

TOPIC NAME : Dimensional Modeling

DAY : _____

TIME : _____

DATE : / /

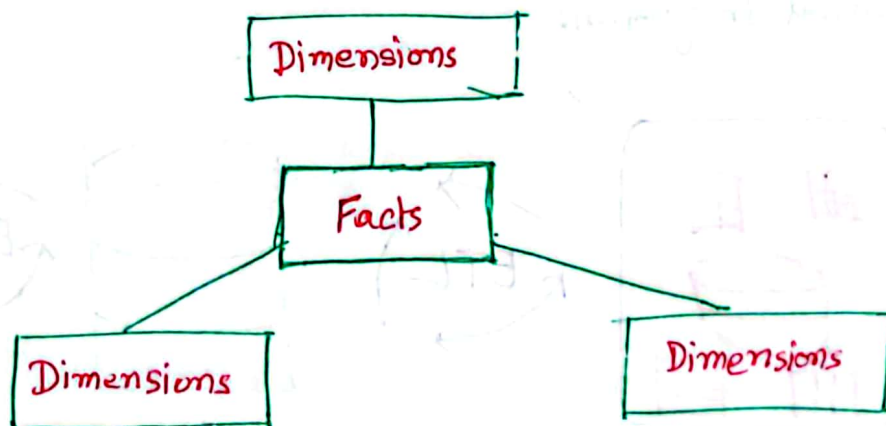
Dimensional Modeling:

- Unique technique of structuring data
- Commonly used in DWH
- Optimized for faster data retrieval
- Oriented around performance and usability
- Designed for reporting / OLAP

Why Dimensional Modeling:

The goal of dimensional modeling is fast data retrieval and oriented around performance and usability. So, we do dimensional modeling to structure data into tables in such a way that we get to do optimized query and no redundant info store in the tables which hampers the performance.

Star Schema:



TOPIC NAME : _____

DAY : _____

TIME : _____ DATE : / /

Facts:

- Aggregatable (numerical values)
- Measurable ~~vs. des~~
- Event - or transactional data
- Date/time columns included
- Consists of PK, FK and Facts.

Dimensions:

- Categorize facts
- Supportive and descriptive
- Filtering, Grouping and labelling
- Non aggregatable
- Consists of PK, Dimensions, FK
- People, products, places, times.

The problem of Star Schema:

- As star schema consists of only facts and dimensions, so, there can be data redundancy issues in the dimension tables.
- This can be solved by snowflake schema which is more normalized.

Benefits of normalization:

- Technique to avoid redundancy
- Minimizes storage
- Performance (write/Update)

Demerits of normalization:

- More joins of data needed
- Query become complex
- Performance get reduced.

Benefits of Star Schema/De normalization:

- Optimized to get data out
- Query performance (read)
- User experience is better.

TOPIC NAME : _____

DAY : _____

TIME : _____

DATE : / /

- Sometimes if needed, we can have multiple fact tables in the star schema.

Snowflake Schema : To reduce the data redundancy in star ~~st~~ schema we use snowflake schema.

Advantages :

- More ~~na~~ normalized
- Less redundant data
- Less space as less data redundancy
- Easier to maintain/update data
- Less risk of corrupted data

Disadvantages :

- More complex query to find data
- More joins
- Less performance