H.T No: | | | | | | | | | | | **R15** **Course Code: A2544**

CMR COLLEGE OF ENGINEERING & TECHNOLOGY

(UGC AUTONOMOUS)

B.Tech VII Semester Supplementary Examinations June-2022

Course Name: Big Data Analytics

(CSE)

Date: 14.06.2022 AN **Time: 3 hours** **Max.Marks: 70**

(Note: Assume suitable data if necessary)

PART-A

Answer all TEN questions (Compulsory)

Each question carries TWO marks. **10x2=20M**

1.	What is Big Data? Justify the need to analyze Bigdata	2 M
2.	Compare and contrast Bigdata with RDBMS	2 M
3.	Is name node same as data node in HDFS.	2 M
4.	Suppose there is file of size 514 MB stored in HDFS (Hadoop 2.x) using default block size configuration and default replication factor. Then, how many blocks will be created in total and what will be the size of each block?	2 M
5.	List two major functionalities of Hadoop YARN.	2 M
6.	Define map reduce?	2 M
7.	Compare Apache Pig and SQL.	2 M
8.	What is Pig Grunt?	2 M
9.	Describe Hive.	2 M
10.	Justify the need of R in data analytics.	2 M

PART-B

Answer the following.Each question carries TEN Marks. **5x10=50M**

11.A).	Discuss the various dimensions of Big Data. Give two examples for the case studies. Indicate which dimension is satisfied by these case studies.	10M
	OR	
11. B).	List the different data architecture patterns in NOSQL. Explain (i) Key value store (ii) Document store	10M
12. A).	What is HDFS? Illustrate the major blocks in HDFS architecture.	10M
	OR	
12. B).	Illustrate the rack awareness in HDFS and how does it minimize bandwidth consumption and read latency	10M
13. A).	Comprehend the functionalities of Hadoop ecosystem components with a neat diagram.	10M
	OR	
13. B).	Explain concept of Map Reduce using an example.	10M
14. A).	Describe the relational operators in PIG Latin.	10M
	OR	
14. B).	Inspect the different mechanisms to define parameters in Pig latin script.	10M
15. A).	Classify primitive and complex data types in Hive.	10M
	OR	
15. B).	Illustrate the Hive architecture and infer the functionalities of the components.	10M

*****