

Based on the provided **DMW GTU Paper Format** and standard Diploma Engineering examination trends, I have analyzed the exam patterns for **Unit 2: Basics of Data Warehousing**.

In a typical 70-mark paper, Unit 2 (linked to **CO2**) carries significant weight, frequently appearing as the primary focus of **Question 1** and **Question 2**, covering roughly **14 to 18 marks**.

Unit 2: Predicted Question Bank (Basics of Data Warehousing)

1. Most Repeated / High-Probability Questions

These questions are structured according to the (a) 3-mark, (b) 4-mark, and (c) 7-mark distribution found in the model paper.

[Short Answer Type - 03 Marks]

- **Define Data Warehouse.** Mention its primary purpose in a business organization.
- **List the four key characteristics** of a Data Warehouse as defined by Bill Inmon.
- **What is a Data Mart?** Briefly explain its relationship with a Data Warehouse.
- **Define 'Non-volatile' and 'Time-variant'** in the context of data warehousing.

[Descriptive Type - 04 Marks]

- **Differentiate between OLTP and OLAP.** Provide at least four technical points of comparison.
- **Explain the concept of 'Subject-Oriented' data.** Give an example of how a retail company organizes its subjects.
- **Describe the role of Metadata** in a Data Warehouse environment.
- **Briefly explain the 'Data Staging' area.** Why is it necessary before loading data into the warehouse?

[Long Answer / Diagram-Based - 07 Marks]

- **Explain the Three-Tier Architecture of a Data Warehouse** with a neat and labeled diagram.
- **Discuss the Top-Down (Inmon) vs. Bottom-Up (Kimball) design approaches.** Which

one is more suitable for quick implementation? Justify.

- **Describe the ETL (Extract, Transformation, Loading) process.** Explain each step's significance in ensuring data quality.

2. Application & Logical Thinking Questions

These questions are designed to test higher-order cognitive levels, specifically **Understanding (U)** and **Application (A)** as highlighted in the paper format.

6. **Scenario Analysis:** A multinational bank wants to analyze its last 10 years of loan data to predict future risks. Explain why a **Data Warehouse** is a better choice for this task than their daily transactional SQL database.
7. **Architecture Choice:** You are tasked with designing a data warehouse for a small startup that needs results in just two months. Would you recommend a **Virtual Warehouse** or a **Physical Data Mart**? Justify your choice based on cost and time.
8. **Data Consistency:** During the **Integration** phase of ETL, you find that "Gender" is stored as "M/F" in one source and "1/0" in another. Explain the specific transformation steps needed to make this data consistent for the warehouse.
9. **System Performance:** If a manager complains that OLAP reports are slowing down the bank's daily ATM transactions, what architectural mistake has been made? Suggest a solution based on the **Separation of OLTP and OLAP**.
10. **Historical Logic:** Explain how the '**Time-Variant**' nature of a warehouse allows a manager to compare the sales of "Festival Season 2023" with "Festival Season 2024." Why is this impossible in a standard operational database?

Exam Success Strategy for Unit 2:

- **The "Three-Tier" Diagram:** This is the most likely 7-mark question. Always draw the source systems, the warehouse server, and the front-end tools clearly.
- **OLTP vs. OLAP Table:** Standardize a comparison table with headers like *User, Function, DB Design, Data, and Usage*.
- **Focus on 'Keywords':** When explaining characteristics, use the terms **Integrated**, **Non-volatile**, and **Summarized** to align with official marking schemes.