

Reg. No.

--	--	--	--	--	--	--	--	--	--



CSS 505

III Semester M.Sc. Degree Examination, December 2018/January 2019
COMPUTER SCIENCE
Big Data Analytics

Time : 3 Hours

Max. Marks : 70

Instructions : 1) Answer **any five** of the following questions.

2) **All** questions carries **equal** marks.

1. a) What are the limitations of distributed file system ? Discuss how they are resolved with HDFS. 7
- b) Define Big Data. Describe the different features of Big Data of real world. 7
2. a) Describe the map and reduce programming components with word count example. 7
- b) Discuss various steps involved in performing matrix addition operations in map reduce programming environment. 7
3. a) What are the different demon virtual possess of apache Hadoop system ? Briefly describe each. 7
- b) What is data redundancy in Hadoop cluster ? Discuss the advantages of it with respect to data high availability. 7
4. a) What is Name Node ? Describe the functions of Name Node. 7
- b) Distinguish between Resource Manager and Application Manager in a Hadoop cluster. 7

P.T.O.



- 5. a) What is Hadoop Ecosystem ? Describe in detail the Hadoop core components of Hadoop Ecosystem. 7
 - b) Briefly describe the different processing layers of Hadoop ecosystems. 7
 - 6. a) Describe the significance differences between Hadoop 1.x and Hadoop 2.x based on their core components. 7
 - b) List the limitations of Hadoop 1.x with respect to Hadoop processing environment. 7
 - 7. a) What is No SQL ? Discuss the advantages of over SQL. 7
 - b) Write any five HBASE statements and describe them briefly. 7
 - 8. Write short notes on the following : **(2×7=14)**
 - a) PIG language.
 - b) Zookeeper.
-