Yash Chikhale
D15C / 50
Batch C

DMBI - 1

Aim : Design Star schema for given problem statement

Theory :

Dimensional Modelling (DM) is a design technique used in data warehouse systems to optimize databases for fast retrieval and easy understanding by business users.
It structures data into facts and dimensions so that it's intuitive for reporting and analytics.

Operational (OLTP) databases are highly normalized to reduce redundancy and ensure consistency. But this makes analytical queries slow and complex because the data is spread across many tables.

Dimensional modelling:

- Reduces query complexity
- Improves performance (fewer joins)
- Makes data business-friendly
- Supports OLAP operations (slice, dice, drill-down, roll-up)

Fact Table

- Definition: Central table in a dimensional model that stores quantitative data (measures).
- Examples: Sales amount, GPA, Total bill, Quantity sold.
- Contains
    - Foreign keys to dimension tables

    - Measures (numeric values for analysis)
    - Granularity (level of detail)

Dimension Table

- Definition: Descriptive attributes (context) related to facts.
- Examples: Date, Product, Customer, Location, Course, Faculty.
- Contains:
    - Primary key
    - Textual attributes
    - Often denormalized for speed

Types of Dimensional Model

- Star Schema
- Snowflake Schema
- Galaxy Schema (a.k.a. Fact Constellation)

Star Schema

- Structure: One fact table in the center, surrounded by dimension tables.
- Advantages: Simple to understand ,fast queries ,fewer joins
- Disadvantages: Can have some data redundancy in dimensions

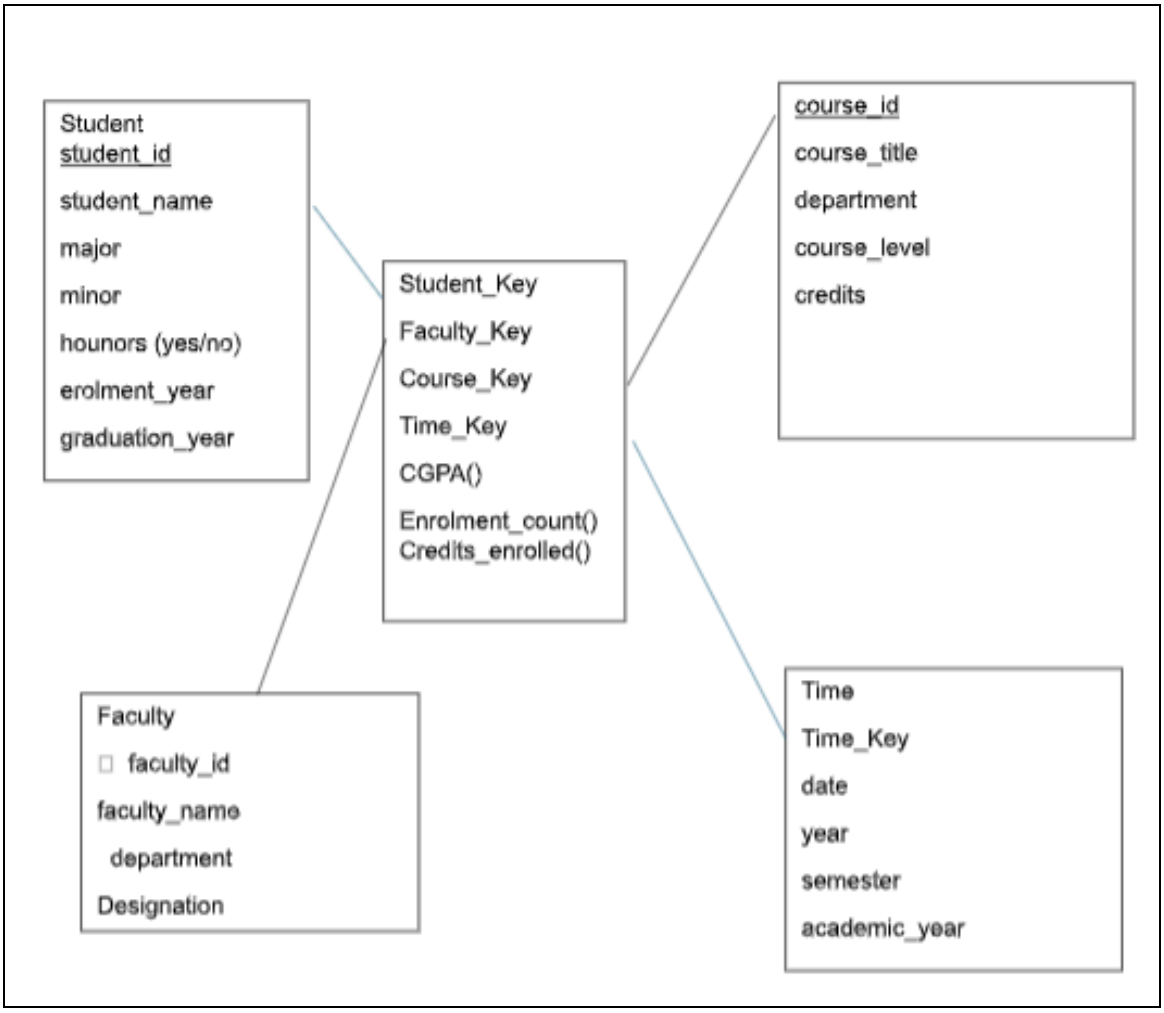
Snowflake Schema

- Structure: Similar to star schema, but dimension tables are normalized into sub-dimensions.
- Advantages: Saves storage space ,removes redundancy
- Disadvantages: More joins → slower queries, Less intuitive for business users

