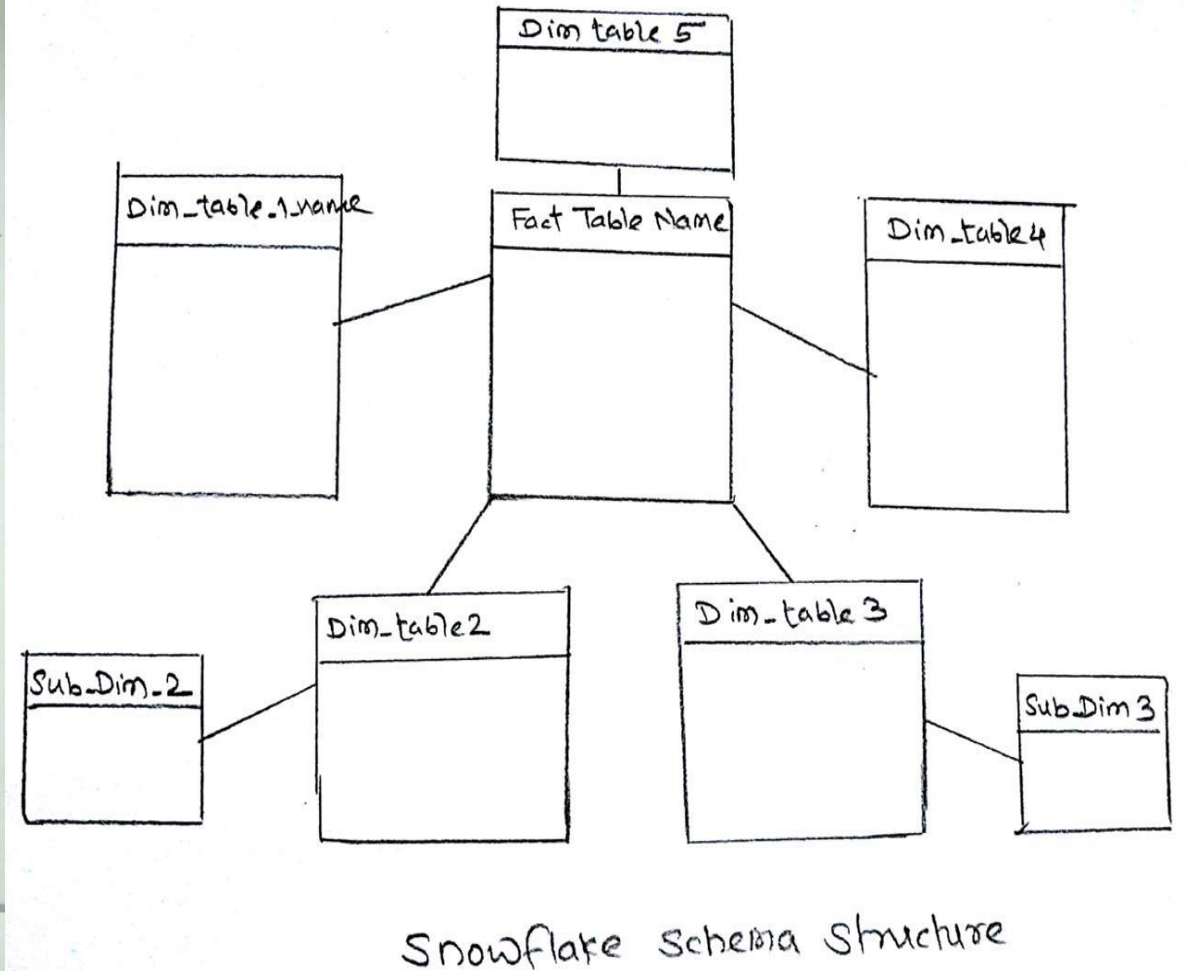
