

GUJARAT TECHNOLOGICAL UNIVERSITY

BE- SEMESTER-VII EXAMINATION – WINTER 2025

Subject Code:3174207

Date:24-11-2025

Subject Name:Big Data Analytics

Time:10:30 AM TO 01:00 PM

Total Marks:70

Instructions:

1. Attempt all questions.
2. Make suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.
4. Simple and non-programmable scientific calculators are allowed.

MARKS

- Q.1**
- | | |
|---|-----------|
| (a) What is data? Explain different types of data. | 03 |
| (b) Compare new data to traditional data. | 04 |
| (c) Explain Descriptive analytics, Diagnostic analytics, Predictive analytics and Perceptive analytics. | 07 |

- Q.2**
- | | |
|---|-----------|
| (a) What is Big data? List types of Big data. | 03 |
| (b) Explain the Vs of Big data in short. | 04 |
| (c) Explain common applications and case study of Big data solutions. | 07 |

OR

- | | |
|--|-----------|
| (c) Explain different sources of Big data and challenges of conventional system. | 07 |
|--|-----------|

- Q.3**
- | | |
|--|-----------|
| (a) What is HDFS? List down any 3 Hadoop's configuration files. | 03 |
| (b) Explain different components of Hadoop in short. | 04 |
| (c) Write and Explain a 'WordCount' Java MapReduce program for input Text: | 07 |

Bus, Car, bus, car, train, car, bus, car, train, bus, TRAIN,BUS,
buS, caR, CAR, car, BUS, TRAIN

OR

- Q.3**
- | | |
|--|-----------|
| (a) Which commands are used to administer Hadoop cluster. | 03 |
| (b) Explain design of HDFS in short. | 04 |
| (c) Explain different step to set up a Hadoop multinode cluster. | 07 |

- Q.4**
- | | |
|---|-----------|
| (a) What is SQL? | 03 |
| (b) What is NoSQL? How it differs from SQL? | 04 |
| (c) Explain the CAP theorem. | 07 |

OR

- Q.4**
- | | |
|---|-----------|
| (a) What is Cassandra? | 03 |
| (b) Differentiate master-slave Vs peer-to-peer distribution model. | 04 |
| (c) Explain Key-value stores, Graph stores, Column family (Bigtable) stores, and Document stores. | 07 |

- Q.5**
- | | |
|---|-----------|
| (a) What is Zookeeper? Explain it in short. | 03 |
| (b) Explain Stream mining in short. | 04 |
| (c) Explain 'WordCount' Scala program for Apache Spark. | 07 |

OR

- | | | | |
|------------|------------|--|-----------|
| Q.5 | (a) | Explain different components of HBase architecture in short. | 03 |
| | (b) | What is metastore in Hive. Explain it in short. | 04 |
| | (c) | Justify the statement: Spark is faster than MapReduce. | 07 |
