

# CBCS SCHEME

USN

--	--	--	--	--	--	--	--	--	--

BDS306C

## Third Semester B.E./B.Tech. Degree Examination, Dec.2023/Jan.2024 Data Analytics with R

Time: 3 hrs.

Max. Marks: 100

Note: 1. Answer any FIVE full questions, choosing ONE full question from each module.

2. M : Marks , L: Bloom's level , C: Course outcomes.

Module – 1			M	L	C
Q.1	a.	Explain with example Environments and Functions.	12	L2	CO1
	b.	Explain with example flow control and loops in R programming.	8	L2	CO1
OR					
Q.2	a.	List and explain different data types in R.	10	L2	CO1
	b.	Explain different steps in Initiating R.	10	L2	CO1
Module – 2					
Q.3	a.	Explain: i) Lists with examples. ii) Data frames with examples.	10	L2	CO2
	b.	Develop a program to create two $3 \times 3$ matrices A and B and perform the following operations i) Transpose of the matrix ii) Addition iii) Subtraction iv) Multiplication.	10	L3	CO2
OR					
Q.4	a.	Explain briefly: i) Factors ii) Strings.	10	L2	CO2
	b.	Develop an R program using functions to find all the prime numbers up to a specified number by the method of sieve of Eratosthenes.	10	L3	CO2
Module – 3					
Q.5	a.	Explain briefly about importing and exporting files.	10	L2	CO3
	b.	Explain with different steps involved in data cleaning and transforming.	10	L2	CO3
OR					
Q.6	a.	Develop a data form in R for storing about 20 employee details. Create a CSV file named "Input CSV" that defines all the required information about the employee such as id, name, salary, start-date, dept. Import into R and do the following analysis. i) Find the total number rows and columns. ii) Find the maximum salary. iii) Retrieve all the details of the employee with maximum salary. iv) Retrieve all the employees working in the IT department. v) Retrieve the employees in the IT department whose salary is greater than 20000 and write these details into another file output CSV".	10	L3	CO3
	b.	Explain briefly with example accessing data bases.	10	L2	CO3

## Module – 4

Q.7	a.	Explain briefly: i) Exploratory data analysis. ii) Scatter plots.	10	L2	CO4
	b.	Explain base graphics and lattice graphics with the help of BOX plots.	10	L2	CO4

## OR

Q.8	a.	Demonstrate the progression of salary with years of experience using a suitable data set (you can create your own dataset). Plot the graph visualizing the best fit line on the plot of the given data points. Plot a curve of actual values vs. predicted values to show their correlation and performance of the model. Interpret the meaning of the slope and y-intercept of the line with respect to the given data. Implement using the lm function. Save the graphs and coefficients in files. Attach the predicted values of salaries as a new column to the original data set and save the data as a new CSV file.	10	L3	CO4
	b.	Explain how histograms are used to base graphics, lattice graphics and ggplot 2 graphics.	10	L2	CO4

## Module – 5

Q.9	a.	Explain briefly basic statistical measures available in R.	10	L2	CO5
	b.	Explain four in-built functions to generate normal distribution in R.	10	L2	CO5

## OR

Q.10	a.	Explain: i) Correlation Analysis ii) Linear Regression.	10	L2	CO5
	b.	Describe about Analysis of Variance (ANOVA).	10	L2	CO5

\*\*\*\*\*