

SRM VALLIAMMAI ENGINEERING COLLEGE

(An Autonomous Institution)
SRM Nagar, Kattankulathur – 603 203

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

QUESTION BANK



V SEMESTER

PIT201-DATA WAREHOUSING

Regulation – 2023

Academic Year 2025-2026 (ODD Semester)

Prepared by

Dr. V. Dhanakoti, Professor / CSE

Dr. A. Samydurai, Professor / CSE



SRM VALLIAMMAI ENGINEERING COLLEGE

(An Autonomous Institution)

SRM Nagar, Kattankulathur – 603 203.

DEPARTMENT OF INFORMATION TECHNOLOGY

QUESTION BANK

SUBJECT : PIT201 DATA WAREHOUSING

SEM / YEAR: V Sem / III Year

UNIT I -INTRODUCTION TO DATA WAREHOUSE			
Data warehouse Introduction - Data warehouse components- operational database Vs data warehouse – Data warehouse Architecture – Three-tier Data Warehouse Architecture - Autonomous Data Warehouse- Autonomous Data Warehouse Vs Snowflake - Modern Data Warehouse			
PART – A			
Q. No.	Questions	BT Level	Competence
1.	Define a Data Warehouse.	BTL1	Remember
2.	List two key characteristics of a data warehouse.	BTL1	Remember
3.	Differentiate between OLTP and OLAP.	BTL2	Understand
4.	State the purpose of a data warehouse.	BTL1	Remember
5.	Name any two components of a data warehouse	BTL1	Remember
6.	Explain the role of ETL in data warehousing.	BTL2	Understand
7.	What is the function of metadata in a data warehouse?	BTL2	Understand
8.	State the difference between a data mart and a data warehouse.	BTL2	Understand
9.	List two differences between operational databases and data warehouses	BTL2	Understand
10.	What is meant by non-volatile in the context of a data warehouse?	BTL1	Remember
11.	Mention one example each of operational database and data warehouse.	BTL1	Remember
12.	Compare data update frequency in operational DB vs data warehouse.	BTL2	Understand
13.	Name the three layers of data warehouse architecture.	BTL1	Remember

14.	What is the role of the staging area in data warehousing?	BTL2	Understand
15.	Differentiate between centralized and distributed data warehouse architecture	BTL2	Understand
16.	Define data warehouse architecture.	BTL1	Remember
17.	List the three tiers in a three-tier data warehouse architecture	BTL1	Remember
18.	Explain the function of the middle tier in the three-tier architecture	BTL2	Understand
19.	Which tier handles data presentation in a three-tier architecture?	BTL1	Remember
20.	Differentiate between bottom and top tier of data warehouse architecture.	BTL2	Understand
21.	What is an autonomous data warehouse?	BTL1	Remember
22.	Mention two benefits of using autonomous data warehouse.	BTL2	Understand
23.	Compare Autonomous Data Warehouse with Snowflake on scalability.	BTL2	Understand
24.	State one difference between ADW and Snowflake in terms of architecture	BTL2	Understand
PART-B			
1.	Apply the concept of OLAP to design a simple reporting system for a retail company.	BTL3	APPLY
2.	Analyze the differences in performance when querying transactional data vs warehoused data.	BTL4	Analyze
3.	Demonstrate how ETL tools can be used to clean and load sales data into a warehouse.	BTL3	Apply
4.	Analyze the role of metadata in enhancing data governance in a Warehouse system.	BTL4	Analyze
5.	Evaluate the effectiveness of a data mart versus an enterprise data warehouse for a mid-sized business.	BTL5	Evaluate
6.	Compare and contrast operational and analytical systems based on workload characteristics.	BTL4	Analyze
7.	Justify the use of a separate data warehouse instead of extending an operational database for reporting.	BTL5	Evaluate
8.	Create a basic architecture for a healthcare data warehouse handling patient and clinical data.	BTL3	Apply
9.	Analyze how data flows through the layers of a data warehouse architecture.	BTL4	Analyze

