Roll No.												

PRESIDENCY UNIVERSITY

BENGALURU

Mid - Term Examinations – March 2026

Date: 13- 03-2026 **Time:**11.45am to 01.15pm

School: SOCSE	**Program:** B.Tech	
Course Code : CSD2002	**Course Name:** Introduction to Data science	
Semester: IV	**Max Marks**: 50	**Weightage:** 25%

CO - Levels	**CO1**	**CO2**	**CO3**	**CO4**	**CO5**
Marks	**12**	**12**	**26**		

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	**List** the six stages of the Data Science Lifecycle.	**2 Marks**	**L1**	**CO1**
2	**Why** is Python preferred over other programming languages for Data Science?	**2 Marks**	**L1**	**CO2**
3	**What** is the function of df.describe() in Exploratory Data Analysis (EDA)?	**2 Marks**	**L1**	**CO3**
4	**Show** how to display only sepal_length and sepal_width from the Iris dataset.	**2 Marks**	**L1**	**CO3**
5	**Name** any two techniques used to handle outliers in a dataset.	**2 Marks**	**L1**	**CO3**

Part B

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

6.	**a.**	**Explain** various real-world applications of Data Science.	**10 Marks**	**L2**	**CO1**
	b.	**Summarize** the role of data science in Healthcare and Education Analytics with reference to hospital and educational platform case scenarios.	**10 Marks**	**L2**	**CO1**
		Or			
7.	**a.**	**Describe** the fundamental operations of NumPy and Pandas, explaining why they are essential for handling large datasets.	**10 Marks**	**L2**	**CO2**
	b.	**Compare** and contrast Jupyter Notebook, Anaconda, and Google Colab as environments for data analysis.	**10 Marks**	**L2**	**CO2**

8.	**a.**	**Apply** appropriate data analysis techniques to load the Titanic dataset and perform basic Exploratory Data Analysis (EDA).	**10 Marks**	**L3**	**CO3**
	b.	**Develop** and utilize Feature Engineering techniques to improve model performance by creating new features from raw data.	**10 Marks**	**L3**	**CO3**
		Or			
9.	**a.**	**Identify** and differentiate data cleaning, data wrangling, and data mining by organizing their scope with suitable examples.	**10 Marks**	**L3**	**CO3**
	b.	**Utilize** Pandas operations to analyze the Iris dataset and extract meaningful insights.	**10 Marks**	**L3**	**CO3**