



**QUESTION BANK**

Year / Semester: **II B.Tech IV Semester**

Regulation: **R23**

Subject and Code: **DATA ENGINEERING, 23CSD242**

**SYLLABUS**

**UNIT-I: INTRODUCTION TO DATA ENGINEERING (9)**

**Introduction to Data Engineering: Definition, Data Engineering Life Cycle, Evolution of Data Engineer, Data Engineering Versus Data Science, Data Engineering Skills and Activities, Data Maturity, Data Maturity Model, Skills of a Data Engineer, Business Responsibilities, Technical Responsibilities, Data Engineers and Other Technical Roles.**

**UNIT-II: DATA ENGINEERING LIFE CYCLE (9)**

**Data Engineering Life Cycle: Data Life Cycle Versus Data Engineering Life Cycle, Generation: Source System, Storage, Ingestion, Transformation, Serving Data. Major undercurrents across the Data Engineering Life Cycle: Security, Data Management, Data Ops, Data Architecture, Orchestration, Software Engineering.**

**UNIT-III: DESIGNING GOOD DATA ARCHITECTURE (9)**

**Designing Good Data Architecture: Enterprise Architecture, Data Architecture, Principles of Good Data Architecture, Major Architecture Concepts. Data Generation in Source Systems: Sources of Data, Files and Unstructured Data, APIs, Application Databases (OLTP), OLAP, Change Data Capture, Logs, Database Logs, CRUD, Source System Practical Details.**

**UNIT-IV: STORAGE AND INGESTION (9)**

**Storage: Raw Ingredients of Data Storage, Data Storage Systems, Data Engineering Storage Abstractions, Data warehouse, Data Lake, Data Lakehouse. Ingestion: Data Ingestion, Key Engineering considerations for the Ingestion Phase, Batch Ingestion Considerations, Message and Stream Ingestion Considerations, Ways to Ingest Data.**

**UNIT-V: QUERIES, MODELING AND TRANSFORMATION (9)**

**Queries, Modeling and Transformation: Queries, Life of a Query, Query Optimizer, Queries on Streaming Data, Data Modelling, Modeling Streaming Data, Transformations, Streaming Transformations and Processing. Serving Data for Analytics, Machine Learning and Reverse ETL: General Considerations for serving Data, Business Analytics, Operational Analytics, Embedded Analytics, Ways to serve data for analytics and ML, Reverse ETL.**



**SREENIVASA INSTITUTE OF TECHNOLOGY AND MANAGEMENT STUDIES  
(AUTONOMOUS)**

**QUESTION BANK**

**Year / Semester: II B.Tech IV Semester**

**Regulation: R23**

**Subject and Code: DATA ENGINEERING, 23CSD242**

**Max Marks: 10**

S.No.	CO	Questions	BT
<b>Unit I: (INTRODUCTION TO DATA ENGINEERING)</b>			
1	1	Compare Data Engineering and Data Science with examples.	L4
2	1	How does Data Engineering help Machine Learning projects?	L3
3	1	Explain problems faced while building large data systems.	L4
4	1	Explain how to improve data quality in an organization.	L3
5	1	Design a simple data team for a small company.	L5
6	1	Explain the stages of Data Engineering Life Cycle clearly.	L4
7	1	Explain the main duties of a Data Engineer in a company.	L3
8	1	Create a simple plan to build a data platform.	L5
9	1	Explain how the role of Data Engineer has changed over time.	L4
10	1	Explain system thinking with an example.	L3
11	1	Explain Data Maturity Model and its levels.	L3



**SREENIVASA INSTITUTE OF TECHNOLOGY AND MANAGEMENT STUDIES  
(AUTONOMOUS)**

**QUESTION BANK**

**Year / Semester: II B.Tech IV Semester**

**Regulation: R23**

**Subject and Code: DATA ENGINEERING, 23CSD242**

S.No.	CO	Questions	BT
<b>Unit II: (DATA ENGINEERING LIFE CYCLE)</b>			
1	2	Compare ETL and ELT.	L4
2	2	Design a simple batch data pipeline.	L3
3	2	Explain common problems in data pipelines.	L4
4	2	Design a simple real-time data pipeline.	L3
5	2	Design a simple secure data pipeline.	L5
6	2	Evaluate batch vs stream processing.	L4
7	2	Explain the role of orchestration tools.	L3
8	2	Explain the importance of security in data systems.	L5
9	2	Explain how data moves from source to user.	L4
10	2	Explain batch processing with example.	L3
11	2	Explain stream processing with example.	L4