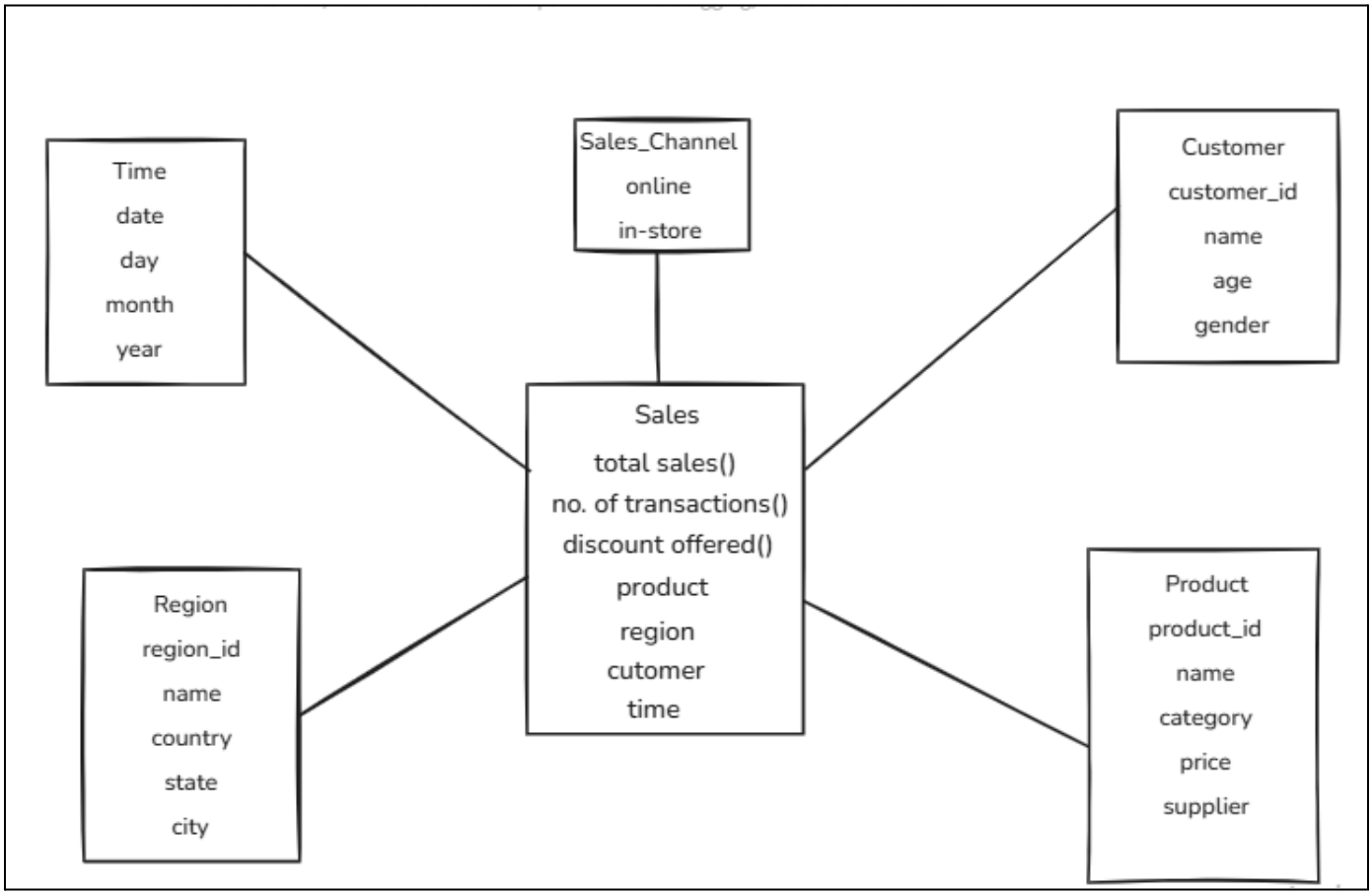
Galaxy Schema (Fact Constellation)

- Structure: Multiple fact tables share dimension tables.
- Advantages: Supports multiple business processes in one model ,reuses dimensions
- Disadvantages: More complex design

1. 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.



2. 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.



3. 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