B.M.S. College of Engineering, Bengaluru-560019

Autonomous Institute Affiliated to VTU

January / February 2025 Semester End Main Examinations

Programme: B.E.   **Semester: V**

Branch: Artificial Intelligence and Machine Learning   **Duration: 3 hrs.**

Course Code: 24AM5PCDAV   **Max Marks: 100**

Course: Data Analysis and Visualization

Instructions: 1. Answer any FIVE full questions, choosing one full question from each unit.
2. Missing data, if any, may be suitably assumed.

			UNIT - I	CO	PO	Marks
1	a)		Elaborate Univariate data and Multivariate Data with a suitable example for each.	CO1	PO1	**10**
	b)		Describe the key issues and challenges faced in data analysis, along with the best practices to address them effectively.	CO1	PO1	**10**
			OR			
2	a)		i. Explain the key elements of structured data. ii. Differentiate structured data from unstructured data.	CO1	PO1	**10**
	b)		Categorize possible data processing errors that can occur during data collection also examine the cause of the same.	CO1	PO1	**10**
			UNIT - II			
3	a)		The district headquarters aims to analyze student performance data of various schools from multiple spreadsheets. How data wrangling strategies can be applied to consolidate this information effectively? Justify.	CO3	PO3	**10**
	b)		Imagine a large dataset with multiple date formats across different columns. Outline a strategy to identify, parse, and standardize inconsistent date formats across a large dataset. Discuss how this transformation could affect subsequent time-series analysis.	CO3	PO3	**10**
			OR			
4	a)		Analyse and Perform an Exploratory Data Analysis (EDA) on a university student grade dataset containing numerical and categorical data, including demographic attributes like gender and age. Identify patterns between demographics and grade distribution and summarize key insights and trends.	CO3	PO3	**10**

		b)	A telecom company aims to reduce customer churn by analyzing customer interaction and satisfaction data. Conduct a structured exploratory data analysis (EDA) including data cleaning, correlation analysis, and distribution assessment. Identify key factors impacting churn and provide actionable insights for customer retention strategies.	CO3	PO3	**10**
			UNIT - III			
5	a)		Analyze how can the foundational Gestalt principles of perception be applied to develop visually cohesive, intuitive, and aesthetically impactful designs?	CO2	PO3	**10**
	b)		i) When dealing with data that has a hierarchical structure or involves relationships in a parent-child format, how can tree representations be used to effectively illustrate the connections and nested categories within the data with suitable example? ii)Given Points on a 2D Plane: A(1, 2) B(2, 4) C(10, 10) D(11, 11) E(50, 50) F(3, 5) G(12, 12) H(51, 49) I(0, 1) J(52, 51) using proximity principle group these points into clusters.	CO2	PO3	**10**
			OR			
6	a)		Analyze the essential structures of visual mapping and explain their significance in creating effective and meaningful visual representations of data	CO2	PO3	**10**
	b)		i. Sales Manager presenting the quarterly sales performance of a company to a group of stakeholders. How could he use metaphorical visualization to represent the sales data in a way that makes it easy for them to understand, using familiar concepts or imagery? ii. Calculate the total continuity cost between consecutive points for the given path (0,0) (1,1) (2,1.8) (3,3).	CO2	PO3	**10**
			UNIT - IV			
7	a)		Explain the significance of clusters, trees, and graphs in data visualization. How do these methods contribute to understanding relationships in networks? Provide examples from interactive data visualizations in news media.	CO3	PO1	**10**
	b)		Analyze how metaphorical visualizations are used in interactive news media to simplify complex multi-dimensional data. Discuss one case study where this technique is effectively applied to engage readers.	CO3	PO3	**10**
			OR			
8	a)		Discuss the classification of visualization systems based on their interaction techniques. How can interaction design mislead the audience when representing one- or two-dimensional data?	CO3	PO1	**10**

		b)	Compare the challenges and techniques for visualizing textual data versus multi-dimensional numerical data. How can grouping and clustering improve the understanding of large text datasets in news media?	*CO3*	*PO3*	**10**
			UNIT - V			
9		a)	Discuss the challenges involved in visualizing volumetric data such as vector fields and simulation processes. How can visualization techniques like streamlines, glyphs, and isosurfaces be applied to better understand the underlying patterns in this type of data?	*CO1*	*PO3*	**10**
		b)	Discuss specific techniques such as heatmaps, choropleth maps, and 3D terrain models, and provide examples of their application in urban planning or environmental studies.	*CO3*	*PO3*	**10**
			OR			
10		a)	Discuss the advantages and potential drawbacks of collaborative visualizations in team-based environments. How can collaboration platforms be integrated with visualization systems to improve the interpretation and communication of complex data?	*CO1*	*PO3*	**10**
		b)	Discuss the criteria for evaluating the effectiveness of visualizations, especially in the context of complex data such as simulations, GIS, or volumetric data. What factors should be considered to ensure that a visualization is accurate, intuitive, and useful for the intended audience?	*CO3*	*PO3*	**10**

******