Name : Manjyot Wadhwa
Roll no : 74
Div: D15C

DBMI PRACTICAL 1

Aim:

Design a Star and Snowflake Schema for the given system

Theory:

Dimensional modelling is a data modelling technique specifically designed to optimize databases for data warehousing and business intelligence (BI). Unlike normalised database design (which aims to reduce redundancy for transactional systems), dimensional modelling focuses on:

- Fast query performance
- Ease of understanding for business users
- Historical tracking
- Aggregated reporting

It was popularized by Ralph Kimball, one of the pioneers of data warehousing
In dimensional modelling:

- Facts = measurable business data (e.g., sales amount, quantity sold).
- Dimensions = descriptive attributes that provide context to the facts (e.g., date, customer, product, store).

The data is organized in a Star Schema or Snowflake Schema, making it easier for analytics tools to query.

Types of Dimensional Modeling Schemas

1. Star Schema

Structure:
A central **Fact Table** is connected directly to multiple **Dimension Tables** .

- **Pros:**
  - Simple and intuitive design
  - Fewer joins → faster queries
  - Easy for non-technical business users to understand
- **Cons:**
  - Denormalized structure → higher storage usage
  - Data redundancy possible

2. Snowflake Schema

Structure:
Dimension tables are **normalized** into multiple related sub-dimension tables.

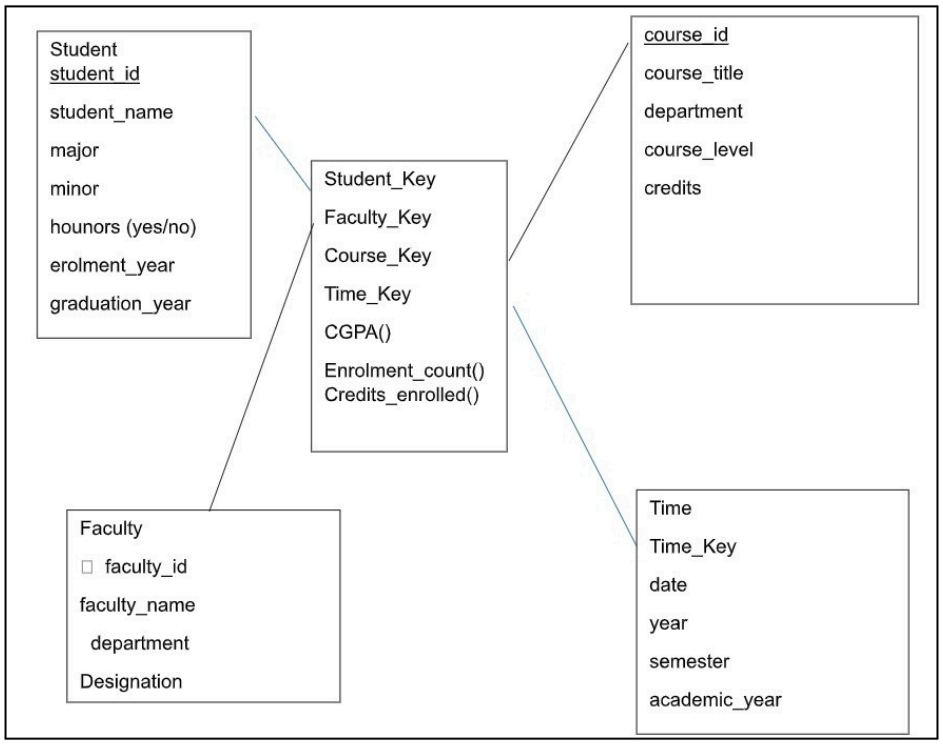
- **Pros:**
  - Saves storage space by eliminating redundancy
  - Better data integrity
- **Cons:**
  - More joins required → slightly slower queries
  - More complex for business users to navigate

