

**BACHELOR OF TECHNOLOGY (CBCS) (2021-COURSE)**  
**B. Tech. Sem - V COMPUTER SCIENCE & ENGINEERING : SUMMER : 2024**  
**SUBJECT: BIG DATA ANALYTICS**

Day : Thursday  
Date : 09/05/2024

**S-25589-2024**

Time : 10:00 AM-01:00 PM  
Max. Marks : 60

**N.B.**

- 1) All questions are **COMPULSORY**.
- 2) Figures to the right indicate **FULL** marks.
- 3) Assume suitable data if necessary.

Q. 1      What is Big Data? Differentiate Big Data and Traditional data.      (10)

**OR**

Q. 1      What are the challenges of conventional system? Write a note on modern data analytic tools.      (10)

Q. 2      What is Hadoop? How do you analyze the data in Hadoop?      (10)

**OR**

Q. 2      Explain the following:      (10)  
            a) Mapper class  
            b) Reduce class  
            c) Scaling out

Q. 3      What are the different types of regression models? Explain any one in detail.      (10)

**OR**

Q. 3      What is classification? Differentiate between correlation and regression.      (10)

Q. 4      Write a short note on      (10)  
            a) counting distinct elements in stream  
            b) finding most popular elements using decaying window

**OR**

Q. 4      What is Real time analytics? Discuss their technologies in detail.      (10)

Q. 5      Discuss overview of clustering. Explain about K-means clustering in detail.      (10)

**OR**

Q. 5      Write a detailed note on classification of Decision tree.      (10)

Q. 6      What is PIG? Explain its installing process and two execution types or modes in PIG.      (10)

**OR**

Q. 6      Describe NoSQL. Explain Hbase and Cassandra Clients in detail.      (10)

\* \* \* \* \*