

CBCS SCHEME

USN

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

BDS306C

Third Semester B.E/B.Tech. Degree Examination, June/July 2025

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																	
1	a.	Discuss briefly the basic data types in R with an example.	6	L2	CO1																	
	b.	Outline a R code to check the given numbers are even or odd.	7	L2	CO1																	
	c.	Illustrate the various operations on vectors.	7	L2	CO1																	
OR																						
2	a.	Discuss the various features of R.	6	L2	CO1																	
	b.	Explain the different types of loops in R with an example.	7	L2	CO1																	
	c.	Interpret the R code to find the factorial of a number (Use recursion).	7	L2	CO1																	
Module – 2																						
3	a.	Explain the different operations on matrices with an example.	7	L2	CO1																	
	b.	Show how to code R program to create a data frame which contain details of five employees and display the details.	7	L2	CO1																	
	c.	Develop the R code to concatenate two strings.	6	L3	CO1																	
OR																						
4	a.	Construct the R program to create the system's idea of the current date with and without time. Explain with a syntax.	6	L3	CO1																	
	b.	Define Factor. Summarize how to create a factor and how to access the components of a factor.	7	L2	CO1																	
	c.	Define an array. Explain how to create an array in R with a suitable example.	7	L1	CO1																	
Module – 3																						
5	a.	List the purpose of data cleaning in R with an example.	7	L2	CO2																	
	b.	Explain how to import and export CSV files in R with an example.	6	L2	CO2																	
	c.	Develop R program to create a data frame with following details and do the following operations : <table border="1"><thead><tr><th>Item Code</th><th>Item Category</th><th>Item Price</th></tr></thead><tbody><tr><td>1001</td><td>Electronics</td><td>700</td></tr><tr><td>1002</td><td>Desktop Supplies</td><td>300</td></tr><tr><td>1003</td><td>Office Supplies</td><td>350</td></tr><tr><td>1004</td><td>USB</td><td>400</td></tr><tr><td>1005</td><td>CD drive</td><td>800</td></tr></tbody></table> i) Subset the Data frame and display the details of only those items whose price is greater than or equal to 350 ii) Subset the data frame and display only the items where the category is either "Office Supplies" or "Desktop Supplies" iii) Create another data frame called "item – details" with three different fields itemcode, ItemQty on Hand and Item ReorderLvl and merge the two frames.	Item Code	Item Category	Item Price	1001	Electronics	700	1002	Desktop Supplies	300	1003	Office Supplies	350	1004	USB	400	1005	CD drive	800	7	L3
Item Code	Item Category	Item Price																				
1001	Electronics	700																				
1002	Desktop Supplies	300																				
1003	Office Supplies	350																				
1004	USB	400																				
1005	CD drive	800																				

OR

6	a.	Explain how to import XML and HTML files into R.	7	L2	CO2
	b.	Summarize the concept of accessing various databases from R.	7	L2	CO2
	c.	Discuss with an example : i) How to create a data frame ii) To add the new variables to data frame iii) How to modify a data frame in R?	6	L2	CO2

Module – 4

7	a.	What is Box plot? Explain briefly the importance of Box plot with an example.	7	L1	CO3
	b.	The built-in data set mammals contain data on body weight versus brain weight. Develop R commands to : i) Find the Pearson and Spearman correlation co-efficient. Are they similar? ii) Plot the data using the plot command iii) Plot the logarithm (Log) of each variable and see if that makes a difference.	6	L3	CO3
	c.	Write about Scatter plot and histogram with examples. Explain its importance.	7	L2	CO3

OR

8	a.	Let us use the built-in data set air quality which has Daily air quality measurements in New York, May to September 1973. Develop R program to generate histogram by using appropriate arguments for the following statements : i) Assigning names, using the air quality data set ii) Change colors of the histogram iii) Remove axis and add labels to histogram iv) Change axis limits of a histogram v) Add density curve to the histogram	10	L3	CO3
	b.	Develop a R code to demonstrate various charts using tree datasets for the following : i) Pie charts ii) Line plots iii) Bar plots iv) Box plots.	10	L3	CO3

Module – 5

9	a.	Explain in detail about the Binomial distribution in R.	7	L2	CO4
	b.	What is linear regression? Explain with an example.	6	L2	CO4
	c.	Describe briefly about the ANOVA test with an example.	7	L2	CO4

OR

10	a.	Write short notes on Normal distribution.	10	L1	CO4
	b.	Discuss briefly about the basic statistical measures in R with an example.	10	L1	CO4
