Printed Pages – 5 Roll No. : 30012821045

D128871(022)

B. Tech. (Hon's) (Eighth Semester) Examination, April-May 2025

(CSE : Data Science Branch)

DATA WAREHOUSING

Time Allowed : Three hours

Maximum Marks : 100

Minimum Pass Marks : 35

Note : *Attempt all questions. Part (a) of each question is compulsory, each of 4 marks. Attempt any* ***two*** *parts from (b), (c) and (d) each of 8 marks.*

Unit-I

1. (a) Differentiate between OLAP and OLTP systems with two examples. 4

(b) Define data warehousing. Explain its need in business intelligence. How does it support decision-making processes? 8

(c) Compare and contrast the architecture of a centralized enterprise data warehouse and a virtual data warehouse. Discuss their pros and cons. 8

(d) Design a high-level architecture for a data warehousing system for a retail company. Your design should identify key components and justify the inclusion of each component. 8

Unit-II

2. (a) Briefly describe Slowly Changing Dimensions (SCD). Give one real-world example for Type 1 and Type 2 SCDs. 4

(b) Differentiate between star schema and snowflake schema. List their advantages and disadvantages with appropriate illustrations. 8

(c) Analyze the role of fact tables and dimension tables in a data warehouse. How do their design choices affect data retrieval and query performance? 8

(d) Design a fact constellation schema for a university management system that includes student admissions, course registrations, and faculty payroll. Justify your schema design and explain how query optimization can be supported. 8

Unit-III

3. (a) What is the role of ETL in a data warehousing system? List its main phases with a brief explanation. 4

(b) Describe different data extraction methods used in ETL processes. How do full extraction and incremental extraction differ? 8

(c) Compare and evaluate at least two popular ETL tools in terms of performance, usability, and industry adoption. 8

(d) Design a basic ETL process for integrating customer data from multiple regional databases into a centralized data warehouse. Your design should outline extraction, transformation, and loading steps along with justifications. 8

Unit-IV

4. (a) Differentiate between MOLAP, ROLAP, and HOLAP based on storage, performance, and flexibility. 4

(b) Explain the OLAP architecture. Discuss the roles of data sources, OLAP server, and front-end tools in the architecture. 8

(c) Evatuate the benefits and limitations of MOLAP, ROLAP, and HOLAP approaches. Which approach is more suitable for real-time analytics in large enterprises and why? 8

(d) Design a multidimensional OLAP model for a retail chain to support sales analysis. Include.appropriate dimensions and measures, and explain how OLAP operations like slice, dice, and drill-down would be used for decision-making. 8

Unit-V

5. (a) List any four key issues encountered during the implementation of a data warehouse and briefly explain them. 4

(b) Explain various data warehouse implementation strategies such as top-down, bottom-up, and hybrid approaches. Discuss their advantages and disadvantages. 8

(c) Discuss the process of data warehouse testing. What are the major types of testing carried out to ensure the reliability and accuracy of data? 8

(d) Evaluate the impact of big data technologies on traditional data warehousing. Propose how a modern enterprise can integrate big data with its existing data warehouse to stay competitive. 8