|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Roll No. |  |  |  |  |  |  |  |  |  |  |  |  |



 **PRESIDENCY UNIVERSITY**

  **Bengaluru**

|  |
| --- |
| **End - Term Examinations – JANUARY 2025** |
| **Date:** 13 – 01- 2025 **Time:** 09:30 am – 12:30 pm |

|  |  |
| --- | --- |
| **School:** SOCSE | **Program:** B Tech-CAI/CSG/CST |
| **Course Code :** CSE2026 | **Course Name** :DATA HANDLING AND VISUALIZATION |
| **Semester**: V | **Max Marks**:100 | **Weightage**:50% |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CO - Levels** | **CO1** | **CO2** | **CO3** | **CO4** |
| **Marks** | **24** | **24** | **26** | **26** |

**Instructions:**

1. *Read all questions carefully and answer accordingly.*
2. *Do not write anything on the question paper other than roll number.*

**Part A**

|  |
| --- |
| **Answer ALL the Questions. Each question carries 2marks. 10Q x 2M=20M** |
| **1** | Explain 'data transformation' in brief with example. | **2 Marks** | **L2** | **CO1** |
| **2** | Differentiate between a bar chart and a histogram. | **2 Marks** | **L2** | **CO1** |
| **3** | Compare scalar and vector data with example | **2 Marks** | **L2** | **CO2** |
| **4** | Explain the methods to displaying Hierarchical Structures | **2 Marks** | **L2** | **CO2** |
| **5** | Describe one challenge faced when visualizing spatial data.  | **2 Marks** | **L2** | **CO3** |
| **6** | List the advantage of using time-series graphs in streaming visualization | **2 Marks** | **L1** | **CO3** |
| **7** | List the visualization tools to represent temporal data | **2 Marks** | **L1** | **CO3** |
| **8** | List the major challenges of streaming data | **2 Marks** | **L1** | **CO4** |
| **9** | Describe one method used in the analysis of streaming data.  | **2 Marks** | **L2** | **CO4** |
| **10** | List the challenges of streaming data | **2 Marks** | **L1** | **CO4** |

**Part B**

|  |
| --- |
| **Answer the Questions Total 80 Marks.** |
| **11.** | **a.** | Define data. List and explain Data in the Real World and Compare data vs. information | **10 marks** | **L2** | **CO1** |
| **b.**  | List the advantages of data preparation and explain key steps in data preparation.  | **10 marks** | **L2** |
| **or** |
| **12.** | **a.** | Explain the role of data cleaning in Exploratory data analysis. Describe the different ways of cleaning the data | **10 Marks** | **L2** | **CO1** |
| **b.**  | Explain different visualization techniques using matplotlib library and iris dataset (minimum of 5 visualization methods) | **10 marks** | **L2** |
|  |  |  |  |  |  |
| **13.** | **a.** | Explain in detail data wrangling using own dataset& List the disadvantage of data with high dimensionality.  | **10 Marks** | **L2** | **CO2** |
| **b.** | Explain how PCA helps in reducing dimensions. Use the python code to reduce the dimensions of breast cancer data to predict 'Benign' or 'Malignant' | **10 marks** | **L3** |
| **or** |
| **14.** | **a.** | Explain Multivariate Data. Compare multivariate data with multidimensional with example.  | **10 Marks** | **L2** | **CO2** |
| **b.**  | Use the python code to visualize different types of Multivariate Data by considering appropriate data. | **10 Marks** | **L3** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **15.** | **a.** | Determine how geospatial data is a spatial type of spatial data. Compare raster and vector spatial data. | **10 Marks** | **L3** | **CO3** |
| **b.** | Use datetime package to visualize and analyze the stock market data with feature set of ('Open', 'High', 'Low', 'Close', 'Volume').  | **10 Marks** | **L3** |
| **Or** |
| **16.** | **a.** | Define time oriented data. List the characteristics of time oriented data. Use a python code to visualize the time oriented data | **10 Marks** | **L3** | **CO3** |
| **b.** | Apply how text data visualization assists in sentiment analysis. Provide a detailed example, including the types of visualizations that could be used to represent data extracted from social media feeds. | **10 Marks** | **L3** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **17.** | **a.** | Explain data-stream-management system with a neat architectural diagram. Compare DSMS with DBMS | **10 Marks** | **L2** | **CO4** |
| **b.** | Use python code to visualize the streaming data by considering random data point in 2D space | **10 Marks** | **L3** |
| **Or** |
| **18.** | **a.** | Describe a comprehensive discussion on the dos and don’ts of data visualization, using examples from election result reporting.  | **10 Marks** | **L2** | **CO4** |
| **b.** | Illustrate how these principles affect public perception and understanding of the results, and suggest improvements based on recent technological advancements in visualization tools.  | **10 Marks** | **L3** |

**\*\*\*\*\* BEST WISHES \*\*\*\*\***