

Based on the **GTU Paper Format** and standard exam patterns for Diploma Engineering, Unit 4: **Data Warehouse Design** focuses on the structural and architectural choices that make data ready for analysis.

In a 70-mark paper, Unit 4 aligns with **CO4** and typically contributes **10–14 marks**. According to the provided paper format, this unit often appears in **Question 3** and **Question 4**, frequently featuring **Remember (R)** level short questions and **Understand (U)** level descriptive or diagram-based questions.

---

## **Unit 4: Predicted Question Bank (Data Warehouse Design)**

### **1. Most Repeated / High-Probability Questions**

*These questions are structured based on the (a) 3-mark, (b) 4-mark, and (c) 7-mark format of the model paper.*

#### **[Short Answer Type - 03 Marks]**

- **Define Dimensional Modeling.** How does it differ from traditional E-R modeling?
- **What is a Fact Table?** List its two main components (measures and foreign keys).
- **Define a Surrogate Key.** Why is it preferred over a natural key in a data warehouse?
- **What is a "Factless" Fact Table?** Give one example of where it is used.

#### **[Descriptive Type - 04 Marks]**

- **Explain the Star Schema** with a neat diagram. Mention one advantage and one disadvantage.
- **Describe the Snowflake Schema.** How does it differ from a Star Schema in terms of normalization?
- **Explain Fact Constellation (Galaxy Schema).** In what scenarios should a designer choose this schema?
- **Differentiate between Fact Tables and Dimension Tables** based on data type, size, and normalization.

#### **[Long Answer / Diagram-Based - 07 Marks]**

- **Explain the Four-Step Dimensional Design Process** (Select business process, Declare grain, Identify dimensions, Identify facts).
- **Compare Star, Snowflake, and Galaxy Schemas.** Provide a detailed comparison table and a diagram for each to illustrate the differences in complexity and performance.
- **Discuss the concept of 'Grain' in a Data Warehouse.** Why is "declaring the grain"

considered the most important step in design?

---

## 2. Application & Logical Thinking Questions

*These questions distinguish high-scoring students by testing their ability to interpret real-world systems.*

16. **Schema Selection Logic:** A retail chain has a very complex hierarchy for its products (Category -> Sub-category -> Brand -> Item). To save storage space and maintain this hierarchy strictly, which schema would you recommend? **Justify your choice** between Star and Snowflake.
17. **Grain Interpretation:** In a "Sales" fact table, if the grain is defined as "**Individual Transaction Item**," can a manager use this table to see "Total Monthly Sales per Store"? Explain how aggregation works in this scenario.
18. **Problem Identification:** If a dimension table for "Customers" is frequently updated with new addresses, it might slow down the warehouse. Suggest a strategy (like **Slowly Changing Dimensions**) to handle this without losing historical data.
19. **System Architecture Reasoning:** You are designing a data warehouse for a hospital. The "Patient" dimension is shared by both the "Lab Tests" fact table and the "Pharmacy Sales" fact table. Which schema is being described here? Explain the benefit of **shared dimensions**.
20. **Real-world Modeling:** Create a simple **Star Schema diagram** for a "University Management System." Identify one central Fact table and at least three relevant Dimension tables.

---

### Exam Success Strategy for Unit 4:

- **Master the Diagrams:** A 7-mark question almost always asks for a schema diagram. Practice drawing the **central fact table** with lines radiating to **dimension tables** to ensure clarity.
- **Normalization vs. Performance:** Remember the core trade-off: **Star Schema** = High performance (less joins), **Snowflake Schema** = High normalization (less redundancy).
- **Outcome Mapping:** This unit maps to **CO4**. Ensure you can "design" a basic schema, as application-level questions often ask you to "suggest a schema for a given scenario".

**Next Step:** Since Unit 4 is highly visual, I can provide a **step-by-step guide on how to convert a standard E-R diagram into a Star Schema**. Would you like me to walk you through that process?