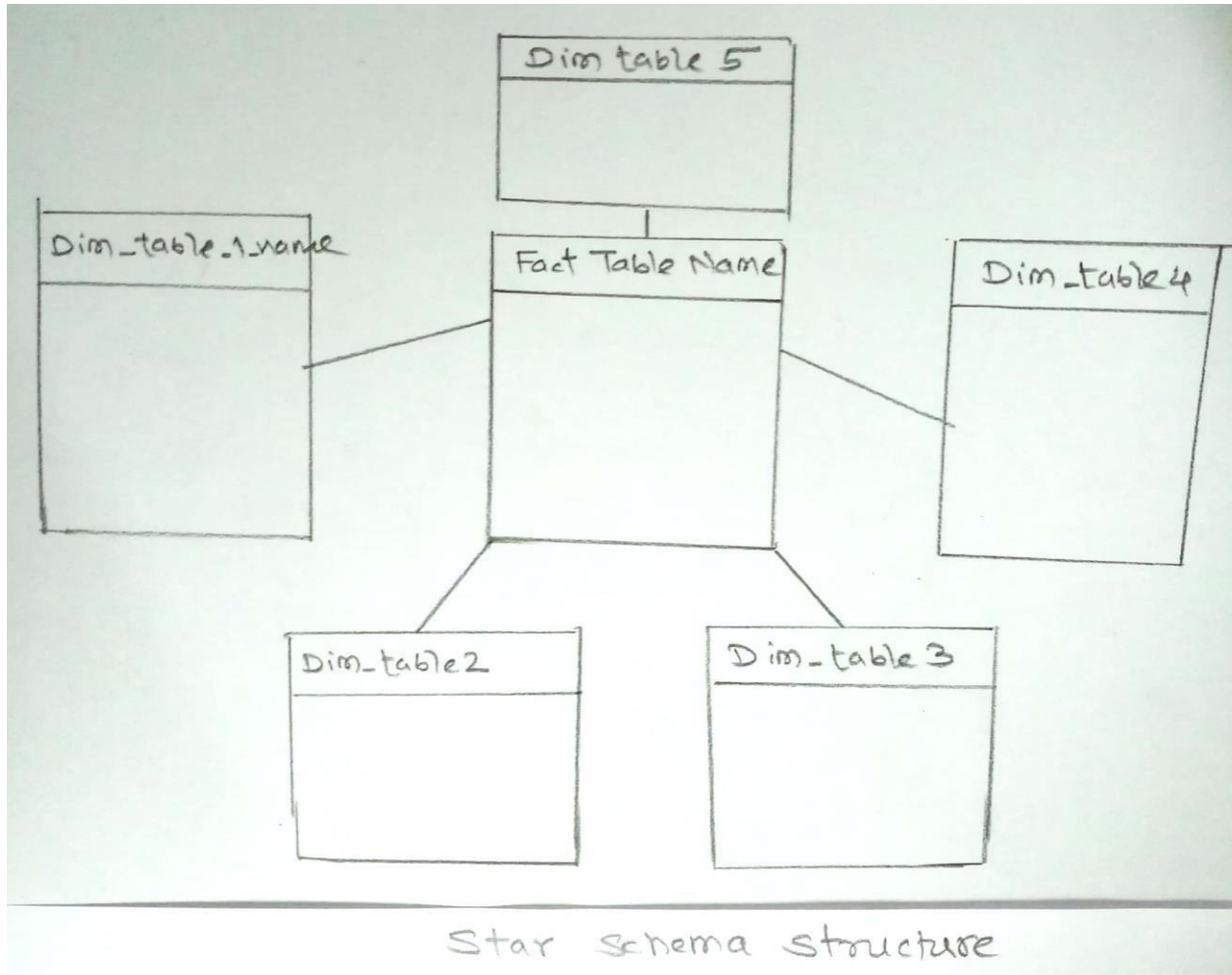


