

ASSIGNMENT-3

CASE STUDY

ADBMSL

Name: Prasad Sanjay Khalkar

Roll No: 33138

TE-09 L-09

Problem Statement: Design conceptual model using Star and Snowflake schema for any one database.

Database:

Employee Database

Fact Table:

A fact table is a primary table in a dimensional model.

A Fact Table contains:

- Measurements/facts
- Foreign key to dimension table

Dimension Table:

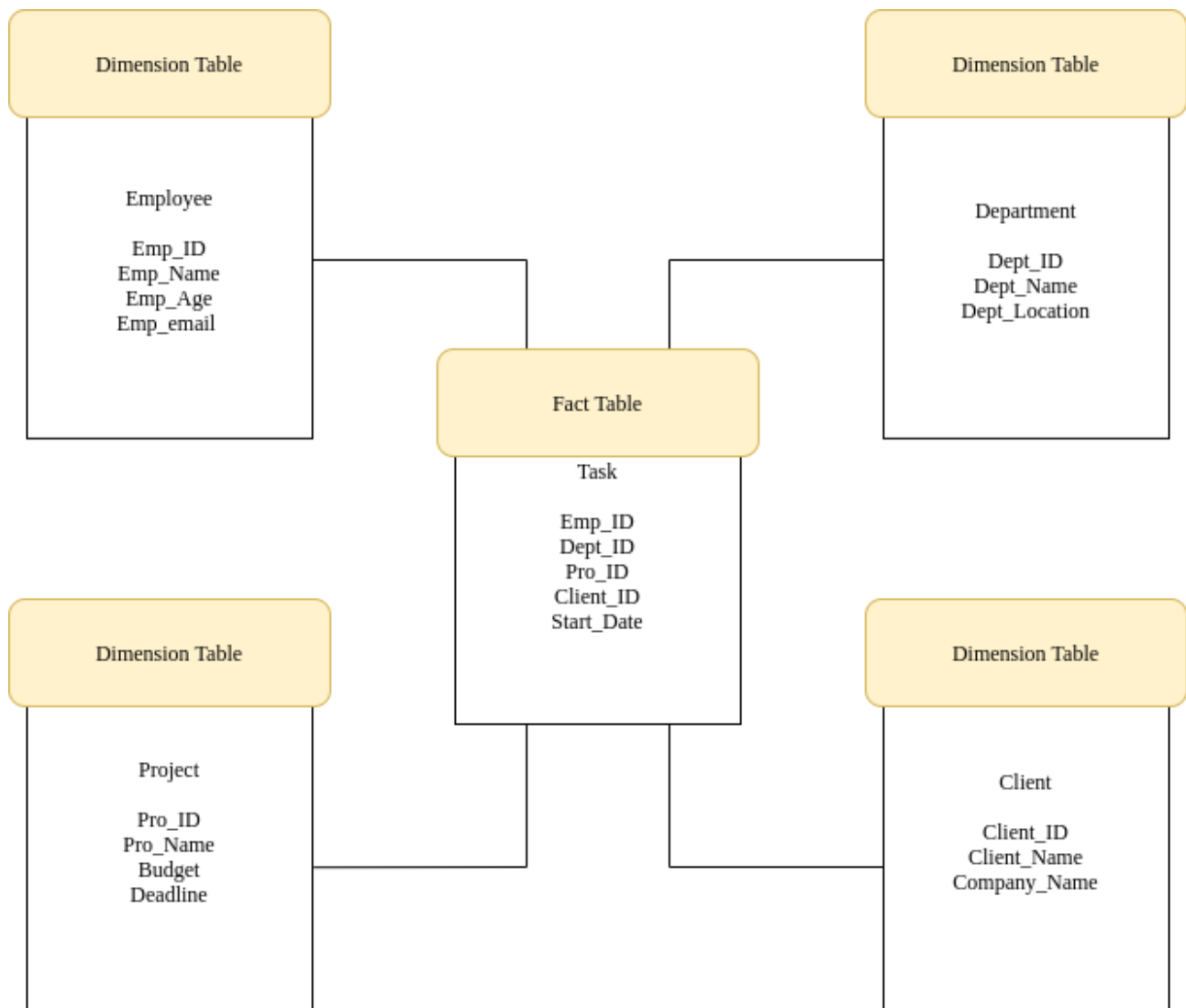
A dimension table contains dimensions of a fact.

- They are joined to fact table via a foreign key.
- Dimension tables are de-normalized tables.
- The Dimension Attributes are the various columns in a dimension table
- Dimensions offers descriptive characteristics of the facts with the help of their attributes
- No set limit set for given for number of dimensions
- The dimension can also contain one or more hierarchical relationships

Star Schema:

This schema is widely used to develop or build a data warehouse and dimensional data marts. It includes one or more fact tables indexing any number of dimensional tables. The star schema is a necessary cause of the snowflake schema. It is also efficient for handling basic queries.

It is said to be star as its physical model resembles to the star shape having a fact table at its center and the dimension tables at its peripheral representing the star's points



In this example, the fact table is Task i.e the task or projects allocated to each employee and to which client these projects belong and to which department this project and the employee belongs

Employee Dimension Table consists of ID, Name, Age, email

Department Dimension Table consists of ID, Name, Location

Project Dimension Table consists of ID, Name, Budget, Deadline

Client Dimension Table consists of ID, Name, CompanyName

The Task Fact table consists of the Emp_ID, Dept_ID, Project_ID, Client_ID and startDate.

