

Total No. of Questions – [06]

Total No. of Printed Pages: 01

G.R. No.	
----------	--

PAPER CODE	P.122-234-DEER
------------	----------------

May 2022 / INSEM+ENDSEM
F. Y. M. TECH. (COMPUTER ENGG) (SEMESTER – II)
COURSE NAME: BIG DATA ANALYTICS
COURSE CODE: CSPA12204B
(PATTERN 2020)

Time: [3 Hours]

[Max. Marks: 60]

Instructions to candidates:

- 1) All Questions are compulsory
- 2) Figures to the right indicate full marks.
- 3) Use of scientific calculator is allowed
- 4) Use suitable data where ever required

- Q.1) a. Explain characteristics of big data. [5]
b. Explain concept of big data in healthcare system. [5]
- Q.2) Define HDFS. Describe namenode, datanode and block. [10]
Explain HDFS operations in detail.
- Q.3) “Spark is superior than Hadoop”. Justify your answer. [10]
- Q.4 a) What is NoSQL database? Explain different types of [5]
NoSQL database
b) Create the database MTech. Create collection student_details
in created database. Insert your information in the database. [5]
- Q.5) Explain HBase architecture with suitable diagram. [10]
- Q.6) Explain two execution types or modes in PIG [10]