Code No.: DS702PC   R20   H.T.No. [ ][ ][8][R][ ][ ][ ][ ]

CMR ENGINEERING COLLEGE: : HYDERABAD
UGC AUTONOMOUS
IV–B.TECH–I–Semester End Examinations (Regular) - November- 2023
BIG DATA ANALYTICS
(CSD)

[Time: 3 Hours]   [Max. Marks: 70]

Note: This question paper contains two parts A and B.
Part A is compulsory which carries 20 marks. Answer all questions in Part A.
Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks and may have a, b, c as sub questions.

PART-A   (20 Marks)

1. a)	List the main characteristics of Big Data.	[2M]
b)	What are the advantages of Hadoop?	[2M]
c)	Define HDFS.	[2M]
d)	What is the advantage of MaPR?	[2M]
e)	Define Map Phase.	[2M]
f)	What are the types of map reduce?	[2M]
g)	Define HIVE.	[2M]
h)	What is Pig Latin?	[2M]
i)	Differentiate between Sql and NoSql.	[2M]
j)	Define HBase.	[2M]

PART-B   (50 Marks)

2. Discuss about the Hadoop ecosystem components in four layers.   [10M]

OR

3. Explain the process of setting up a hodoop cluster.   [10M]

4. Explain about HDFS architecture.   [10M]

OR

5. Explain the working of Map reduce with a relevant example.   [10M]

6. Describe the MapReduce execution steps with neat diagram.   [10M]

OR

7. Write the deference between input formats and output formats for MapReduce framework.   [10M]

8. Explain working of Hive with proper steps and diagram.   [10M]

OR

9. Explain the pig Latin data types and examples.   [10M]

10. Explain about the various data types supported by HiveQL with an example.   [10M]

OR

11. Explain the NoSQL databases and their characteristic and features.   [10M]

****