-shirt, pant, jeans etc., if he sale the above the product quantities between 20 and 30, then the print "Average day". If quantity is > 30 then print "What a great day!", otherwise "Not enough for today". Write the R code using if-else-if control statements.
3. Create R matrix x=3,2,4,5 and y= 6,7,5,4. Find the following and check ur answers in R
   a. x<-matrix (3 2 4 5, 2:2)
   b. x%*%y
   c. y*y
   d. 2*x
4. i) Describe how histogram charts are created in R. Create a histogram chart for the below given age attribute.

   Age : 5,45,23,30,33,32,34,35,42,41,28,29

   ii) Create a 3D Pie Chart for the dataset "political Knowledge" with suitable labels and colour.

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. Write R Program To Print The Sequence Of Numbers (1,2,….,10) Using Repeat loop.
2. Write a R program to create inner, outer, left, right join(merge) from given two data frames.
   df1 = data.frame(numid = c(12, 14, 10, 11))
   df2 = data.frame(numid = c(13, 15, 11, 12))
3. . The table below shows one year of marketing spend and company sales by month.

Month	1	2	3	4	5	6	7	8	9	10	11	12
Spends	1000	4000	5000	4500	3000	4000	9000	11000	15000	12000	7000	3000
Sales	9914	40487	54324	50044	34719	42551	94871	118914	158484	131348	78504	36284

Create a regression model to show the amount of sales(Sales) based on the how much the company spends (Spends) in advertising. Predict the Sales if Spend=13500.

4. For this exercise, use the (built-in) dataset Titanic.
   a. Draw a Bar chart to show details of “Survived” on the Titanic based on passenger Class
   b. Modify the above plot based on gender of people who survived
   c. Draw histogram plot to show distribution of feature “Age”

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours Marks: 100

1. Write a R program to create a matrix - m(c(1:16), nrow=4,byrow = TRUE) taking a given vector of numbers as input and define the column and row names. Access the element at 3rd column and 2nd row, only the 3rd row and only the 4th column of a given matrix.
2. Using linear regression analysis establish a relationship between height and weight of a person using the input vector given below.
   # values of height -> 151, 174, 138, 186, 128, 136, 179, 163, 152, 131
   # values of weight. -> 63, 81, 56, 91, 47, 57, 76, 72, 62, 48
3. Explore the USArrests dataset, contains the number of arrests for murder, assault, and rape for each of the 50 states in 1973. It also contains the percentage of people in the state who live in an urban area.
   (i) a. Explore the summary of Data set, like number of Features and its type. Find the number of records for each feature. Print the statistical feature of data
   b. Print the state which saw the largest total number of rape
   c. Print the states with the max & min crime rates for murder
4. (i) Write a R program to extract the five of the levels of factor created from a random sample from the LETTERS (Part of the base R distribution.)
   (ii)Write R function to find the range of given vector. Range=Max-Min
   Sample input, C<-(9,8,7,6,5,4,3,2,1),
   output=8

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING

PROGRAMME MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. Write the R CODE for the loop next and break statements.
2. Create a data frame for the given table.

Name	Score	Attempts	Qualify
Shan	12.5	1	Yes
Ash	9.0	NA	No
Malu	16.5	2	Yes
Viji	12.5	NA	No

Write a R program to extract 2rd and 4th rows with 1st and 3rd columns from a given data frame.

3. For this exercise, use the (built-in) dataset Titanic.

a. Draw a Bar chart to show details of "Survived" on the Titanic based on passenger Class

b. Modify the above plot based on gender of people who survived

c. Draw histogram plot to show distribution of feature "Age"

4. a. Create a 6 × 10 matrix of random integers chosen in the range of from 1:10

b. Find the number of entries in each row which are greater than 4.

c. Which rows contain exactly two occurrences of the number 7?

