

(Time: 2½ Hours)

[Total Marks: 75]

- N.B.:** (1) All questions are compulsory.
 (2) Make suitable assumptions wherever necessary and state the assumptions made.
 (3) Answers to the same question must be written together.
 (4) Numbers to the right indicate marks.
 (5) Draw neat labeled diagrams wherever necessary.
 (6) Use of Non-programmable calculators is allowed.

1. Attempt any Three of the following: 15

- State and Explain the stages of Data Engineering Life Cycle.
- What is the role and significance of Data Engineering in today's technological landscape?
- What are architectural tiers? Explain single tier and multitier.
- What skills and activities are essential for professionals in the field of Data Engineering to effectively navigate the complexities of their work?
- What role do Data Engineers play within an organizational framework?
- Discuss major undercurrent across the data engineering life-cycle.

2. Attempt any Three of the following: 15

- Discuss principles of good architecture.
- Discuss the following terms for choosing technologies across the data engineering life-cycle:
 - Cost optimization and business value
 - Build vs Buy
- What roles and expertise are typically involved in the collaborative process of designing a Data Architecture within an organization?
- What are the fundamental concepts, types, and examples that define the landscape of Data Architecture within information systems?
- State the difference between mutable and transistor technology.
- How do underlying trends and influences throughout the technological landscape impact the decision-making process when choosing technologies, particularly within the context of data engineering?

3. Attempt any Three of the following: 15

- What is data? How is data created? Also describe the types of data.
- State and explain CRUD operations.
- What are the major concepts and emerging trends that define the landscape of storage solutions, reflecting the significant ideas shaping the field of data storage?
- How do storage abstractions contribute to the field of data engineering?
- Define source system practical details:
 - Databases
 - RDBMS
 - Key value stores
 - document stores.
- What are the key considerations and mechanisms associated with the systematic transfer and processing of data in batches?

4. Attempt any Three of the following:

15

- a How can data ingestion be defined, and what processes and mechanisms are involved in the ingestion of data into information systems?
- b What is batch ingestion? How does batch ingestion play a vital role in the data engineering process?
- c What are data pipelines? Explain ETL process in the data pipeline.
- d How do queries function in the context of data processing? What role do they play in extracting, manipulating, and retrieving information from databases or data sets?
- e Explain the term data scraping.
- f Explain the difference between push vs poll vs pull data.

5. Attempt any Three of the following:

15

- a. What a Data Engineer Should Know About ML?
 - b. Define analytics. Explain the types of analytics.
 - c. State and explain the working of reverse ETL.
 - d. State and explain the important factor required for security in data analytics.
 - e. What are the general considerations for serving data?
 - f. What is Serialization? Explain Columnar and Hybrid Serialization.
-