

Seat No.: _____

Enrolment No. _____

GUJARAT TECHNOLOGICAL UNIVERSITY
MCA - SEMESTER-V EXAMINATION – WINTER 2018

Subject Code: 3650010

Date: 28/11/2018

Subject Name: Big Data Tools

Time: 10.30 am to 1.00 pm

Total Marks: 70

Instructions:

1. Attempt all questions.
2. Make suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.

- Q.1 (a) Explain Following. 07**
1. PigLoader
 2. REGEX_EXTRACT
 3. Tokenizer
 4. Pig Vs. Hive
 5. FLATTEN
 6. JobTracker
 7. ResourceManager
- (b) Explain 4V's of Big Data with Example. 07**
- Q.2 (a) What is Hadoop? Describe Hadoop Eco System with details. 07**
- (b) What is NoSQL? Different Between SQL Vs. NoSQL. Also explain advantages of NoSQL. 07**

OR

- (b) Describe NameNode, Secondary NameNode, and DataNode with example. 07**
- Q.3 (a) Explain any 7 MongoDB Function with example. 07**
- (b) How Replication and fault tolerance important in Hadoop? Discuss it. 07**

OR

- Q.3 (a) What is MapReduce? Explain Mapper and Reducer. 07**
- (b) Discuss Any 5 HDFS Command with example. 07**
- Q.4 (a) What is Hive? Explain Type of Partitions with example. 07**
- (b) How piggybank help in pig? Discuss with example of piggybank. 07**

OR

- Q.4 (a) Explain Hive Complex data Type with example. 07**
- (b) What is Pig? Describe its Complex Data types with example. 07**
- Q.5 (a) Write a short note on MLib and Tensor Flow. 07**
- (b) What is Spark? Explain RDD in details. 07**

OR

Q.5 (a) Write a Hive query for word count. **07**

(b) Write a Pig Query for following Operation on HDFS. **07**

Data Set:

(person, dstore, spent)

A, S, 3.3

A, S, 4.7

B, S, 1.2

B, T, 3.4

C, Z, 1.1

C, T, 5.5

D, R, 1.1

a) List total sales per department stores:

b) List total sales per customer

c) List Total No. of customer visited per department
