

INDIRA UNIVERSITY, PUNE

SCHOOL OF INFORMATION TECHNOLOGY- MSC CS

Term End Examination (2025 Pattern) April- May – 2026- Semester – II

Course Name: Big Data
Course Code: 25PSC205AT

Max. Marks: 25
Time: 1 Hr. 30 Minutes

Instructions

- All questions are compulsory
- Figures to the right indicate full marks.

Q1.	Attempt any 5 out of 7. (1 Mark each) a) Define Big Data. b) What is DataNode in Hadoop? c) What is Hive Metastore? d) Define RDD in Apache Spark. e) What is Sqoop used for? f) What is the role of Zookeeper in Hadoop ecosystem? g) List any two characteristics of Big Data	(5 Marks)	CO 1
Q2.	Attempt any 2 out of 4. (5 marks each) a) Explain Hadoop Architecture with neat diagram. b) Describe the 3Vs of Big Data with suitable examples. c) Differentiate between Hive and HBase. d) Explain the working of MapReduce with an example.	(10 Marks)	CO 2
Q.3	Attempt all questions (5 marks each) a) Write a note on Spark RDD Operations. b) Analyze and explain HDFS features in detail. OR	(10 Marks)	CO 3
Q.3	Attempt all questions (5 marks each) a) Sketch and explain Spark Architecture. b) Write a note on Hive Partitioning and Bucketing.	(10 Marks)	CO 3