10.	Design a data warehouse architecture that supports real-time analytics for a logistics company.	BTL5	Create
11.	Examine the responsibilities of each tier in a three-tier data warehouse and how they interact.	BTL4	Analyze
12.	Assess the advantages of a three-tier architecture compared to a two-tier model.	BTL5	Evaluate
13.	Use the features of Oracle ADW to automate data backup and scaling in a given scenario.	BTL3	Apply
14.	Evaluate the benefits and limitations of using Autonomous Data Warehouse in large-scale enterprise environments.	BTL5	Evaluate
15.	Analyze the differences between ADW and Snowflake in terms of performance tuning and storage architecture.	BTL4	Analyze
16.	Evaluate which platform—ADW or Snowflake—is better suited for a startup with variable workloads.	BTL5	Evaluate
17.	Design a modern data warehouse solution using cloud-native tools to support a multi-channel e-commerce business.	BTL6	Create

UNIT II - ETL AND OLAP TECHNOLOGY

What is ETL – ETL Vs ELT – Types of Data warehouses - Data warehouse Design and Modeling - Delivery Process - Online Analytical Processing (OLAP) - Characteristics of OLAP - Online Transaction Processing (OLTP) Vs OLAP - OLAP operations- Types of OLAP- ROLAP Vs MOLAP Vs HOLAP.

PART – A

Q. No.	Questions	BT Level	Competence
1.	Define ETL in data warehousing.	BTL1	Remember
2.	Explain the role of ETL in data processing.	BTL2	Understand
3.	What does ELT stand for in data integration?	BTL1	Remember
4.	Distinguish between ETL and ELT with one key difference.	BTL2	Understand
5.	List any two types of data warehouses.	BTL1	Remember
6.	Describe the characteristics of a virtual data warehouse.	BTL2	Understand
7.	What is meant by star schema in data modeling?	BTL1	Remember
8.	Differentiate between star and snowflake schema in warehouse design.	BTL2	Understand

9.	State what is meant by the delivery process in data warehousing.	BTL1	Remember
10.	Summarize the steps involved in the delivery process of a data warehouse.	BTL2	Understand
11.	Define OLAP.	BTL1	Remember
12.	Explain how OLAP supports decision-making.	BTL2	Understand
13.	Mention any two key characteristics of OLAP systems.	BTL1	Remember
14.	Describe the concept of multidimensional data in OLAP.	BTL2	Understand
15.	Write any one difference between OLTP and OLAP systems.	BTL1	Remember
16.	Compare OLTP and OLAP based on data storage structure.	BTL2	Understand
17.	What is the roll-up operation in OLAP?	BTL1	Remember
18.	Illustrate the difference between drill-down and roll-up with an example.	BTL2	Understand
19.	List the three major types of OLAP systems.	BTL1	Remember
20.	Explain the basic idea behind Hybrid OLAP (HOLAP).	BTL2	Understand
21.	State one difference between ROLAP and MOLAP.	BTL1	Remember
22.	Contrast ROLAP, MOLAP, and HOLAP in terms of data storage.	BTL2	Understand
23.	Identify any one OLAP operation and define it briefly.	BTL1	Remember
24.	Interpret how a data warehouse supports OLAP functionality.	BTL2	Understand
PART-B			
1.	Apply the ETL process to a business scenario and explain each phase—Extract, Transform, and Load—with examples.	BTL3	Apply
2.	Analyze the differences between ETL and ELT processes. Justify which is more suitable for cloud-based data warehousing.	BTL4	Analyze
3.	Illustrate the types of data warehouses—Enterprise Data Warehouse, ODS, and Data Mart—with practical use cases	BTL-3	Apply
4.	Evaluate top-down and bottom-up approaches in data warehouse design. Discuss their pros and cons.	BTL-5	Evaluate
5.	Compare star schema and snowflake schema. Analyze which model is Better for performance and query optimization.	BTL-4	Analyze
6.	Analyze the major stages in the delivery process of a data warehouse. Explain how it ensures successful deployment.	BTL-4	Analyze
7.	Apply the OLAP concept to a business intelligence system. Explain how it enhances analytical decision-making.	BTL-3	APPLY

