Roll No. 22DEC02	**Experiment No.** 01	**Marks :**
BATCH - C		**Sign :**

Aim: One case study on building Data warehouse/Data Mart

- Write Detailed Problem statement and design dimensional modeling (creation of star and snowflake schema)
- Implementation of all dimension table and fact table

Apparatus: Paint Tool software.

Theory:

What is a data mart?

A data mart is a data storage system that contains information specific to an organization's business unit. It contains a small and selected part of the data that the company stores in a larger storage system. Companies use a data mart to analyze department-specific information more efficiently. It provides summarized data that key stakeholders can use to quickly make informed decisions.
For example, a company might store data from various sources, such as supplier information, orders, sensor data, employee information, and financial records in their data warehouse or data lake. However, the company stores information relevant to, for instance, the marketing department, such as social media reviews and customer records, in a data mart.

Database

A database is organized storage that computer systems use to store, search, retrieve, and analyze information. There are various types of databases, such as relational databases. A relational database stores information in tables consisting of rows and columns. Data in different tables is connected by a unique identifier known as a key. Keys are the non-repetitive values in specific columns.

Data mart vs. database

A data mart serves as the front-facing element for a department's data. You can use a data mart to retrieve and analyze information. Meanwhile, a database collects, manages, and stores information. You can then use tools to process, format, and transfer the stored information to a data mart.

Data warehouse

A data warehouse is an extensive database system that stores information for an entire business. It collects raw information from various sources, such as business software and social media feeds, and processes it into structured data stored in a tabular format. Businesses can connect an enterprise data warehouse to business intelligence tools to make smarter decisions.

Data mart vs. data warehouse

A data mart shares many of the qualities of a data warehouse. Where they differ is that a data warehouse contains enterprise-wide data about various topics. Meanwhile, a data mart stores information closely related to a specific subject. For example, a data warehouse might store information for the marketing, human resources, procurement, and customer support departments. However, a data mart might store only transactional data relevant to a single department. The appeal of building a data mart is that departments who manage their data marts have complete control over the loading and management of their data.

Many organizations are using technologies like data sharing to publish their data marts to a central data warehouse. By doing so they can be more agile by distributing ownership and isolating workloads. Similarly, data sharing allows departmental data marts to consume data shared from a data warehouse or other data marts.

Why is a data mart important?

1. **Retrieve data more efficiently**
   By using a data mart, companies can access specific information more efficiently. Compared to a data warehouse, a data mart contains relevant and detailed information that a department accesses frequently. Therefore, business managers don't need to search the entire data warehouse to generate performance reports or graphics.

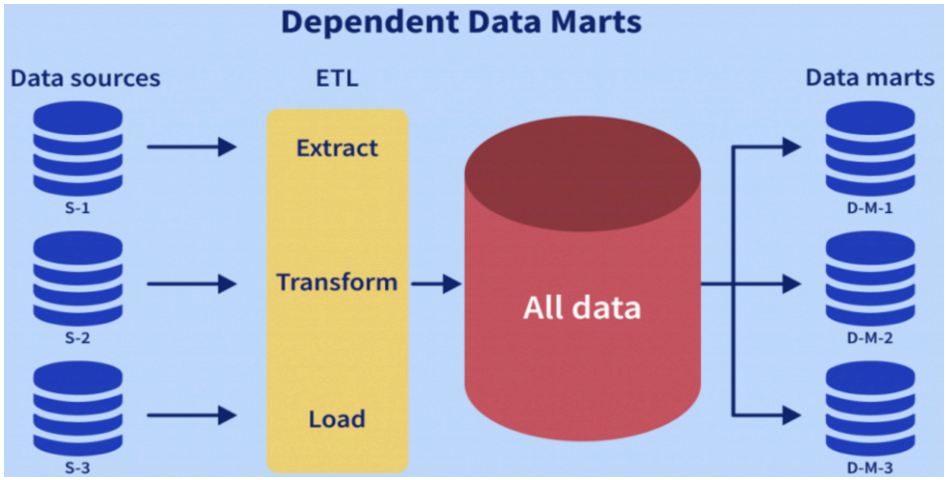
2. **Streamline decision-making**
   Companies can create a subset of data from a data warehouse with a data mart. Employees within the department can then analyze the data and make decisions based on the same set of information.

