



SAVEETHA SCHOOL OF ENGINEERING
SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES
DEPARTMENT OF DATA SCIENCE
INSTITUTE OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE



MODEL EXAMINATION

Course Code	ITA0444	Course Title	Statistics with R programming for Data Visualization			
Branch	B.Tech- CSE/IT/AI&DS/AI &ML		Year\Academic year		I, II, III, IV Year/2022-23	
Date	22-09-2022		Time	AN	Marks	100

- 1.(i) Write a R program to create a vector of a specified type and length. Create vector of numeric, complex, logical and character types of length 10.
(ii) Write a R program to sort a Vector of you made in ascending and descending order with if else and loop statements.
(iii) Convert vector you made to list and Extract the values at positions 2,3,7 and 9 and print as a single vector.
(iv) Consider $A = \text{matrix}(c(2,0,1,3), \text{ncol}=2)$ and $B = \text{matrix}(c(5,2,4,-1), \text{ncol}=2)$.
a) Find $A + B$
b) Find $A - B$
(v) Find the transpose matrix of A, don't allowed to use built-in function.
- 2.(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.
(iii) create a vector 1:10 and Find a square of each number in the vector using loop statement and save the resultant vector as a list and print the list.
(iv) A group of customer service surveys were sent out at random. Write suitable R code to compute the average, median and mode of the following values
The scores were 90, 50, 70, 80, 90, 70, 60, 20, 30, 80, 90, and 20
3. There is a popular built-in data set in R called "**mtcars**" (Motor Trend Car Road Tests), which is retrieved from the 1974 Motor Trend US Magazine.
(i) Find the dimension of the data set
(ii) Give the statistical summary of the features.
(iii) Find the largest and smallest value of the variable hp (horsepower).

- (iv) Create a Boxplot graph for the relation between "mpg"(miles per gallon) and "cyl"(number of Cylinders) .
 - (v) Generate a multiple regression model using the built-in dataset mtcars. Establish the relationship between "mpg" as a response variable with "disp", "hp" and "wt" as predictor variables. Predict the mileage of the car with dsp=221, hp=102 and wt=2.91.
4. (i) Melt 'airquality' data set which is built-in dataset in 'R' and display as a long-format data?
- (ii) Melt air quality data and specify month and day to be "ID variables"?
 - (iii). Use cast function appropriately and compute the average of Ozone, Solar, Wind and temperature per month ?
 - (iv). Create a boxplot for wind reading of 'airquality' dataset for month of may and july. Add title, label and color.
 - (v) Draw a suitable Plot for to interrupt air quality due to Ozone, Solar, Wind in day-wise in the month of June from 'airquality' dataset



SAVEETHA SCHOOL OF ENGINEERING
SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES
DEPARTMENT OF DATA SCIENCE
INSTITUTE OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE



MODEL EXAMINATION

Course Code	ITA0444	Course Title	Statistics with R programming for Data Visualization		
Branch	B.Tech- CSE/IT/AI&DS/AI &ML	Year\Academic year	I, II, III, IV Year/2022-23		
Date	22-09-2022	Time	AN	Marks	100

1. a. Create a data frame from four given vectors
name = c ('Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas')
score = c (12.5, 9, 16.5, 12, 9, 20, 14.5, 13.5, 8, 19)
attempts = c (1, 3, 2, 3, 2, 3, 1, 1, 2, 1)
qualify = c ('yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes')
b. Write a R program to extract first two rows from a given data frame.
c. Write a R program to extract 3rd and 5th rows with 1st and 3rd columns from a given data frame
d. find the average score with respect to first, second, and third attempts. Don't use any special in build function for this task.
e. Write a R program to create a list containing a vector, a matrix and a list and give names to the elements in the list. Access and print the first and second element of the list
2. a. Write a function called kelvin_to_celsius() that takes a temperature in Kelvin and returns that temperature in Celsius (**Hint:** To convert from Kelvin to Celsius you subtract 273.15).
b. Write an R program to print "R programming Lab" 10 times using for, while and repeat until.
c. Write an R program to generate Fibonacci sequence using Recursion in R.
d. Find the L.C.M of two numbers entered by the user by

creating a user-defined function.

e. Write a function to find a Mode in a given vector and test with sample vector contains following value `c(14, 12,17,13,14,13,14.2,18,20,15,14, 16,18,15)`.

3.
 - a. Melt 'airquality' data set which inbuilt 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.
 - 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. 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.
 - c. Visualize the regression graphically.
 - d. Call 'mtcars' which is built in dataset in 'R' and plot distribution of mpg feature. Make x axis range from 10 to 35 and plot title as "More trends in 70's Vehicles".
 - e. Find statistical summary of the 'mtcars' dataset.