



VIT

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

REG.NO.:

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

SLOT: C1

Programme Name & Branch : BTech-IT
Course Code and Course Name : Data Mining - BITE312E
Faculty Name(s) : Dr.E.Sathiyamoorthy & Dr. Neelu Khare
Class Number(s) : VL2025260502165 & VL2025260502163
Date of Examination : 29/01/2026
Exam Duration : 90 minutes Maximum Marks: 50

General instruction(s): Answer All Questions

• Course Outcomes

Analyze various real-time problems and design the data visualizations.
Apply the data mining concepts to conduct data pre-processing.

Q. No	Question	M	CO	BL															
1	Suppose a public transportation data warehouse is used to analyse passenger travel patterns. The warehouse consisting of four dimensions Date, Passenger type, Location and Route. There are two measures, Count and Fare where count is the number of passengers travelling and fare is the ticket amount paid by the passenger for a journey on a given date. (a) Draw a star schema diagram for the data warehouse. (b) What specific OLAP operations should be performed to find the total fare collected per route per month and yearly revenue to daily fare collection for a specific route.	10	1	2															
2.	Compare and contrast OLAP and OLTP systems with respect to data model, types of operators used, query complexity and performance requirements. Support your answer with suitable examples.	10	1	1															
3.	A random sample of 395 people was surveyed and each person was asked to report the highest education level they obtained. The data resulted from the survey are summarized in the following table: <table border="1" data-bbox="221 1899 1271 2092"> <tr> <th></th><th>High School</th><th>Bachelors</th><th>Masters</th><th>Ph.D</th></tr> <tr> <td>Female</td><td>60</td><td>54</td><td>46</td><td>41</td></tr> <tr> <td>Male</td><td>40</td><td>44</td><td>53</td><td>57</td></tr> </table> Whether the gender of an individual and the level of education are correlated? Use appropriate correlation measures to state your answer.		High School	Bachelors	Masters	Ph.D	Female	60	54	46	41	Male	40	44	53	57	10	2	3
	High School	Bachelors	Masters	Ph.D															
Female	60	54	46	41															
Male	40	44	53	57															
4.	Consider the below data and normalize by applying following methods: 250, 550, 700, 1150 (a) Min-Max normalization by setting min=0 and max=1. (b) Z-Score Normalization.	10	2	3															
5.	Let the age of patients are given to analyse their BMI index as; 15, 17, 19, 21, 27, 25, 40, 33, 35, 37, 39, 39, 40, 42, 40, 43, 44, 46, 46, 48, 49, 53, 51, 57. Apply Equi-depth binning and smooth the bins by bin means and bin boundaries, let bin depth of 6.	10	2	3															