

BACHELOR OF TECHNOLOGY (CBCS) (2020 COURSE)
B.Tech.Sem - VIII INFORMATION TECHNOLOGY : SUMMER : 2025
SUBJECT: DATA ENGINEERING

Day : Tuesday
Date : 27/05/2025

S-24773-2025

Time : 02:30 PM-05:30 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 data architecture? Enlist and elaborate the principles of good data architecture. [10]

OR

Q.1 Are data science and data engineering same or different? Justify your answer elaborately. [10]

Q.2 Categorize the major source systems in data engineering. Elaborate any two in detail. [10]

OR

Q.2 Determine how the APIs, REST, data sharing, message queues and event streaming platforms are source systems for data creation and collection. [10]

Q.3 Validate the importance of Data Storage Lifecycle and Data Retention in the data engineering life cycle. [10]

OR

Q.3 Compare and contrast data warehouse vs data lake vs data lakehouse. [10]

Q.4 Explain the concept of batch data ingestion and discuss three common methods used for batch data ingestion. [10]

OR

Q.4 Outline three best practices for effective data ingestion implementation, highlighting their importance in ensuring data quality and reliability. [10]

Q.5 Review in detail the process and significance of reverse ETL. [10]

OR

Q.5 Identify and explain three sophisticated data transformation techniques or feature engineering methods used to enhance model performance. [10]

Q.6 Examine the purpose and significance of Live Data Stack in Data Engineering. [10]

OR

Q.6 Discuss three key features or capabilities of cloud-scale data OS that contribute to enhancing interoperability across diverse data platforms and services. [10]

* * * * *