# **DATA WAREHOUSING AND MINING**

T.E. CSE-DS, Sem V  
Academic Year: 2022-23

Data Warehousing Fundamentals: Snowflake Schema, Aggregate Fact  
Table, Constellation Schema

# Star and Snowflakes Schema



# Aggregate tables

- Aggregate fact tables contain pre-calculated, summarized, derived data from the most detailed (granular) fact table
- They are mainly designed to reduce runtime processing
- Aggregate tables store data needed for multiple executions of the same query or execution of multiple similar queries using the same data subset
- They are good for repetitive actions

# Aggregate tables

## Need for building aggregate fact tables

✓ **Large size of the fact table** - retail Industry , Telecommunication industry

✓ **To speed up query execution**

eg: How did the four new store in Delhi performed during the last quarter compared to national average.

How do the 24 December sales by product categories compared to that of last year's?

These queries required selection and summation from the rows of the base fact tables.

To summarize the fact table's rows, detailed data based on one or more dimensions and summary totals based on other dimensions is needed.

# Limitations of Aggregate Fact Tables

- It must be re-aggregated every time there is a change in the source data to reflect those changes in the data warehouse .
- Aggregate data do not support exploratory analysis and hence have narrow applicability.

# Fact Constellation Schema –(Family of star)

- It is possible to construct fact constellation schema for each star or snowflake schema.
- Almost every organization has multiple star schema structure where each star serves specific purpose.
- More complicated design
- Family of star could be required for various purpose – family of starts could be generated by adding aggregate tables or for supporting value chain

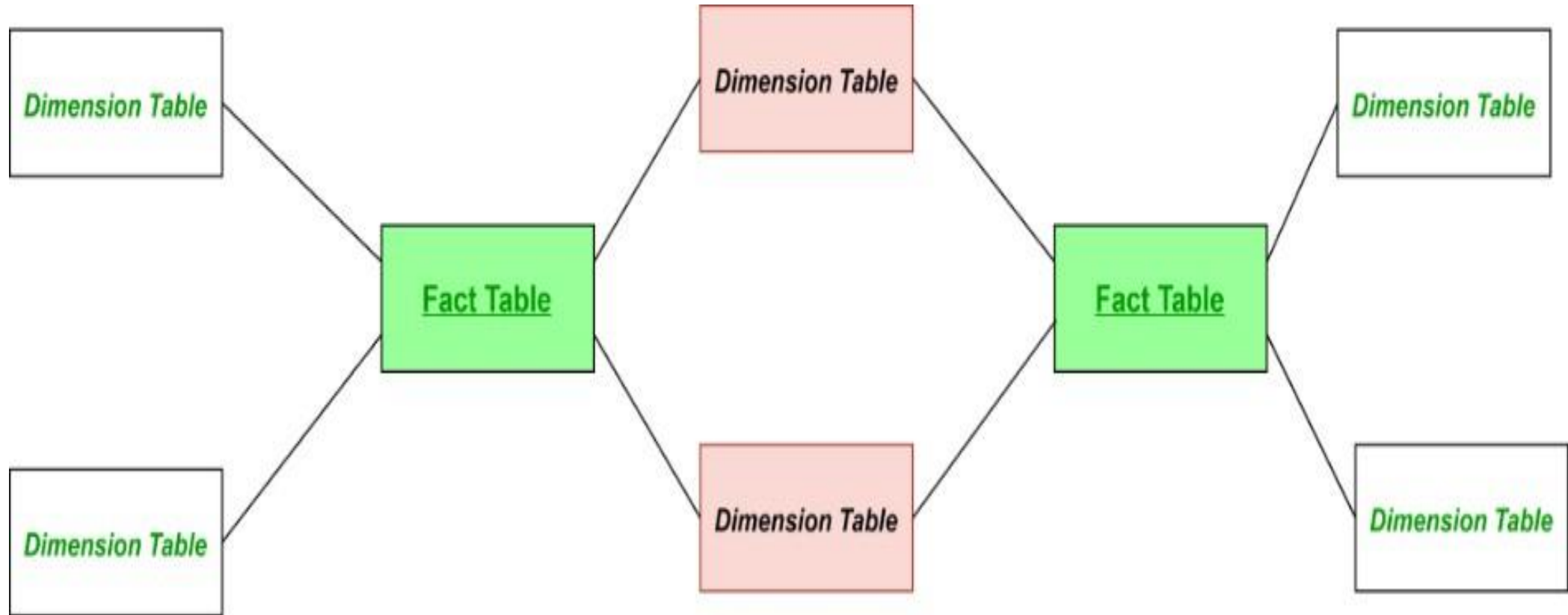
# Fact Constellation Schema

- Dimensions are shared among the fact tables.
- For this the attributes of the shared dimensions must have common meaning in every fact table.
- It is also needed to standardize facts – right/ same unit of measurement, same algorithm is used to derived data in each fact table.

