Name: **Enrolment No:**	

UPES

End Semester Examination, May 2025

Course: Data Visualization **Semester: 2nd**

Program: MCA in Data Science **Time : 03 hrs.**

Course Code: CSDS7003 **Max. Marks: 100**

Instructions:

SECTION A
(5Qx4M=20Marks)

S. No.		Marks	CO
Q 1	Calculate the mean, median, and mode for the dataset: {5, 10, 15, 10, 20, 10, 25} and explain their significance in data analytics	**5**	**CO1**
Q 2	Discuss how scatter plots and regression plots can be used together to interpret relationships between two variables.	**5**	**CO1**
Q 3	What are the key elements that should be included while creating a Data Entry Standards Document (DES).	**5**	**CO1**
Q 4	Write the difference between Descriptive and Prescriptive Analysis on the basis of some characteristics.	**5**	**CO2**
Q 5	Discuss the pros and cons of using static charts vs. dynamic dashboards for executive reporting	**5**	**CO3**
	SECTION B **(4Qx10M= 40 Marks)**		
Q 6	What are the major task in data preprocessing. Explain them briefly.	**10**	**CO1**
Q 7	Discuss the role of outliers in data analysis and calculate the interquartile range (IQR) for the dataset: 3, 5, 7, 10, 12, 15, 18, 21 OR A company wants to analyze employee salaries using box plots. Explain how this helps in data visualization and perform an IQR-based outlier analysis on the salaries: {25k, 30k, 35k, 40k, 50k, 55k, 70k, 80k, 150k}.	**10**	**CO1**
Q 8	What is univariate analysis? Explain with an example using a histogram.	**10**	**CO2**
Q 9	In an election poll analysis scenario, what statistical tests and visualizations would you use to compare voter preferences across regions?	**10**	**CO2**
	SECTION-C **(2Qx20M=40 Marks)**		
Q 10	A healthcare provider wants to analyze patient data and create an interactive report showing: i. Age distribution of patients	**20**	**CO3**

<table>
<tr>
<td></td>
<td>
ii. Diagnosis categories
iii. Hospital readmission rates
iv. Comparison of treatments using hypothesis testing
Design a dashboard using colored pens in your exam sheet for the analysis.
OR
Create a dynamic dashboard to visualize the following dataset:
<table>
<tr><th>Product</th><th>Sales (Jan)</th><th>Sales (Feb)</th><th>Sales (Mar)</th><th>Sales (Apr)</th></tr>
<tr><td>A</td><td>500</td><td>600</td><td>700</td><td>750</td></tr>
<tr><td>B</td><td>400</td><td>450</td><td>470</td><td>500</td></tr>
<tr><td>C</td><td>700</td><td>750</td><td>780</td><td>800</td></tr>
</table>
The dashboard should include line charts, bar charts, and a comparison table.
</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q 11</td>
<td>Propose a complete analytical workflow—from preprocessing to visualization—to uncover meaningful insights after choosing a publicly available dataset.</td>
<td>20</td>
<td>CO1, CO3</td>
</tr>
</table>