8.	Analyze the key characteristics of OLAP systems. How do these features support multidimensional data analysis?	BTL-4	Analyze
9.	Evaluate the architectural and functional differences between OLTP and OLAP with suitable examples.	BTL-5	Evaluate
10.	Demonstrate and explain OLAP operations—roll-up, drill-down, slice, dice, and pivot—using business data examples.	BTL-3	Apply
11.	Analyze different types of OLAP systems: MOLAP, ROLAP, and HOLAP. Explain their architecture and application areas.	BTL-4	Analyze
12.	Evaluate ROLAP and MOLAP based on performance, scalability, and storage techniques	BTL-5	Evaluate
13.	Compare MOLAP and HOLAP. Analyze the trade-offs involved in their implementation in modern data environments	BTL-4	Analyze
14.	Evaluate the advantages and disadvantages of ROLAP and HOLAP. Suggest where each type is best used.	BTL-5	Evaluate
15.	Apply OLAP tools in a real-time application (e.g., sales or healthcare) and explain their role in trend analysis.	BTL-3	APPLY
16.	Evaluate the importance of metadata in a data warehouse. How does it support data consistency and usability?	BTL-5	Evaluate
17.	Create a solution strategy to overcome key challenges in data warehouse implementation, including data integration and performance.	BTL-6	Create

**UNIT III - META DATA, DATA MART AND PARTITION
STRATEGY**

Meta Data – Categories of Metadata – Role of Metadata – Metadata Repository – Challenges for Meta Management - Data Mart – Need of Data Mart- Cost Effective Data Mart- Designing Data Marts- Cost of Data Marts- Partitioning Strategy – Vertical partition – Normalization – Row Splitting – Horizontal Partition

PART – A

Q. No.	Questions	BT Level	Competence
1.	What is metadata?	BTL1	Remember
2.	how metadata is used in data warehousing.	BTL2	Understand
3.	List the main categories of metadata.	BTL1	Remember
4.	Differentiate between technical and business metadata.	BTL2	Understand
5.	State any two roles of metadata in a data warehouse.	BTL1	Remember

6.	Explain the importance of metadata in data integration.	BTL2	Understand
7.	What is a metadata repository?	BTL1	Remember
8.	Describe the function of a metadata repository in data warehousing.	BTL2	Understand
9.	Mention two challenges in managing metadata.	BTL1	Remember
10.	Explain any one difficulty in maintaining metadata across platforms.	BTL2	Understand
11.	Define a data mart.	BTL1	Remember
12.	Differentiate between data mart and data warehouse.	BTL2	Understand
13.	Why is a data mart needed?	BTL1	Remember
14.	Explain how a data mart supports departmental decision-making.	BTL2	Understand
15.	What makes a data mart cost-effective?	BTL1	Remember
16.	Identify two ways to reduce the cost of building a data mart.	BTL2	Understand
17.	List any two key steps in designing a data mart.	BTL1	Remember
18.	Describe the importance of user requirements in designing data marts.	BTL2	Understand
19.	Name any two factors influencing the cost of a data mart.	BTL1	Remember
20.	Explain how scalability affects the cost of data marts.	BTL2	Understand
21.	What is a partitioning strategy in data warehousing?	BTL1	Remember
22.	Explain the significance of partitioning in query optimization.	BTL2	Understand
23.	Define vertical partitioning.	BTL2	Understand
24.	Differentiate between vertical and horizontal partitioning.	BTL1	Remember
PART-B			
1.	Apply the concept of metadata in a data warehouse environment. How does metadata support the functioning of the warehouse?	BTL3	APPLY
2.	Analyze the different categories of metadata. How do technical, business, and operational metadata work together?	BTL4	Analyze
3.	Evaluate the role of metadata in improving data warehouse performance, consistency, and user experience.	BTL5	Apply
4.	Analyze the structure and functions of a metadata repository. Explain its importance in a large-scale data system.	BTL4	Analyze
5.	Evaluate the challenges involved in metadata management across distributed systems. Suggest solutions to address these challenges.	BTL5	Evaluate
6.	Apply the concept of data mart to a specific business unit (e.g., finance or marketing) and explain how it supports decision-making.	BTL3	Apply
7.	Analyze the reasons for using a data mart instead of a full data warehouse. When is it more beneficial?	BTL4	Analyze