Flaw:

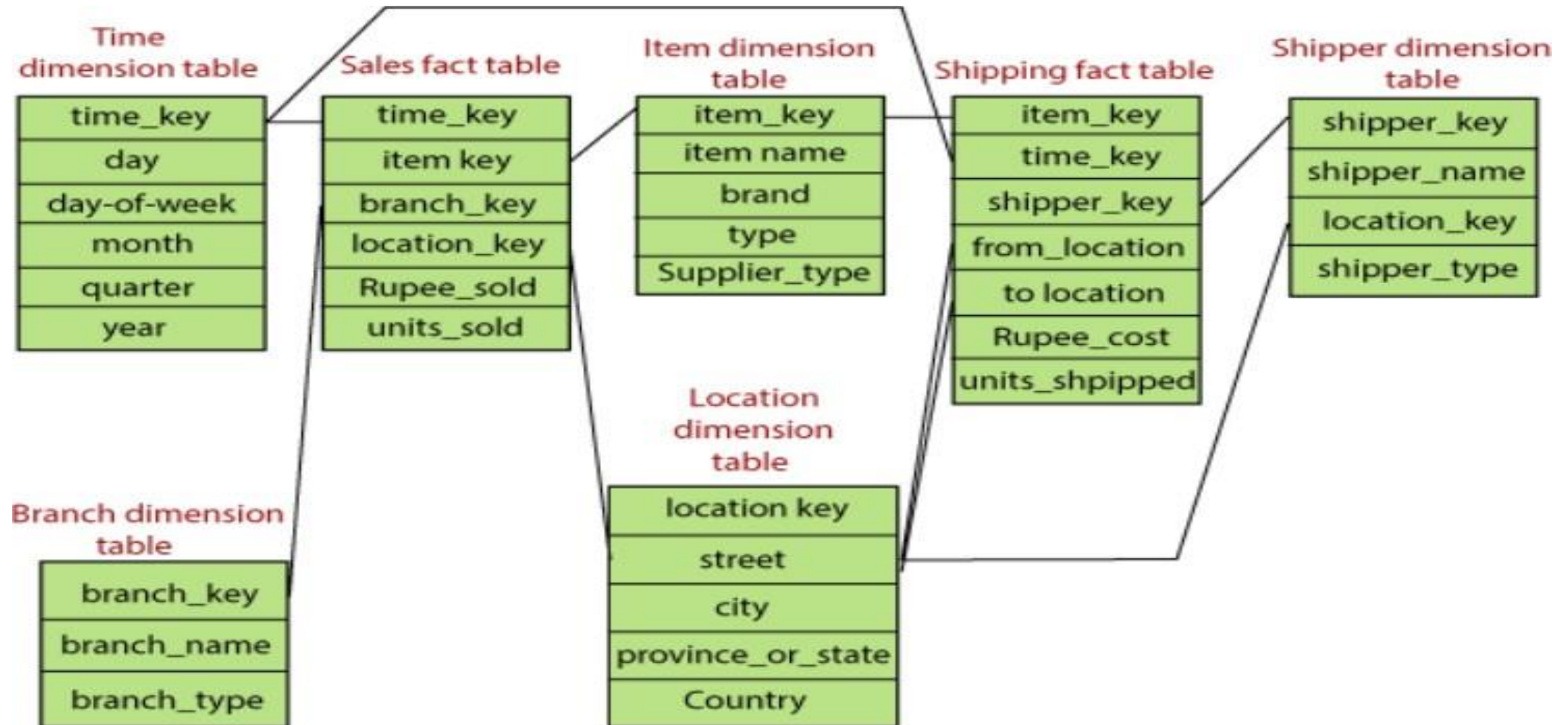
Too complicated compared to other two schems.

