VIT
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

REG.NO.: 23BCE2093

SLOT: C1+TC1

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - I
WINTER SEMESTER 2025-2026

Programme Name & Branch	: B.Tech & Computer Science and Engineering
Course Code and Course Name	: BCSE206L Foundations of Data Science
Faculty Name(s)	: Dr. Kalyanaraman P, Dr. Rajasekhara Babu M, Dr. Nalini N, Dr. Badrinath N
Class Number(s)	: 2056, 2530, 2533, 2535
Date of Examination	: 29-1-2026
Exam Duration	: 90 minutes

Maximum Marks: 50

General instruction(s):

- Answer All Questions
- M - Max mark; CO – Course Outcome; BL – Blooms Taxonomy Level (1 – Remember, 2 – Understand, 3 – Apply, 4 – Analyse, 5 – Evaluate, 6 – Create)
- Course Outcomes: (Type the CO statements covered in this question paper. Use the CO number as per the syllabus copy)
  1. Ability to obtain fundamental knowledge on data science.
  2. Demonstrate proficiency in data analytics.

Q. No	Question	Module	Mark	CO	BL
1.	A hospital wants to reduce patient wait times in its emergency department. It has collected data on patient arrivals, treatment durations, and staff availability. Define data science in this context and describe how it can be applied to optimize hospital operations.	1	10	1	4
2.	Consider the Courses and Enrollments Database Schema : - Courses (CourseID, CourseName, Credits, DepartmentID) - Enrollments (EnrollID, StudentID, CourseID, Marks) (a) Retrieve students who scored above the average marks in a course. (b) List students enrolled in courses with more than 20 credits. (c) Identify students who scored higher than the course average. (d) Display courses along with the number of students enrolled.	2	10	2	3
3.	Consider the Sales Database - Apply only sql window functions, Schema: Sales (SaleID, ProductID, Region, Quantity, Revenue) (a) Write a query to display the ProductID, Region, and the median revenue where the region is 'South'. [4 marks] (b) Write a query to display the SaleID, Region, Revenue, and their rank (based on revenue) within the respective region. [2 marks] (c) Write a query to display the percentile rank of each sale based on revenue. [2 marks] (d) For each sale, display the current SaleID, Region, Revenue, and also show the revenue of the next sale in the same region (ordered by SaleID). [2 marks]	2	10	2	3
4.	A startup wants to hire its first data scientist to analyze user engagement on its mobile app. List the key prerequisites, tools, and skills the candidate should possess, and explain why these are important for solving the startup's problem.	1	6	1	4

Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

REG.NO.:

SLOT: C1+TC1

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - I
WINTER SEMESTER 2025-2026

	(b) Find the right NoSQL database type for the following scenarios: - Store user profiles with varying attributes. - Represent friend connections and recommendations. Justify which NoSQL database type is best suited for each requirement and explain why traditional SQL may not be sufficient	2	4		
5.	A telecom company wants to predict customer churn. It has raw data on call records, billing, complaints, and service usage. Provide at least five sample instances of this data. Explain how you would prepare the data and plan the model before building it, including visualization techniques, business implications, and recommendations for deployment.	3	10	2	4

****