BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI
(END SEMESTER EXAMINATION)

CLASS: BBA **SEMESTER : V**
BRANCH: BBA **SESSION : MO/2025**

SUBJECT: MN437 DATA SCIENCE USING R

TIME: 3 Hours **FULL MARKS: 50**

INSTRUCTIONS:
1. The question paper contains 5 questions each of 10 marks and total 50 marks.
2. Attempt all questions.
3. The missing data, if any, may be assumed suitably.
***4*. Before attempting the question paper, be sure that you have got the correct question paper.**
5. Tables/Data hand book/Graph paper etc. to be supplied to the candidates in the examination hall.

Q.1(a)	Explain how different data types in R influence the kind of operations and functions that can be performed on them. Include examples to support your explanation.	[5]	1	3
Q.1(b)	Describe the role of control structures and looping in R programs. Illustrate how they help in automating repetitive tasks.	[5]	1	2
Q.2(a)	Discuss how measures of central tendency and dispersion together provide a comprehensive understanding of a dataset.	[5]	2	3
Q.2(b)	Explain the importance of skewness and kurtosis in understanding the shape and nature of data distribution in R.	[5]	2	2
Q.3(a)	Compare the usefulness of histograms and box plots in representing one-dimensional data distributions.	[5]	3	4
Q.3(b)	Demonstrate how two-dimensional graphical techniques in R can reveal relationships between variables. Support your explanation with an example	[5]	3	5
Q.4(a)	Illustrate how data frames are manipulated in R to add or modify data. Include an example using R code.	[5]	4	4
Q.4(b)	Discuss how factors are used in R to represent categorical data. Give an example showing how to create and use a factor.	[5]	4	3
Q.5(a)	Discuss the process of building a linear regression model in R, highlighting how model parameters are interpreted.	[5]	5	
Q.5(b)	Compare decision tree and random forest approaches in terms of interpretability, performance, and practical use in prediction tasks	[5]	5	

::::::21/11/2025::::::M