PRESIDENCY UNIVERSITY

BENGALURU

Mid - Term Examinations – October 2025

Date: 10-10-2025　　　　　　　　　　　　　　　　　**Time:** 11.45am to 01.15pm

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

CO - Levels	CO1	CO2	CO3	CO4
Marks	25	25		

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 any two roles in the data science job ecosystem.	2 Marks	L1	CO1
2	List two examples of data sources used in data science.	2 Marks	L1	CO1
3	List the types of qualitative data.	2 Marks	L1	CO1
4	List two differences between data and information.	2 Marks	L1	CO1
5	List two applications of data science in healthcare.	2 Marks	L1	CO1

Part B

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

6.	a.	**Discuss how data wrangling is applied in real-world scenarios such as healthcare, social media, or e-commerce, and interpret its impact.(5 marks)**	10 Marks	L2	CO2
	b.	Outline the steps involved in data exploration and describe how summary statistics and visualizations help in understanding datasets. (5 Marks)		L2	CO2
Or					

7.	a.	Explain why data wrangling is important in data science and summarize how it leads to quality insights. (5 Marks)	10 Marks	L2	CO2
	b.	Differentiate between data cleaning and data wrangling by describing their scope with suitable examples. (5 Marks)		L2	CO2

8.	a.	**Discuss the use of functions like sqrt(), sum(), and transpose (T) in matrix manipulation with an example. (5 Marks)**	15 Marks	L1	CO1
	b.	Describe the process of creating a NumPy array using lists, tuples, and explicit data types with suitable examples (5 Marks)		L1	CO1
Or					
9.	a.	Interpret the output of basic NumPy array operations such as addition, subtraction, and reshaping with examples. (5 Marks)	15 Marks	L1	CO1
	b.	Summarize different matrix operations (add, subtract, multiply, divide, dot) and highlight their significance in data science. (5 Marks)		L1	CO1

10.	a.	**Explain how the Standard Deviation Method and the IQR Method are used for outlier detection with suitable examples. (7 Marks)**	15 Marks	L2	CO2
	b.	Compare normalization and standardization as data transformation techniques and explain when each should be used. (8 Marks)		L2	CO2
Or					
11.	a.	Discuss univariate, bivariate, and multivariate analysis in exploratory data analysis (EDA) and illustrate the techniques used in each. (7 Marks)	15 Marks	L2	CO2
	b.	Classify the processes involved in feature engineering and describe how each contributes to building better features. (8 Marks)		L2	CO2