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

# PRESIDENCY UNIVERSITY

## BENGALURU

**Mid - Term Examinations – October 2025**

**Date:** 28 – 10-2025 **Time:** 11:00am – 12:30pm

| | | |
|---|---|---|
| **School:** SOCSE | **Program**: M.Tech. CSE specialization in Data Science | |
| **Course Code :** DSC4001 | **Course Name:** Programming in Data Science | |
| **Semester**: I | **Max Marks**: :50 | **Weightage**: 25% |

| **CO - Levels** | **CO1** | **CO2** | **CO3** | **CO4** | **CO5** |
|---|---|---|---|---|---|
| **Marks** | **21** | **29** | | | |

**Instructions:**

*(i) Read all questions carefully and answer accordingly.*

*(ii) Do not write anything on the question paper other than roll number.*

## Part A

**Answer ALL the Questions. Each question carries 2marks.** **5Q x 2M=10M**

| | | | | |
|---|---|---|---|---|
| **1** | Define Data Science . | **2 Marks** | **L 2** | **CO1** |
| **2** | Explain the difference between structured and unstructured data. | **2 Marks** | **L 2** | **CO1** |
| **3** | What is the role of inferential statistics in decision-making? | **2 Marks** | **L 2** | **CO1** |
| **4** | How can you handle missing or NaN values in a high-dimensional NumPy array for statistical computations? | **2 Marks** | **L 3** | **CO2** |
| **5** | Name any three placeholder functions in NumPy and their purpose. | **2 Marks** | **L 3** | **CO2** |

## Part B

**Answer the Questions.** **Total Marks 40M**

| | | | | | |
|---|---|---|---|---|---|
| **6.** | **a.** | Analyze how different disciplines such as computer science, statistics, and domain expertise converge within the field of Data Science. Illustrate with an example. | **10 Marks** | **L** 2 | **CO1** |
| | | **Or** | | | |
| **7.** | **a.** | Create a Numpy array and perform the following operations on Sorting, Swapping, Dealing with Missing values | **10 Marks** | **L 2** | **CO2** |
| **8.** | **a.** | Evaluate methods like feature aggregation, encoding, and dimensionality reduction in data preprocessing and recommend an approach for diverse, large-scale data. | **10 Marks** | **L** 2 | **CO1** |
| | | **Or** | | | |
| **9.** | **a.** | Assess differences between descriptive and inferential statistics with an example. | **10 Marks** | **L2** | **CO1** |
| **10.** | **a.** | Assess the use of Numpy for array operations, Renaming, Iteration, Sorting, Statistical functions in large datasets. | **10 Marks** | **L** 3 | **CO2** |
| | | **Or** | | | |
| **11.** | **a.** | Illustrate the use of Pandas window functions, aggregation, and sorting for a time-series dataset. | **10 Marks** | **L 3** | **CO2** |
| **12.** | **a.** | Explain the use of Numpy for basic statistics, copying, slicing, and subsetting of arrays. | **10 Marks** | **L3** | **CO2** |
| | | **Or** | | | |
| **13.** | **a.** | Demonstrate how data can be read, summarized, and inspected using Pandas tools. | **10 Marks** | **L 3** | **CO2** |