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. Write R code to create a bar chart where Bike is assigned red , car is assigned yellow , bus is assigned blue , auto is assigned black , and train is assigned white. Use respective parameters such xlab, ylab, main, labels, color etc.,
2. Create the vector with following numbers. v <- c(2,1,2,3,1,2,3,4,1,5,5,3,2,3)
   Apply the mean, median, and mode in the above vector using R.
3. a. Write a program for creating a pie-chart in R using the input vector (21,62,10,53). Provide labels for the chart as 'London', 'New York', 'Singapore', 'Mumbai'. Add a title to the chart as 'city pie-chart' and add a legend at the top right corner of the chart.
   b. Write a program for creating a bar chart using the vectors H=c(7,12,28,3,41) and M=c("mar", "apr", "may", "jun", "jul"). Add a title to the chart as "Revenue chart"
4. a. Write suitable R code to compute the mean, median, mode of the following values
   c(90, 50, 70, 80, 70, 60, 20, 30, 80, 90, 20,75,70,10,60,70,85,95,55,15)
   b. Write R code to find 2nd highest and 4th Lowest value of above problem

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. Create First Dataset with variables -> Surname, nationality
Create Second Dataset with variables -> Surname, movies
The common key variable is surname. How to merge both data and check if the dimensionality is 7x3.

2. Write a R program to call the (built-in) dataset airquality. Check whether it is a data frame or not? Order the entire data frame by the first and second column.

3. Load dataset named ChickWeight,
i.Order the data frame, in ascending order by feature name "weight" grouped by feature "diet" and Extract the last 6 records from order data frame.
(ii).a Perform melting function based on "Chick", "Time", "Diet" features as ID variables
b. Perform cast function to display the mean value of weight grouped by Diet
c. Perform cast function to display the mode of weight grouped by Diet
(iii)a. Create Box plot for "weight" grouped by "Diet"
b. Create a Histogram for "weight" features belong to Diet- 1 category
c. Create Scatter plot for " weight" vs "Time" grouped by Diet

4. (i)Write R function to find the range of given vector. Range=Max-Min
Sample input, C<-(9,8,7,6,5,4,3,2,1), output=8
(ii)Wirte the R function to find the number of vowels in given string
Sample input c<- "matrix", output<-2

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. Write a R program to add new row(s) to an existing data frame.

2. Write an R program to create an array with three columns, three rows, and two "tables", taking two vectors as input to the array. Print the array and perform the following.
   a) A+B b) A-B c) t(A)
3. Randomly Sample the iris dataset such as 80% data for training and 20% for test and create Logistics regression with train data,use species as target and petals width and length as feature variables ,Predict the probability of the model using test data, Create Confusion matrix for above test model
4. a. Create a data frame based on below table. Month

Month	1	2	3	4	5	6	7	8	9	10	11	12
Spends	1000	4000	5000	4500	3000	4000	9000	11000	15000	12000	7000	3000
Sales	9914	40487	54324	50044	34719	42551	94871	118914	158484	131348	78504	36284

   b. Create a regression model for that data frame table to show the amount of sales(Sales) based on the how much the company spends (Spends) in advertising
   c. Predict the Sales if Spend=13500

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. A student recorded his/her scores on weekly R programming quizzes that were marked out of a possible 10 points. His/Herscores were as follows:
   8, 5, 8, 5, 7, 6, 7, 7, 5, 7, 5, 5, 6, 6, 9, 8, 9, 7, 9, 9, 6, 8, 6, 6, 7
   What is the mean, median, mode of his/her scores on the weekly R programming quizzes?
2. Following table contains name and score of the cricket players.

	name	score
1	Raina	67
2	Bravo	56
3	Dhoni	87
4	Virat	91

a) Create a data frame for the above table1.
b) Write a R program to add a new column in a given data frame
c) Find the dimensions and structure of the data frame.

3. Create a 6 × 10 matrix of random integers chosen in the range of from 1:10
   b. Find the number of entries in each row which are greater than 4.
   c. Which rows contain exactly two occurrences of the number 7?

