



VIT

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

REG.NO.: 23BCED760

SLOT: C2+TC2

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. K.Santhi Dr. Radhakrishnan Delhibabu, Dr. S. Vinila Jinny, Dr. A. Devipriya
Class Number(s) : 2058, 2532, 2534, 2057
Date of Examination : 29/01/2026
Exam Duration : 90 minutes **Maximum Marks: 50**

General instruction(s):

- Answer All Questions
- Course Outcomes
 1. Ability to obtain fundamental knowledge on data science.
 2. Demonstrate proficiency in data analytics.

Q. No	Question	Module	Marks	CO	BL																																												
1.	A bank wants to minimize credit risk. It has access to customer transaction history, loan repayment records, credit scores, and demographic data. Define Data Science in this context and explain how it can be used to improve loan approval and risk assessment processes.	1	10	1	4																																												
2.	A large retail organization operates through both online platforms and physical stores across multiple regions. The company collects massive volumes of data such as customer purchase transactions, website browsing behavior, loyalty program details, social media feedback, and inventory records. Despite this, the management is unable to derive actionable insights for improving customer retention and optimizing sales strategies. (a) Explain the need for Data Science in the above scenario and how it supports effective business decision-making. (3) (b) Define Data Science and describe the Data Science process that can be applied to analyze the given business problem.(3) (c) Compare Business Intelligence (BI) and Data Science with respect to data types handled, analytical techniques used, and outcomes, in the context of the given scenario.(4)	1	10	2	4																																												
3.	i) An online bookstore maintains the following tables to analyze its sales performance. BOOK Table <table><tr><th>Book_ID</th><th>Title</th><th>Category</th><th>Price</th></tr><tr><td>B101</td><td>Data Science 101</td><td>Technology</td><td>550</td></tr><tr><td>B102</td><td>Python Basics</td><td>Technology</td><td>450</td></tr><tr><td>B103</td><td>Business Trends</td><td>Management</td><td>600</td></tr><tr><td>B104</td><td>SQL Essentials</td><td>Technology</td><td>500</td></tr></table> SALES Table <table><tr><th>Sale_ID</th><th>Book_ID</th><th>Quantity</th><th>City</th></tr><tr><td>S201</td><td>B101</td><td>3</td><td>Chennai</td></tr><tr><td>S202</td><td>B102</td><td>5</td><td>Coimbatore</td></tr><tr><td>S203</td><td>B101</td><td>2</td><td>Chennai</td></tr><tr><td>S204</td><td>B103</td><td>4</td><td>Madurai</td></tr><tr><td>S205</td><td>B104</td><td>6</td><td>Chennai</td></tr></table>	Book_ID	Title	Category	Price	B101	Data Science 101	Technology	550	B102	Python Basics	Technology	450	B103	Business Trends	Management	600	B104	SQL Essentials	Technology	500	Sale_ID	Book_ID	Quantity	City	S201	B101	3	Chennai	S202	B102	5	Coimbatore	S203	B101	2	Chennai	S204	B103	4	Madurai	S205	B104	6	Chennai	2	6	2	3
Book_ID	Title	Category	Price																																														
B101	Data Science 101	Technology	550																																														
B102	Python Basics	Technology	450																																														
B103	Business Trends	Management	600																																														
B104	SQL Essentials	Technology	500																																														
Sale_ID	Book_ID	Quantity	City																																														
S201	B101	3	Chennai																																														
S202	B102	5	Coimbatore																																														
S203	B101	2	Chennai																																														
S204	B103	4	Madurai																																														
S205	B104	6	Chennai																																														



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

	(a) Write an SQL query to display the title and total quantity sold for each book. (b) Write an SQL query to find the total sales amount category-wise. (c) Write an SQL query to display books sold in Chennai only.								
	ii) Differentiate between Document Databases, Wide-column Databases, and Graph Databases with one example use case for each.						4		
4.	student_id	student_name	course_id	test_date	score	2	10	2	3
	101	Arun	C101	2024-01-10	78				
	102	Beena	C101	2024-01-10	85				
	103	Charan	C101	2024-01-10	85				
	101	Arun	C101	2024-02-15	82				
	102	Beena	C101	2024-02-15	88				
	103	Charan	C101	2024-02-15	80				
	104	Divya	C102	2024-01-12	72				
	105	Eshan	C102	2024-01-12	90				
	104	Divya	C102	2024-02-18	75				
	105	Eshan	C102	2024-02-18	92				
	An online learning platform stores students' assessment results in the Student_Scores table shown above. The academic team wants to perform analytical queries on student performance without losing individual test records. a) Write an SQL query using a window function to assign a rank to each student within their course for each test date, such that the highest score gets Rank 1. b) Using a window function, display the average score of each course along with each student's score. c) Write an SQL query to display the previous test score for each student, ordered by test_date. d) Explain why window functions are more suitable than GROUP BY for this scenario.								
5.	A city health department aims to reduce hospital readmissions by applying data analytics to patient data collected from multiple hospitals. Explain how the Data Analytics Lifecycle can be used to address this problem, including the type of data involved, the data preparation process, model planning considerations, suitable visualization techniques, and how the results can be translated into actionable healthcare decisions and deployed within hospital systems.					3	10	2	4
