



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

PRESIDENCY UNIVERSITY

BENGALURU

Mid - Term Examinations –October2025

Date: 09-10-2025

Time: 09.30am to 11.00am

School: SOCSE	Program: CBD	
Course Code :CBD-3406	Course Name: Introduction to Data Science & Big Data	
Semester: V	Max Marks:50	Weightage:25%

CO - Levels	CO1	CO2	CO3	CO4	CO5
Marks	24	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 x2M=10M

1	Summarize the benefits of using Big Data in modern businesses.	2 Marks	L2	CO1
2	List two common sources of unstructured data.	2 Marks	L2	CO1
3	Define outliers.	2 Marks	L2	CO2
4	Illustrate the two consequences of missing data	2 Marks	L2	CO2
5	List two types of correlation commonly observed between variables.	2 Marks	L2	CO2

Part B

Answer the Questions.

Total Marks 40M

6.	a.	Apply Python and R program code to real world use case	10Marks	L3	CO1
Or					
7.	a.	Apply word count example to Hadoop Map Reduce	10 Marks	L3	CO1

8.	a.	Illustrate the various data science Roles.	10Marks	L2	CO1
Or					
9.	a.	Illustrate the 5 Vs of big data with real world example	10 Marks	L2	CO1

10.	a.	Apply one-hot encoding during feature transformation with example	10 Marks	L3	CO2
Or					
11.	a.	Apply descriptive statistics (mean, median, standard deviation) to summarize to a sample dataset.	10 Marks	L3	CO2

12.	a.	Explain histogram representation with dataset	10 Marks	L2	CO2
Or					
13.	a.	Explain in details with sample data to various visualization charts	10 Marks	L2	CO2