3. Galaxy Schema (Fact Constellation)

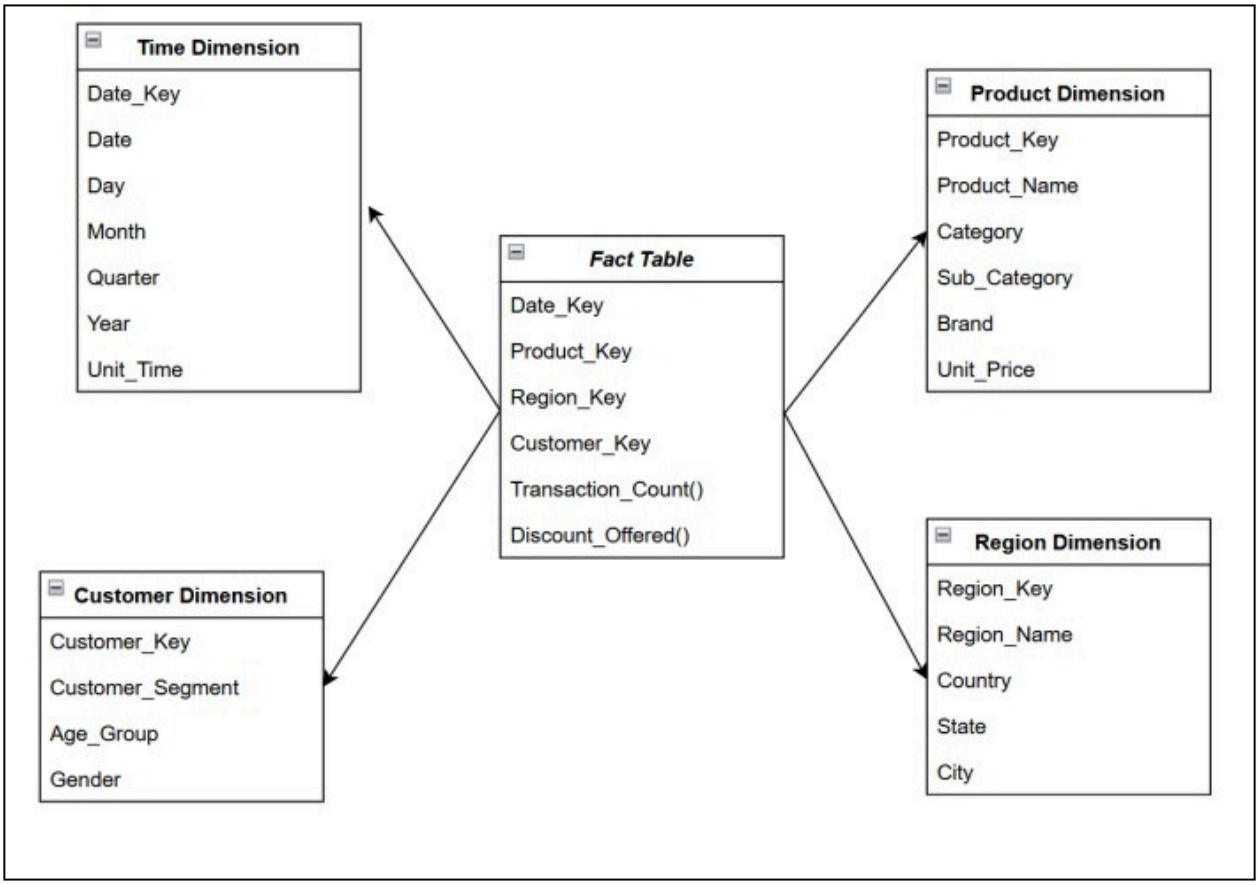
Structure:
Multiple **Fact Tables** share common **Dimension Tables** .

- **Use Case:**
  Enterprise-level data warehouses combining **multiple business processes** or **subject areas** .

Q1) A university wants to design a data warehouse to analyze student performance, course enrollments, and faculty workload. The university's operational database is highly normalized, making it difficult to perform analytical queries.



Q2) A retail company wants to analyze its sales performance across different regions, time periods, products, and customer segments. The company wants to track total sales, number of transactions, and discount offered.



Q3) A hospital management wants to create a data warehouse to analyze patient admissions, procedures, and billing information. The goal is to improve operational efficiency and patient care.