4. a. Write a program for creating a pie-chart in R using the input vector (21,62,10,53). Provide labels for the chart as 'London', 'New York', 'Singapore', 'Mumbai'. Add a title to the chart as 'city pie-chart' and add a legend at the top right corner of the chart.
   b. Write a program for creating a bar chart using the vectors H=c(7,12,28,3,41) and M=c("mar", "apr", "may", "jun", "jul"). Add a title to the chart as "Revenue chart"

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. Generate the following matrix

1	6	21	26
2	7	22	27
3	8	23	28
4	9	24	29
5	10	25	30

(i)find the mean of each row of the above matrix .
(ii)compute the median
(iii)compute the sum of first 2 columns.

2. Using linear regression analysis establish a relationship between height and weight of a person using the input vector given below.

values of height -> 151, 174, 138, 186, 128, 136, 179, 163, 152, 131
values of weight. -> 63, 81, 56, 91, 47, 57, 76, 72, 62, 48

3. . Randomly Sample the iris dataset such as 80% data for training and 20% for test and create Logistics regression with train data,use species as target and petals width and length as feature variables ,Predict the probability of the model using test data, Create Confusion matrix for above test model
4. For this exercise, use the (built-in) dataset Titanic.
   a. Draw a Bar chart to show details of "Survived" on the Titanic based on passenger Class
   b. Modify the above plot based on gender of people who survived
   c. Draw histogram plot to show distribution of feature "Age"

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. Create First Dataset with variables -> Surname, nationality
   Create Second Dataset with variables -> Surname, movies
   The common key variable is surname. How to merge both data and check if the dimensionality is 7x3.
2. Consider the below vector x
   x <- c(12,7,3,4.2,18,2,54,-21,8,-5,NA)
   a. Find mean, median for the above "x" vector variable.
   b. Apply the trim function to remove the negative values and find the mean value
   c. Remove the NA value and find the mean value.
3. Load dataset named ChickWeight,
   i.Order the data frame, in ascending order by feature name "weight" grouped by feature "diet" and Extract the last 6 records from order data frame.
   (ii).a Perform melting function based on "Chick", "Time", "Diet" features as ID variables
   b. Perform cast function to display the mean value of weight grouped by Diet
   c. Perform cast function to display the mode of weight grouped by Diet
4. Write suitable R code to compute the mean, median ,mode of the following values
   c(90, 50, 70, 80, 70, 60, 20, 30, 80, 90, 20,75,70,10,60,70,85,95,55,15)
   b. Write R code to find 2nd highest and 4th Lowest value of above problem

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. How break control statement works. Draw the flowchart and write the code for the below output using R

   [1] "WELCOME" "CSE"
   [1] "WELCOME" "CSE"
   [1] "WELCOME" "CSE"
   [1] "WELCOME" "CSE"

2. Write a program for creating a Histogram in R using the input vector(v←c(9,13,21,8,36,22,12,41,31,33,19).
   Use respective parameters such as xlim,ylim, xlab,ylab and main

3. a. Write a program for creating a pie-chart in R using the input vector (21,62,10,53). Provide labels for the chart as 'London', 'New York', 'Singapore', 'Mumbai'. Add a title to the chart as 'city pie-chart' and add a legend at the top right corner of the chart.
   b. Write a program for creating a bar chart using the vectors H=c(7,12,28,3,41) and M=c("mar", "apr", "may", "jun", "jul"). Add a title to the chart as "Revenue chart"

4. The table above shows one year of marketing spend and company sales by month.

Month	1	2	3	4	5	6	7	8	9	10	11	12
Spends	1000	4000	5000	4500	3000	4000	9000	11000	15000	12000	7000	3000
Sales	9914	40487	54324	50044	34719	42551	94871	118914	158484	131348	78504	36284

Create a regression model to show the amount of sales(Sales) based on the how much the company spends (Spends) in advertising. Predict the Sales if Spend=13500

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. Consider A=matrix(c(2,0,1,3),ncol=2) and B=matrix(c(5,2,4,1),ncol=2) and find the following using R.

a) Addition of two matrix
b) Subtraction of two matrix
c) Multiplication of A and B
d) Transpose of matrix
e) Modulus of A and B