# Structure of Fact Constellation Schema





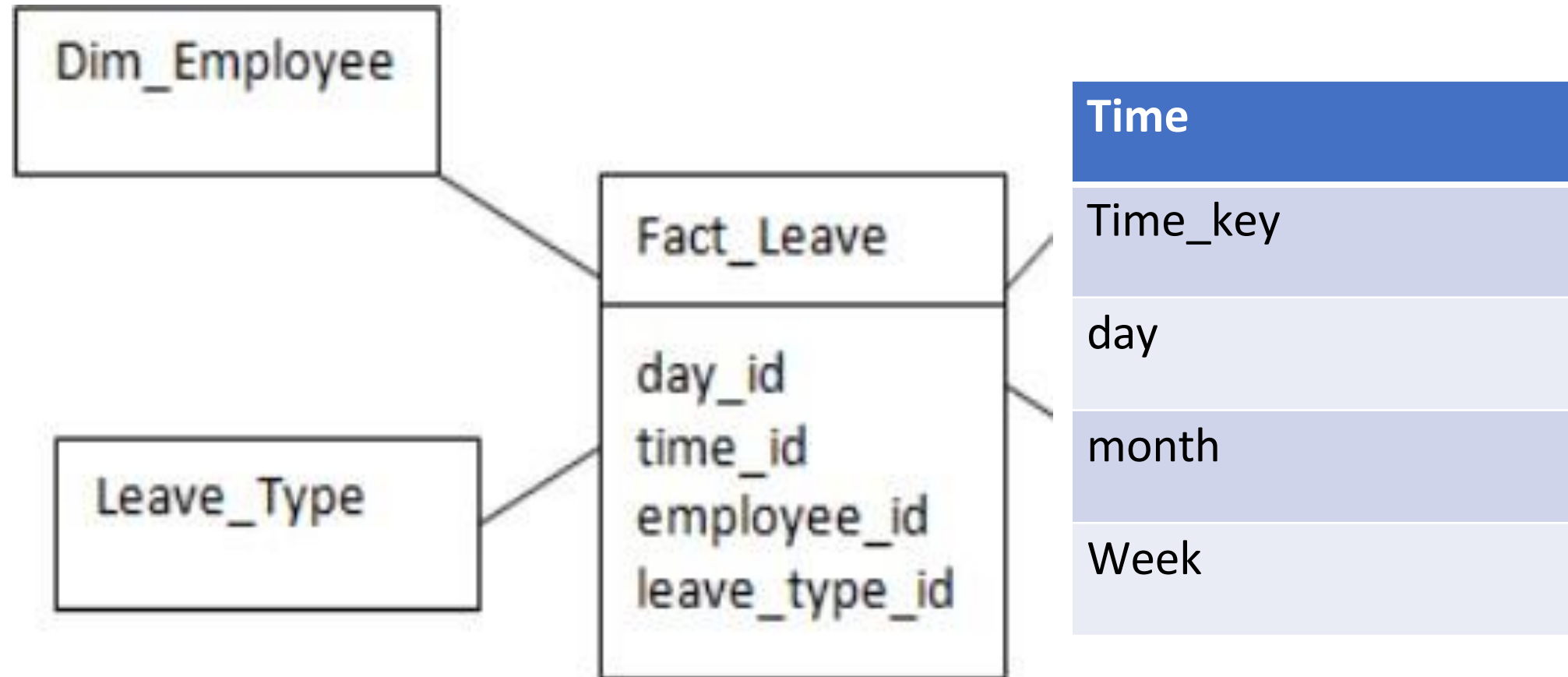
# Fact Constellation Schema



# Factless Fact Table

- Some fact tables do not have any facts at all and may consist of nothing but keys.
- These are called factless fact tables.
- Such tables are used to record events.
- Even though there are no facts stored in the fact table, the event can be counted to produce very meaningful process measurements.

# Factless Fact Table



- This fact table can help us identify events such as leaves consumed by the employees.
- ✓ How many leaves consumed by the employee
- ✓ Which types of leaves are consumed
- ✓ Which day/ month