[This question paper contains 4 printed pages.]

Your Roll No. 2400356

Sr. No. of Question Paper : 5328 **J**

Unique Paper Code : 2514001202

Name of the Paper : Data Visualization Technique

Name of the Course : **B.Sc. (H) Electronics (GE)**

Semester : II

Duration : 3 Hours Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. There are **seven** questions in all, out of which you have to attempt any **five** questions.
3. **All** questions carry equal marks.
4. **First** Question is Compulsory.

1	a	Explain any one method to handle the outliers.	3
	b	How can dashboards enhance the data analysis? Mention two key features of Plotly dashboards.	3
	c	What is a Gantt chart, and how is it useful in project management?	3
	d	Explain how missing data is handled during data cleaning. Provide an example.	3
	e	Discuss the advantages of using regression plots for predictive analysis.	3
	f	What are advanced chart types in Tableau, and why are they useful for data storytelling?	3
2	a	Provide real-world examples of structured, semi-structured, and unstructured data. Suggest appropriate visualization techniques for each type and justify your choices.	6
	b	Discuss different techniques for handling missing and noisy data in a dataset. How do these techniques impact the accuracy and reliability of visualizations?	6
	c	How does missing data influence the accuracy and effectiveness of visualizations? Demonstrate with an example how improper handling of missing values can lead to misleading insights in a graphical representation.	6
3	a	math_scores = 78, 85, 90, 72, 88, 95, 80, 76, 89, 92 science_scores = 82, 87, 91, 74, 85, 97, 78, 79, 90, 94 For the given data: Plot a **scatterplot** using Matplotlib to show the relationship between Mathematics and Science scores. Label the axes and give an appropriate title.	6
	b	Generate a regression plot to visualize the relationship between **total_bill** and **tip**, using the **sex** column as the hue parameter for **tips dataset**. Interpret how the regression lines differ for male and female customers.	6

P.T.O.

c Create a **waffle chart** using Seaborn where each product category is visually distinguishable. Customize it by: 6

- Assigning different colors to each category
- Displaying a legend
- Adding appropriate labels

data = {'Product': ['Laptop', 'Mobile', 'Tablet', 'Headphones'], 'Sales': [40, 35, 15, 10]}

5 a Differentiate between Matplotlib, Seaborn and Plotly Visualisation tool. Give at least one example of each. 6

b A manufacturing company is analyzing the temperature distribution across a metal sheet during a heating process. The temperature is recorded at specific points along the X and Y axes. 6

X(cm)	Y(cm)	Temperature Range(°C)	Temp(°C)
0	0	20-30	25
0	5	20-30	30
0	10	30-40	35
5	0	20-30	28
5	5	30-40	33
5	10	30-40	38
10	0	40-50	41
10	5	30-40	36
10	10	40-50	42

Describe the steps to create a Sunburst Chart using the temperature data, where each layer represents different temperature ranges categorized by X and Y coordinates. Explain how a Sunburst Chart can provide insights about the relationships in the data.

c You have been provided with a dataset that contains information about the sales performance of a company, including monthly revenue, product categories, and geographical distribution of sales. 6

Month	Product Category	Revenue (Rs.)	Location	Latitude	Longitude
January	Electronics	25000	New Delhi	28.613	77.209
January	Furniture	15000	Mumbai	19.076	72.877
February	Clothing	20000	Bengaluru	12.971	75.594
February	Electronics	30000	Kolkata	22.572	88.363
March	Furniture	18000	Chennai	13.082	80.270
March	Clothing	22000	Hyderabad	17.385	78.486
April	Electronics	35000	Ahmedabad	23.022	72.571
April	Furniture	25000	Pune	18.520	73.856

Design an interactive dashboard using **Plotly** to showcase the following:

i) A line chart for monthly revenue trends.

ii) A map displaying geographical sales data with markers for each location.

5 a You are working with a dataset that contains sales data across various product categories, regions, and time periods. 6

Month	Product Category	Revenue (Rs.)	Region	Year
January	Electronics	25000	North	2023
February	Furniture	15000	East	2023
March	Clothing	20000	South	2023
April	Electronics	30000	North	2023
June	Furniture	18000	West	2023
July	Clothing	22000	North	2023
August	Electronics	35000	East	2023
September	Furniture	25000	South	2023

Explain how you can use cross filtering in Tableau to link visualizations in a dashboard, such as a bar chart for product category revenue and a map showing regional sales distribution.

b Describe the advantages of cross filtering for interactive dashboards and decision-making. Provide an example of how applying a filter on "South Region" impacts other visualizations dynamically. 6

c Explain the key differences between Tableau Desktop and other data visualization tools. What types of data sources can Tableau Desktop connect to, and how does this flexibility benefit users? 6

6 a Using the dataset given: 9

i) Perform a **univariate analysis** of "Marks" using an appropriate Seaborn plot.

ii) Conduct a **bivariate analysis** between "Marks" and "Study_Hours". Visualize and interpret the relationship.

iii) Perform a **multivariate analysis** including "Gender" to analyze how gender influences the relationship between study hours and marks. Use Seaborn's hue parameter to help.

Provide visualizations and insights for each part.

Dataset:

Marks: [88, 76, 92, 85, 67, 73, 90, 81, 77, 69],
Study_Hours: [6, 4, 7, 5.5, 3, 4.5, 6.5, 5, 4, 3.5],
Gender: ['F', 'M', 'F', 'F', 'M', 'M', 'F', 'M', 'F', 'M']

b Create a **Folium map** that uses **circle markers** to represent the **Air Quality Index** of each city. 9

i. Vary the radius and color of the circles based on AQI
ii. Add **popup text** to show city name and AQI
iii. Add a tooltip or legend explaining the color scale

Dataset:
City: ['Mumbai', 'Chennai', 'Kolkata', 'Bengaluru']
Lat: [19.0760, 13.0827, 22.5726, 12.9716]
Lon: [72.8777, 80.2707, 88.3639, 77.5946]
AirQualityIndex: [160, 120, 180, 140]

7 a A dataset contains monthly sales data for three categories: Electronics, Furniture, and Clothing, along with total revenue. 9

i. Explain how bubble plots can help visualize differences in revenue across categories dynamically.
ii. Write Python code using Plotly to create a bubble plot where the size of bubbles reflects revenue.

Month	Product Category	Revenue (Rs.)
January	Electronics	25000
February	Furniture	15000
March	Clothing	20000
April	Electronics	30000
June	Furniture	18000
July	Clothing	22000
August	Electronics	35000
September	Furniture	25000

b Discuss how data storytelling improves stakeholder engagement and aids in better decision-making. Provide an example of how this feature helps convey actionable insights to non-technical audiences. Provide steps to create dashboards in Tableau. 9