8.	Evaluate techniques to design a cost-effective data mart. How can organizations balance performance and budget?	BTL5	Evaluate
9.	Design a data mart for a university or retail store. Justify your design choices with schema and access needs.	BTL6	Create
10.	Analyze the cost factors associated with data mart development, deployment, and maintenance.	BTL4	Analyze
11.	Apply a partitioning strategy to a large customer data table. Explain how it improves query performance.	BTL3	Apply
12.	Analyze the use of vertical partitioning in data storage. How does it affect query speed and storage efficiency?	BTL4	Analyze
13.	Analyze how normalization impacts data redundancy and integrity in a warehouse environment.	BTL4	Analyze
14.	Apply the row-splitting technique in data partitioning and explain how it helps in better data access and maintenance.	BTL3	Apply
15.	Compare horizontal and vertical partitioning techniques. Analyze which is suitable for large time-series data.	BTL4	Analyze
16.	Evaluate the interrelationship between metadata and data marts in the data warehousing lifecycle.	BTL5	Evaluate
17.	Propose a metadata management strategy for a multi-source enterprise data warehouse system.	BTL6	Create

UNIT IV-DIMENSIONAL MODELING AND SCHEMA

Dimensional Modeling- Multi-Dimensional Data Modeling – Data Cube- Star Schema- Snowflake schema- Star Vs Snowflake schema- Fact constellation Schema- Schema Definition - Process Architecture- Types of Data Base Parallelism – Datawarehouse Tools.

PART - A

Q. No.	Questions	BT Level	Competence
1.	What is dimensional modeling in data warehousing?	BTL1	Remember
2.	Explain the purpose of dimensional modeling.	BTL2	Understand
3.	Define multi-dimensional data modeling.	BTL1	Remember
4.	Explain how multi-dimensional modeling helps in OLAP operations.	BTL2	Understand
5.	What is a data cube?	BTL1	Remember
6.	Describe the use of slicing and dicing in a data cube.	BTL2	Understand
7.	What is a star schema?	BTL1	Remember
8.	Explain the structure of a star schema with an example.	BTL2	Understand
9.	Define snowflake schema in dimensional modeling.	BTL1	Remember
10.	How does snowflake schema handle dimension normalization?	BTL2	Understand

11.	List two differences between star and snowflake schema.	BTL1	Remember
12.	Compare star and snowflake schema based on performance.	BTL2	Understand
13.	What is a fact constellation schema?	BTL1	Remember
14.	Explain how fact constellation schema supports complex analysis.	BTL2	Understand
15.	Define schema in the context of a data warehouse.	BTL1	Remember
16.	Why is schema design important in data warehousing?	BTL2	Understand
17.	What is process architecture in a data warehouse?	BTL1	Remember
18.	Describe the role of ETL in process architecture.	BTL2	Understand
19.	List the types of parallelism in databases.	BTL1	Remember
20.	Explain inter-query and intra-query parallelism with examples.	BTL2	Understand
21.	Name any two popular data warehouse tools.	BTL1	Remember
22.	Explain how data warehouse tools support analytics and reporting.	BTL2	Understand
23.	Define a fact table and a dimension table.	BTL1	Remember
24.	Explain the relationship between OLAP tools and dimensional modeling.	BTL2	Understand
PART-B			
1.	Apply dimensional modeling to a retail business scenario. Explain the structure with fact and dimension tables.	BTL3	Apply
2.	Analyze the principles of multi-dimensional data modeling. How does it support OLAP operations?	BTL4	Analyze
3.	Discuss the concept of a data cube. Analyze slicing, dicing, drill-down, and roll-up operations with examples.	BTL4	Analyze
4.	Construct a star schema for a sales organization. Explain the role of fact and dimension tables in your design.	BTL3	Apply
5.	Analyze a snowflake schema structure. What are the advantages and disadvantages compared to star schema?	BTL4	Analyze
6.	Evaluate star and snowflake schemas in terms of performance, query complexity, and normalization.	BTL5	Evaluate
7.	Analyze the structure of a fact constellation schema. How does it manage multiple fact tables efficiently?	BTL4	Analyze
8.	Define and apply different data warehouse schema types in a real-world business context.	BTL3	Apply

9.	Analyze the architecture of a data warehouse process including ETL, data staging, and presentation layers.	BTL4	Analyze
10.	Explain and analyze the types of database parallelism—interquery, intraquery, intraprocessor, and interprocessor—with examples.	BTL4	Analyze
11.	Apply interquery and intraquery parallelism concepts in a large database query execution context.	BTL3	Apply
12.	Evaluate the effectiveness of parallel processing in improving data warehouse query performance.	BTL5	Evaluate
13.	Apply any two data warehouse tools (e.g., Informatica, Power BI, Tableau) to support analytics and decision-making.	BTL3	Apply
14.	Evaluate different data warehouse tools based on usability, integration capabilities, and performance.	BTL5	Evaluate
15.	Design a data warehouse schema for a hospital or banking system using dimensional modeling. Justify your schema selection.	BTL6	Create
16.	Evaluate various schema designs in terms of normalization, storage efficiency, and query performance.	BTL5	Evaluate
17.	Create a dimensional model for an e-commerce platform. Include data cube operations and justify the use of schema type.	BTL6	Create