2. Write a program for creating a pie-chart in R using the input vector (21, 62, 10, 53).Provide labels for the chart as 'London', 'New York','Singapore','Mumbai'.Add a title to the chart as 'city pie-chart' and add a legend at the top right corner of the chart.
3. Randomly Sample the iris dataset such as 80% data for training and 20% for test and create Logistics regression with train data,use species as target and petals width and length as feature variables ,Predict the probability of the model using test data, Create Confusion matrix for above test model
4. Suppose you track your commute times for two weeks (10 days) and you find the following times in minutes 17 16 20 24 22 15 21 15 17 22 Enter this into R
   a. create function "maxi" to find the longest commute time, the function "avger" to find the average and the function "mini" to find the minimum.
   b. Oops, the 24 was a mistake. It should have been 18. How can you fix this? Do so, and then find the new average.
   c. How many times was your commute 20 minutes or more?

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. Suppose that the data for analysis includes the attribute age. The age values for the data tuples are (in increasing order) 13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70. Find the mean, median, mode using R CODE.
2. Using linear regression analysis establish a relationship between height and weight of a person using the input vector given below.
   # values of height -> 151, 174, 138, 186, 128, 136, 179, 163, 152, 131
   # values of weight. -> 63, 81, 56, 91, 47, 57, 76, 72, 62, 48
   Predict the weight of a person with height 170. Visualize the regression graphically
3. Explore the USArrests dataset, contains the number of arrests for murder, assault, and rape for each of the 50 states in 1973. It also contains the percentage of people in the state who live in an urban area.
   (i) a. Explore the summary of Data set, like number of Features and its type. Find the number of records for each feature. Print the statistical feature of data
   b. Print the state which saw the largest total number of rape
   c. Print the states with the max & min crime rates for murder
   (ii). a. Find the correlation among the features
   b. Print the states which have assault arrests more than median of the country
   c. Print the states are in the bottom 25% of murder
4. Write a R program to generate the Fibonacci Series

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. Write R program to print the sequence of numbers (1,2,….,10) using repeat loop.
2. Create a Data frame as shown below

ID	Items	Store	Price
110	book	TRUE	2.5
220	pen	FALSE	8.0
330	textbook	TRUE	10.0
440	Color pen	FALSE	25.0

   a. Write a R program to extract specific column from a data frame using column name.
   b. Write a R program to extract first two rows from a given data frame.
3. Randomly Sample the iris dataset such as 80% data for training and 20% for test and create Logistics regression with train data,use species as target and petals width and length as feature variables ,Predict the probability of the model using test data, Create Confusion matrix for above test model
4. Load dataset named ChickWeight,
   i.Order the data frame, in ascending order by feature name “weight” grouped by feature “diet” and Extract the last 6 records from order data frame.
   (ii).a Perform melting function based on “Chick", "Time", "Diet" features as ID variables
   b. Perform cast function to display the mean value of weight grouped by Diet
   c. Perform cast function to display the mode of weight grouped by Diet
   (iii)a. Create Box plot for “weight” grouped by “Diet”
   b. Create a Histogram for “weight” features belong to Diet- 1 category
   c. Create Scatter plot for “ weight” vs “Time” grouped by Diet

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. Following table contains name and score of the cricket players.

	name	score
1	Raina	67
2	Bravo	56
3	Dhoni	87
4	Virat	91

   a) Create a data frame for the above table.
   b) Write a R program to add a new column in a given data frame
   c) Find the dimensions and structure of the data frame.
2. Write a program for creating a bar chart using the vectors H=c(7,12,28,3,41) and M=c("mar", "apr", "may", "jun", "jul"). Add a title to the chart as "Remainder chart".
3. Using linear regression analysis establish a relationship between height and weight of a person using the input vector given below.
   # Values of height -> 151, 174, 138, 186, 128, 136, 179, 163, 152, 131
   # Values of weight -> 63, 81, 56, 91, 47, 57, 76, 72, 62, 48
   Predict the weight of a person with height 170 and Visualize the regression graphically
4. a. Create a 3x4 matrix with 12 random numbers between 1-100; have the matrix be filled our row-by-row, instead of column-by-column. Name the columns of the matrix uno, dos, tres, cuatro, and the rows x, y, z. Scale the matrix by 10 and save the result.
   b. Extract the column called "uno" as a vector from the original matrix and save the result
   c. Extract the row called 'y' as a vector from the original matrix and print the sum of the vector.