**SREENIVASA INSTITUTE OF TECHNOLOGY AND MANAGEMENT STUDIES  
(AUTONOMOUS)**

**QUESTION BANK**

**Year / Semester: II B.Tech IV Semester**

**Regulation: R23**

**Subject and Code: DATA ENGINEERING, 23CSD242**

S.No.	CO	Questions	BT
<b>Unit III: (DESIGNING GOOD DATA ARCHITECTURE)</b>			
1	3	Explain the difference between enterprise and data architecture.	L4
2	3	Design a simple data architecture for an online shopping system.	L3
3	3	Compare centralized and distributed systems.	L4
4	3	Explain OLTP and OLAP with examples.	L3
5	3	Design a cloud-based data architecture.	L5
6	3	Explain how APIs are used in data systems.	L4
7	3	Explain how CDC works.	L3
8	3	Create a simple event-driven architecture.	L5
9	3	Explain challenges in source systems.	L4
10	3	Explain schema evolution.	L3
11	3	Explain different types of data sources.	L3



**SREENIVASA INSTITUTE OF TECHNOLOGY AND MANAGEMENT STUDIES  
(AUTONOMOUS)**

**QUESTION BANK**

**Year / Semester: II B.Tech IV Semester**

**Regulation: R23**

**Subject and Code: DATA ENGINEERING, 23CSD242**

S.No.	CO	Questions	BT
<b>Unit IV: (STORAGE AND INGESTION)</b>			
1	4	Compare Data Warehouse, Data Lake and Lakehouse.	<b>L4</b>
2	4	Design a simple batch ingestion system.	<b>L3</b>
3	4	Explain problems in large data storage systems.	<b>L4</b>
4	4	Design a simple real-time ingestion system.	<b>L3</b>
5	4	Create a storage plan for a company.	<b>L5</b>
6	4	Explain cost problems in data storage.	<b>L4</b>
7	4	Explain batch ingestion clearly.	<b>L3</b>
8	4	Design a disaster recovery plan.	<b>L5</b>
9	4	Evaluate advantages of Lakehouse.	<b>L4</b>
10	4	Explain stream ingestion clearly.	<b>L3</b>
11	4	Explain security issues in storage systems.	<b>L3</b>



**SREENIVASA INSTITUTE OF TECHNOLOGY AND MANAGEMENT STUDIES  
(AUTONOMOUS)**

**QUESTION BANK**

**Year / Semester: II B.Tech IV Semester**

**Regulation: R23**

**Subject and Code: DATA ENGINEERING, 23CSD242**

S.No.	CO	Questions	BT
<b>Unit V: (QUERIES, MODELING AND TRANSFORMATION)</b>			
1	5	Analyze query optimization techniques.	L4
2	5	Design a star schema for a sales system.	L3
3	5	Compare batch and streaming transformations.	L4
4	5	Explain the life cycle of a query step by step.	L3
5	5	Design a simple analytics platform.	L5
6	5	Evaluate normalization versus denormalization.	L4
7	5	Explain data modeling with simple example.	L3
8	5	Design a fraud detection data model.	L5
9	5	Evaluate different data models for ML systems.	L4
10	5	Explain data transformation with example.	L3
11	5	Explain Reverse ETL with example.	L3

Note: L1-Remembering, L2-Understanding, L3-Applying, L4-Analyzing, L5-Evaluating, and L6-Creating



**Year / Semester: II B.Tech IV Semester**

**Regulation: R23**

**Subject and Code: DATA ENGINEERING, 23CSD242**

## **Instruction to Faculty Members:**

### **The Six Levels of Bloom's Taxonomy:**

1. **Remembering:** Retrieving, recognizing, and recalling relevant knowledge from long-term memory (e.g., list, define, name, locate).
2. **Understanding:** Constructing meaning, explaining ideas, or concepts (e.g., summarize, interpret, classify, compare).
3. **Applying:** Using information in new situations or implementing procedures to solve problems (e.g., solve, use, demonstrate, implement).
4. **Analyzing:** Breaking material into constituent parts, determining how the parts relate to one another and to an overall structure (e.g., contrast, categorize, distinguish, diagram).
5. **Evaluating:** Making judgments based on criteria and standards through checking and critiquing (e.g., judge, critique, justify, defend, argue).
6. **Creating:** Putting elements together to form a coherent or functional whole; reorganizing elements into a new pattern or structure (e.g., design, construct, develop, formulate).