

# GUJARAT TECHNOLOGICAL UNIVERSITY

## MCA INTEGRATED – SEMESTER VIII- EXAMINATION –WINTER-2024

**Subject Code:2688601**

**Date: 20/11/2024**

**Subject Name: Big Data Analytics**

**Time:02:30 PM TO 05:00 PM**

**Total Marks: 70**

**Instructions:**

1. Attempt all questions.
2. Make Suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.
4. Use of simple calculators and non-programmable scientific calculators are permitted.

- Q.1** (a) Discuss replication process in MongoDB. **03**  
(b) Explain characteristics of big data. **04**  
(c) Write short note on Hadoop Ecosystem. **07**
- Q.2** (a) Discuss the working of map task and reduce task. Solve following word count **07**  
problem using map-reduce.  
Big Data Mongo  
PIG Mongo Hadoop  
Big PIG Hadoop  
Mongo Data Big
- (b) Define NoSQL. Discuss the advantages of NoSQL. **07**  
**OR**
- (b) Explain following HDFS commands with example: **07**  
I. mkdir II. copyFromLocal III. cat
- Q.3** (a) Discuss Hive architecture in detail. **07**  
(b) What is HBase? Explain HBase tables and regions. **07**  
**OR**
- Q.3** (a) Compare Hive with traditional databases. **07**  
(b) Enlist and explain the types of NoSQL databases. **07**
- Q.4** (a) Explain execution modes of PIG. **07**  
(b) List the features of SPARK. Discuss RDD in detail. **07**  
**OR**
- Q.4** (a) Define zookeeper. Discuss zookeeper characteristics in detail. **07**  
(b) Discuss data serialization in hadoop. **07**
- Q.5** (a) Discuss creation of MongoDB databases, collections and documents with example. **07**  
(b) Explain following MongoDB aggregate operators with example: **07**  
1. \$project 2. \$group  
**OR**
- Q.5** (a) What is database schema design? Discuss the principles of schema design. **07**  
(b) Explain following: **07**  
1. Scaling in MongoDB 2. Document Data Model

\*\*\*\*\*