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examine**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. Write R Program To Find The Given Number Is Positive Or Negative.
2. Write an R program to create an array with three columns, three rows, and two "tables", taking two vectors as input to the array. Print the array and perform the following.
   a) A+B b) A-B c) t(A)
3. a. Write a program for creating a pie-chart in R using the input vector (21,62,10,53). Provide labels for the chart as 'London', 'New York', 'Singapore', 'Mumbai'. Add a title to the chart as 'city pie-chart' and add a legend at the top right corner of the chart.
   b. Write a program for creating a bar chart using the vectors H=c(7,12,28,3,41) and M=c("mar", "apr", "may", "jun", "jul"). Add a title to the chart as "Revenue chart"
4. The table above shows one year of marketing spend and company sales by month.

Month	1	2	3	4	5	6	7	8	9	10	11	12
Spends	1000	4000	5000	4500	3000	4000	9000	11000	15000	12000	7000	3000
Sales	9914	40487	54324	50044	34719	42551	94871	118914	158484	131348	78504	36284

Create a regression model to show the amount of sales(Sales) based on the how much the company spends (Spends) in advertising. Predict the Sales if Spend=13500.

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4					
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	GitHub (10)	Viva (10)	Total (100)

Internal Examiner **External Examiner**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. How to expand the data frame by adding rows and columns in data frame for employee data set.
2. Make a histogram for the "AirPassengers "dataset, start at 100 on the x-axis, and from values 200 to 700, make the bins 150 wide
3. Randomly Sample the iris dataset such as 80% data for training and 20% for test and create Logistics regression with train data,use species as target and petals width and length as feature variables ,Predict the probability of the model using test data, Create Confusion matrix for above test model
4. Suppose you track your commute times for two weeks (10 days) and you find the following times in minutes 17 16 20 24 22 15 21 15 17 22 Enter this into R
   a. create function "maxi" to find the longest commute time, the function "avger" to find the average and the function "mini" to find the minimum.
   b. Oops, the 24 was a mistake. It should have been 18. How can you fix this? Do so, and then find the new average.
   c. How many times was your commute 20 minutes or more?

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**

SAVEETHA

INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

COMPUTER SCIENCE AND ENGINEERING PROGRAMME

MODEL PRACTICAL EXAMINATION

FEB – 2023

Subject Code : ITA0443 **Subject Name: STATISTICS WITH R PROGRAMMING FOR REAL TIME PROBLEMS**

Time : 3 Hours **Marks: 100**

1. Write a program for creating a pie-chart in r using the input vector(21,62,10,53). Provide labels for the chart as 'london', 'new york', 'singapore', 'mumbai'. Add a title to the chart as 'city pie-chart' and add a legend at the top right corner of the chart.
2. Write the R CODE for the loop next and break statements with neat flowchart.
3. Randomly Sample the iris dataset such as 80% data for training and 20% for test and create Logistics regression with train data, use species as target and petals width and length as feature variables ,Predict the probability of the model using test data, Create Confusion matrix for above test model.
4. a. Write suitable R code to compute the mean, median, mode of the following values c(90, 50, 70, 80, 70, 60, 20, 30, 80, 90, 20,75,70,10,60,70,85,95,55,15)
   b. Write R code to find 2nd highest and 4th Lowest value of above problem.

PROGRAM 1			PROGRAM 2			PROGRAM 3			PROGRAM 4			GitHub (10)	Viva (10)	Total (100)
Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)	Aim & Program (10)	Debug-ging (5)	Output & Result (5)			

Internal Examiner **External Examiner**