3. **Control information more effectively**
   A data mart gives employees highly granular access privileges. This means the company can authorize a certain person to view or retrieve specific data. It helps companies to improve data governance and enforce information access policies. For example, you can use data marts to provide user access to employees for specific information in a data warehouse.

4. **Manage data flexibly**
   A data mart is smaller and contains fewer tables than a data warehouse. This means data engineers can manage and change information in a data mart without causing major database changes.

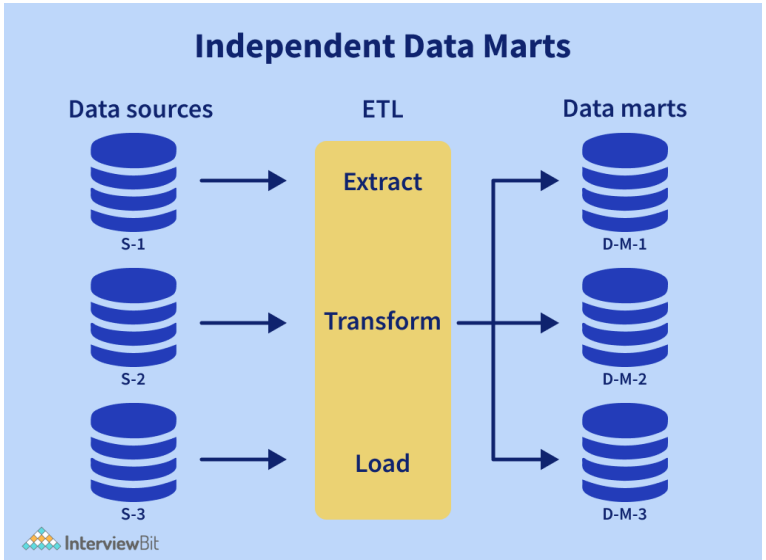
What are the types of data marts?

1. **Dependent data mart**
   A dependent data mart populates its storage with a subset of information from a centralized data warehouse. The data warehouse gathers all the information from data sources. Then, the data mart queries and retrieves subject-specific information from the data warehouse**.**

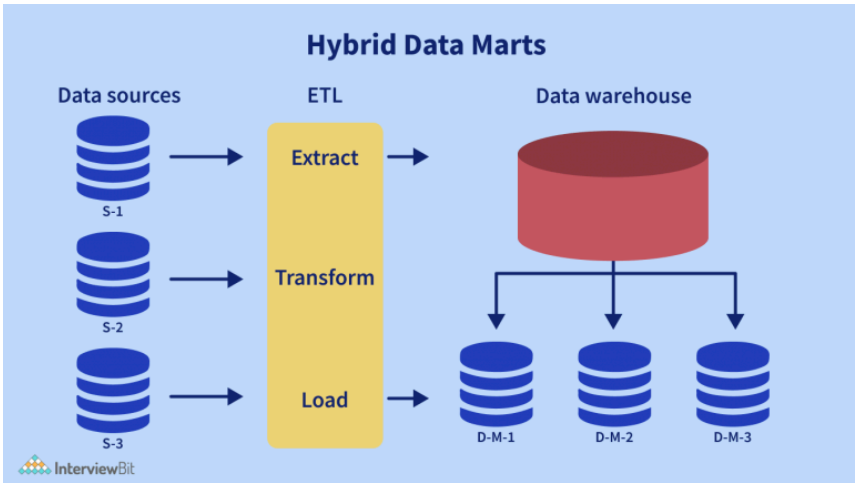


2. **Independent data mart**

   An independent data mart does not rely on a central data warehouse or any other data mart. Each data mart collects information from its sources instead of from a data warehouse. Independent data marts are suitable for smaller companies, but only specific departments need to access and analyze information.