UNIT V SYSTEM & PROCESS MANAGERS

Data Warehousing System Managers: System Configuration Manager- System Scheduling Manager - System Event Manager - System Database Manager - System Backup Recovery Manager - Data Warehousing Process Managers: Load Manager – Warehouse Manager- Query Manager – Tuning – Testing.

PART - A

Q. No.	Questions	BT Level	Competence
1.	What is the role of a system configuration manager in data warehousing?	BTL1	Remember
2.	how the configuration manager supports system resource allocation.	BTL2	Understand
3.	Define system scheduling in a data warehouse.	BTL1	Remember
4.	Describe the function of a scheduling manager in task automation.	BTL2	Understand
5.	What is an event manager in a data warehouse system?	BTL1	Remember
6.	Explain the significance of event handling in warehouse operations.	BTL2	Understand

7.	State the role of a system database manager.	BTL1	Remember
8.	how the database manager ensures data consistency.	BTL2	Understand
9.	What is the purpose of the backup recovery manager?	BTL1	Remember
10.	Discuss how backup and recovery are handled in case of system failure.	BTL2	Understand
11.	Define the function of a load manager in data warehousing.	BTL1	Remember
12.	Explain the role of the load manager in data preprocessing.	BTL2	Understand
13.	What is the function of a warehouse manager?	BTL1	Remember
14.	Describe how the warehouse manager maintains data consistency.	BTL2	Understand
15.	Who is a query manager in a data warehouse environment?	BTL1	Remember
16.	Explain how the query manager optimizes user queries.	BTL2	Understand
17.	What is tuning in the context of a data warehouse?	BTL1	Remember
18.	Explain the importance of tuning for performance enhancement.	BTL2	Understand
19.	What is the objective of testing a data warehouse system?	BTL1	Remember
20.	Describe any two key areas of focus during data warehouse testing.	BTL2	Understand
21.	List any two components of the data warehouse system management layer.	BTL1	Remember
22.	how system and process managers work together in a warehouse environment.	BTL2	Understand
23.	Name two data warehousing process managers.	BTL1	Remember
24.	Explain how the load manager and query manager contribute to user experience.	BTL2	Understand
PART-B			
1.	How would you apply configuration settings to improve data warehouse performance?	BTL3	Apply
2.	Analyze the impact of incorrect system configuration in a data warehouse environment.	BTL4	Analyze
3.	Apply a scheduling strategy for running ETL jobs during off-peak hours.	BTL3	Apply
4.	Analyze how scheduling failures can affect warehouse operations.	BTL4	Analyze
5.	Analyze the role of the event manager in handling data warehouse alerts and system events.	BTL4	Analyze
6.	Evaluate the effectiveness of event-driven automation in reducing downtime.	BTL5	Evaluate
7.	How does a database manager maintain metadata and indexes for	BTL4	Analyze

	optimization?		
8.	Evaluate the contribution of the database manager to query performance.	BTL5	Evaluate
9.	Apply a disaster recovery plan using backup recovery manager functionalities.	BTL3	apply
10.	Evaluate the pros and cons of automated vs manual recovery processes.	BTL5	Evaluate
11.	Analyze the responsibilities of the load manager in staging and data cleansing.	BTL4	Analyze
12.	Evaluate load balancing strategies to improve loading performance.	BTL5	Evaluate
13.	Analyze how the warehouse manager handles aggregations and data summarization.	BTL4	Analyze
14.	Evaluate the role of query manager in optimizing query execution plans.	BTL5	Evaluate
15.	Propose a tuning technique to resolve slow-running queries in a warehouse.	BTL6	Create
16.	Create a checklist for data warehouse testing after ETL completion.	BTL6	Create
17.	Evaluate the importance of regression testing in the warehouse update process.	BTL5	Evaluate