

INDIRA UNIVERSITY, PUNE

SET- 2

SCHOOL OF INFORMATION TECHNOLOGY- BSC DS

Backlog Examination (2025 Pattern) April/ May – 2026 - Semester – I

Subject Name: Foundations of Data Science
Subject Code: 25BDS112T

Max.Marks: 25
Time: 1 ½ Hours

Instructions

- All questions are compulsory.

CO #	Cognitive Ability	Course Outcome
CO1	Remember	Recall and explain fundamental concepts of data science, including the data science lifecycle, statistical measures, and data types.
CO2	Understand	Interpret and describe various statistical and mathematical foundations, data collection methods, and visualization techniques.
CO3	Apply	Implement data preprocessing, statistical analysis, and data visualization methods to solve real-world data science problems
CO4	Analyze	Analyze different data types, data distributions, and outliers using statistical tools and explore relationships through visualizations
CO5	Evaluate	Evaluate the effectiveness of data preprocessing, statistical models, and visualization techniques for meaningful data interpretation.
CO6	Create	Design and develop data science solutions by integrating data collection, preprocessing, statistical analysis, and visualization techniques.

Q1.	Attempt any 5 out of 7. (1 Mark Each) a) Define Data Science. b) What is Variety in data? c) What is variance ? d) Define structured data . e) What is graph-based data ? f) Define nominal attribute . g) What do you mean by temporal data ?	(5 Marks)	CO1
Q2.	Attempt any 2 out of 4 (5 Marks Each) a) Compare mean, median, and mode with examples. b) Illustrate the characteristics of graph-based and streaming data . c) Illustrate the need for data quality and preprocessing in the data science workflow. d) Compare bar charts, dot plots, and histograms in terms of use and data type.	(10 Marks)	CO2

Q. 3.	Attempt any 1 out of 2 (10 Marks) 1. a) Apply label encoding and one-hot encoding on a categorical dataset of your choice and explain the result. (5 Marks) b) Explain the structure and interpretation of box-and-whisker plots. (5 Marks) 2. Prepare a data preprocessing pipeline for handling numeric and categorical data — from collection to transformation. (10 Marks)	CO3
-------	--	-----
