



DCBD403

Reg. No.

--	--	--	--	--	--	--	--

IV Semester B.Com. (BDA) Degree Examination, June/July - 2025

COMMERCE

Data Science Life Cycle

(NEP Scheme)

Time : 2½ Hours

Maximum Marks : 60

*Instructions to Candidates:*

Answers should be written only in English.

**SECTION - A**

Answer any **Five** of the following questions. Each question carries **2** marks. (5×2=10)

1. a) What is Apache Hadoop?
- b) Mention the stages in Data Science.
- c) What is bias in Machine Learning?
- d) What is Agreement in NLP?
- e) Mention few innovative tools used for data visualization.
- f) What is clustering?
- g) Mention two problems in Overfitting in Machine Learning.

**SECTION - B**

Answer any **Four** of the following questions. Each question carries **5** marks. (4×5=20)

2. Explain Data Analysis Pipeline.
3. Explain the challenges in NLP.
4. Briefly explain IBM Big Data Strategy.
5. Explain the tools used in Data Science.
6. How does Machine Learning Works?

[P.T.O.]



(2)

DCBD403

SECTION - C

Answer any **Two** of the following questions. Each question carries **12** marks. **(2×12=24)**

7. Explain Data Extraction and Exploration.
8. Explain different types of charts in Data Visualization.
9. Explain Analysing Data in Hadoop.

SECTION - D

Answer any **One** of the following questions. Each question carries **6** marks. **(1×6=6)**

10. Explain various types of Machine Learning.
11. What are the applications of NLP? Explain.