K. J. Somaiya Institute of Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)

Nov–Dec 2025-26
Program: B.Tech Scheme : III
Examination: TY Semester: V
Course Code: AIC502 and Course Name: Data Warehousing, Mining & Visualization
Date of Exam: 26/11/2025 Duration: 2.5 Hours Max. Marks: 60

Instructions:
(1)All questions are compulsory.
(2)Draw neat diagrams wherever applicable.
(3)Assume suitable data, if necessary.

Q. No.	Question	Max. Marks	CO	BT level
Q 1	**Solve any two questions out of three: (05 marks each)**	**10**		
a)	What are the different characteristics of unstructured data?		CO1	U
b)	Compare: D3.js and Google Charts		CO6	U
c)	Explain in brief box plot with suitable diagram. Construct a box plot for the following data: 12, 5, 22, 30, 7, 36, 14, 42, 15, 53, 25		CO3	Ap
2	**Solve any two questions out of three: (05 marks each)**	**10**		
a)	Define data warehouse. Explain in detail the definition components of data warehouse.		CO2	U
b)	Draw Apriori algorithm flowchart and explain it.		CO4	U
c)	Explain in detail acquiring and visualizing data steps.		CO5	U
Q.3	**Solve any two questions out of three. (10 marks each)**	**20**		
a)	Construct FP tree with minimum support 3 and write frequent patterns generated.		CO4	Ap
b)	Explain in detail text data visualization. Justify with the Examples		CO5	U
c)	Describe the Knowledge Discovery in Databases (KDD) process. What are the key stages involved, and why is each stage important in		CO3	U

TID	Items bought	Ordered frequent Items
100	{f, c, a, d, g, i, m, p}	{f, c, a, m, p}
200	{a, b, c, f, l, m, o}	{ f, c, a, b, m}
300	{b, f, h, j, o, w}	{f, b }
400	{b, c, k, s, p}	{c, b, p }
500	{a, f, c, e, l, p, m, n}	{f, c, a, m, p}

Seat No. :

	data mining?			
Q.4	**Solve any two questions out of three. (10 marks each)**	**20**		
a)	Justify the need of Strategic information. Compare: Operational vs Decision Support Systems		CO1	An
b)	The sugar_sales_data table contains multidimensional data for analyzing sugar product sales (SalesAmount in lakhs)across different years (2023, 2024), quarters, months, product categories (Raw Sugar, Refined Sugar, Brown Sugar, Jaggery), and regions (4 Indian states). Draw data cube for the given Data. Perform OLAP operations for below mentioned queries using SQL. Identify which OLAP operation is performed for the respective query? Draw all OLAP operation diagrams.		CO2	Ap

Year	Quarter	Month	ProductCategory	Region	SalesAmount
2023	Q1	January	Raw Sugar	Maharashtra	2200
2023	Q2	April	Refined Sugar	Uttar Pradesh	2600
2023	Q3	July	Brown Sugar	Karnataka	2100
2023	Q4	October	Jaggery	Tamil Nadu	1900
2024	Q1	February	Refined Sugar	Maharashtra	2800
2024	Q2	May	Raw Sugar	Uttar Pradesh	2500
2024	Q3	August	Jaggery	Karnataka	2300
2024	Q4	November	Brown Sugar	Tamil Nadu	2000

Answer the following queries.

1. What is the total sales amount for each year?
2. What are the total quarterly sales figures for Maharashtra in 2024?
3. Show the sales of all sugar types in the second quarter (Q2) across all regions.
4. Show sales data for Raw Sugar and Jaggery during the first two quarters of 2023 in Karnataka and Uttar Pradesh.
5. Display a table showing sales amounts by product category.

c)	Develop a simple transaction manager that enables users to add financial transactions by entering a description and amount. The application should update the total balance dynamically based on the transactions, distinguishing between income (positive values) and expenses (negative values). Additionally, the transactions should be categorized and displayed in a list format, with income shown in green and expenses in red for better clarity.		CO6	Ap

****

Seat No. :