



Term Evaluation Theory (Even) Semester Regular Examination February 2020

Roll no. **252409**Name of the Course: **BCA / BSc IT**Semester: **2nd**Name of the Paper: **Introduction to Data Science**Paper Code: **TBD 201 / CAD 201**Time: **1.5 hour**Maximum Marks: **50****Note:**

- (i) Answer all the questions by choosing any one of the sub-questions
- (ii) Each question carries 10 marks.

Q1.

(10 Marks)

- a. Explain the concept of Exploratory Data Analysis (EDA). Discuss how statistical measures and graphical methods are used in the EDA process. **CO 2**

OR

- b. How has Data Science evolved over time? Discuss the key stages of its growth and explain its role in the present AI-based era. **CO 1**

Q2.

(10 Marks)

- a. Describe the Hadoop framework and its architecture with a diagram. Explain how Hadoop processes large volumes of data using a distributed system. **CO 1**

OR

- b. Discuss data retrieval as a step in the Data Science process. Explain the different data sources and data collection techniques used for analysis. **CO 2**

Q3.

(10 Marks)

- a. Explain the facets of Data Science. How do statistics, computer science, domain knowledge, and visualization contribute to the Data Science process? **CO 1**

OR

- b. Explain data cleaning in detail. Discuss common data quality issues and techniques used to handle missing values, noise, outliers, and inconsistencies. **CO 2**

Q4.

(10 Marks)

- a. Discuss the applications of Data Science in various domains. Explain how Data Science is used in healthcare, finance, and e-commerce. **CO 1**

OR

- b. Discuss the methods involved in preparing raw data for analysis. Explain how data cleaning, integration, and transformation improve data quality. **CO 2**

Q5.

(10 Marks)

Term Evaluation Theory (Even) Semester Regular Examination February 2026

- a. Describe the complete lifecycle of Data Science. Explain each phase starting from problem identification to model implementation with a suitable flow diagram. **CO 1**

OR

- b. Explain the Data Science Process in detail. Describe each stage involved in transforming raw data into meaningful insights. **CO 2**