

LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034



M.Sc. DEGREE EXAMINATION – COMPUTER SCIENCE

THIRD SEMESTER – NOVEMBER 2019

16/17/18PCS3MC01 – BIG DATA ANALYTICS

Date: 29-10-2019

Dept. No.

Max. : 100 Marks

Time: 09:00-12:00

PART-A

Answer all the Questions:-

10X2=20

1. What is structured and unstructured data?
2. Define Radom variable with example.
3. What is stream processing in big data?
4. What is mean by GIPOMS?
5. Write down the highlights of map reduce.
6. What are the services provided by Hcatalog?
7. What is the role of the Map reduce cluster?
8. List the Hadoop log commands messages to Log4J.
9. Draw the HBase cluster members.
10. Write down the operations in ZooKeeper.

PART-B

Answer all the Questions:-

5X8=40

11 a) Explain the need for Big Data Analytics.

Or

b) Explainin detail about Hypothesis testing.

12 a) Explain about counting distinct elements in a stream.

Or

b) Explain about Filtering Streams in detail.

13 a) Explain the different scheduling methods in Hadoop YARN.

Or

b) ExplainShuffle and Sort phase in Hadoop Map Reduce.

14 a) Explain Configuring Hadoop in fully-distributed mode.

Or

b) Explain Routine Administration Procedures in Hadoop.

15 a) Write short note on ZooKeeper.

Or

b) Discuss the various tools available in Hadoop ecosystem.

PART-C

Answer Any TWO Questions:-

2X20=40

16 a) Explain in detail about statistical inference and probability axioms.

b) Discuss about Stream Queries and ad-hoc queries.

17 a) Explain Hadoop Distributed File System architecture.

b) Explain about Hadoop benchmarks.

18 a) Explain LOAD and TRANSFORM commands in Pig.

b) Explain the steps involved in using BigSheets to perform big data analysis.