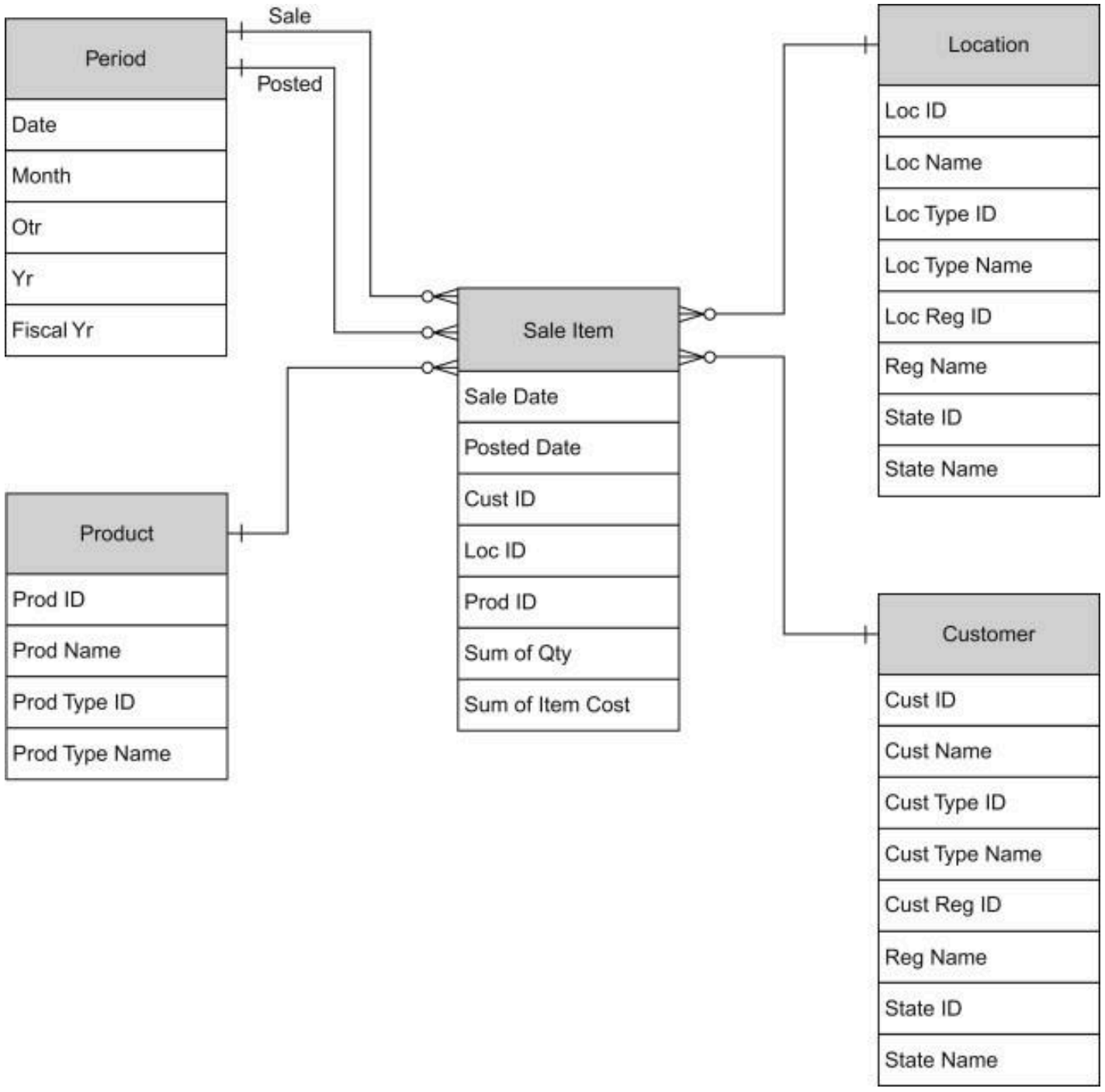
3. Hybrid data mart



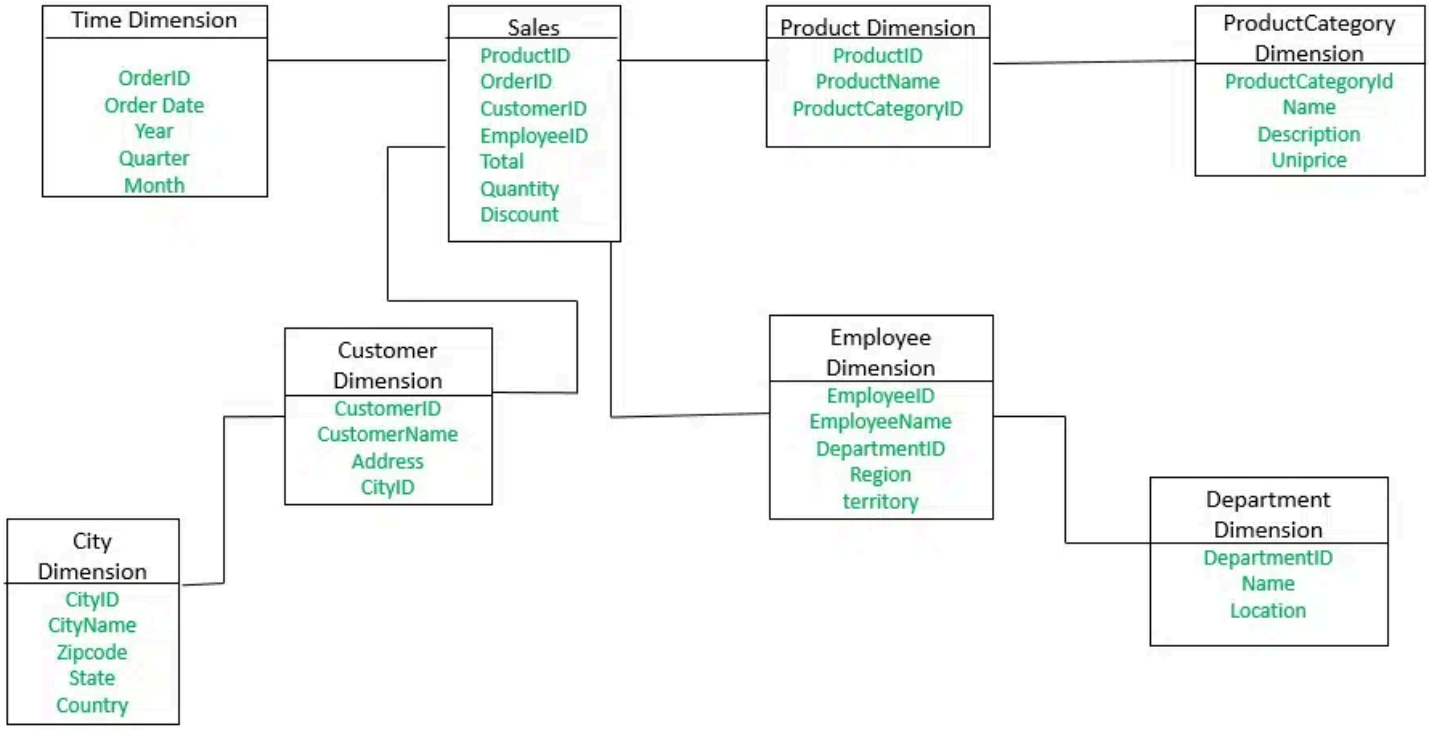
Hybrid data marts collect information from a data warehouse and from external sources. This allows companies the flexibility to test independent data sources before they direct the data to the data warehouse. For example, suppose you launch a new product and want to analyze its initial sales data. The data mart uses sales information that comes directly from the e-commerce software and retrieves sales records for other products from the data mart. After the product becomes a permanent fixture in your store, you channel the transaction details to the data warehouse.

Problem Statement:	Dimensional Fact table for Supermarket.

Star Schema



SnowFlake Schema



Conclusion:

In conclusion, the implementation of a data mart using either a star or snowflake schema offers the supermarket chain a powerful tool to enhance decision-making processes. By efficiently organizing and contextualizing sales data through dimension tables, such as product, store, and date dimensions, the company can gain valuable insights into sales performance and inventory management. Whether opting for the simplicity of the star schema or the enhanced normalization of the snowflake schema, the data mart will empower the sales team with quick access to relevant information, ultimately driving strategic decision-making and improving overall performance.