Snowflake Schema:

The snowflake schema is a variant of the star schema. Here, the centralized fact table is connected to multiple dimensions. In the snowflake schema, dimensions are present in a normalized form in multiple related tables. The snowflake structure materialized when the dimensions of a star schema are detailed and highly structured, having several levels of relationship, and the child tables have multiple parent tables. The snowflake effect affects only the dimension tables and does not affect the fact tables.

Example:

The same example as of Star Schema is taken to see how Snowflake Schema works if more and more attributes are added to the Dimension Tables.

The Task Fact table is same as of the Star Schema.

The Employee Dimension Table consists of a sub-dimensional table as Address i.e. the address of the employee(city, street, state, country, pincode, etc)

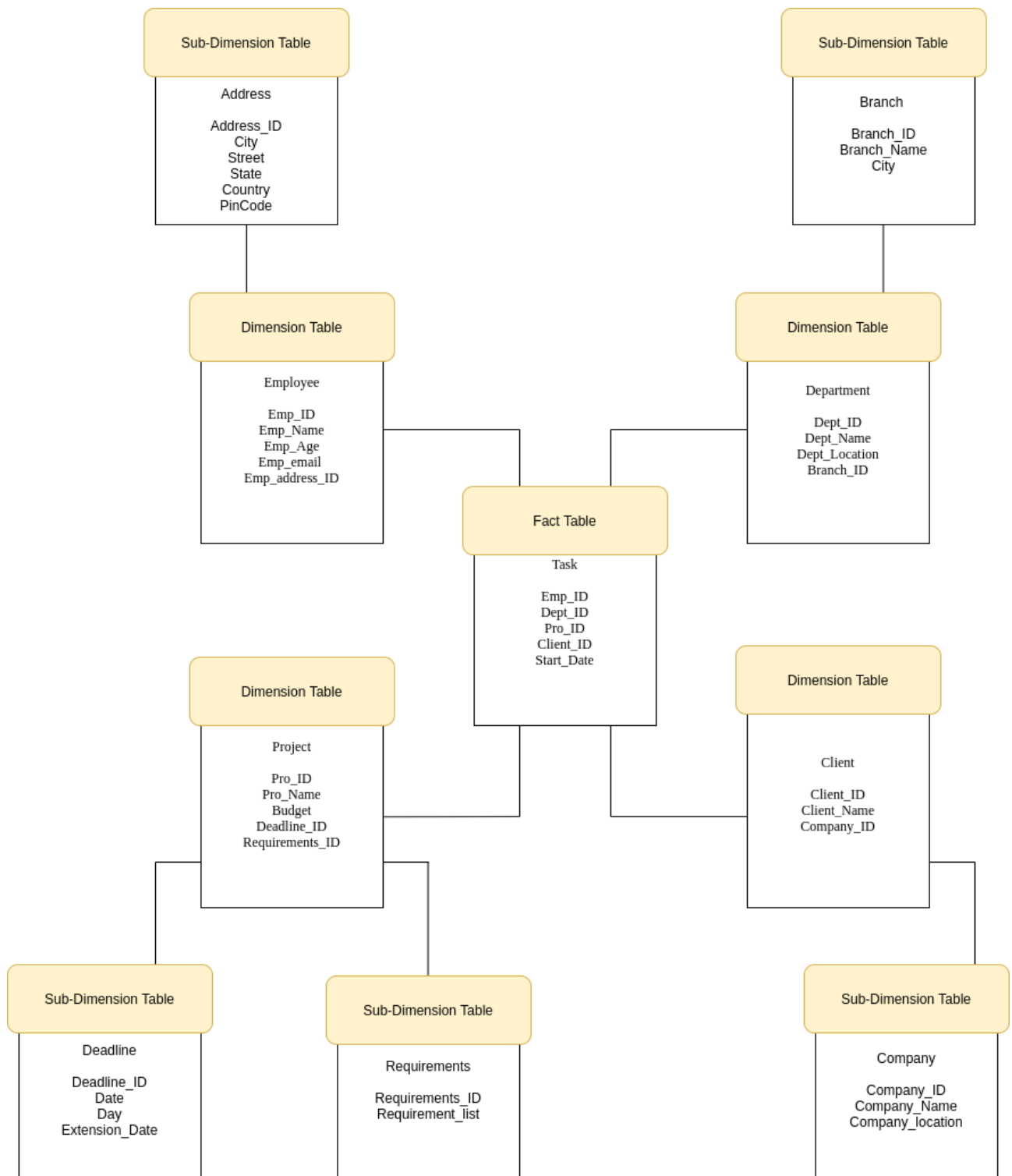
The Department Dimension Table consist of a sub-dimensional table as Branch i.e. the BranchName, City, etc.

The Client Dimension Table consist of a sub-dimensional table as Company i.e. CompanyName, Location, etc

The Project Dimension Table consist of 2 sub-dimensional table as follows:

- 1) Deadline, which consists of date, day of the deadline and the extensionDate if any
- 2) Requirements, which consists of list of requirements provided by the client for this project

This shows how Snowflake Schema normalises the denormalised data in Star Schema when more complex, multi dimensional attributes are added to the tables, and how it speeds up the processes.



Comparison Between Star and Snowflake Schema:

S.NO	Star Schema	Snowflake Schema
1.	Star schema is a top-down model.	While it is a bottom-up model.
2.	Star schema uses more space.	While it uses less space.
3.	In star schema, Normalization is not used.	While in this, Both normalization and denormalization are used.
4.	It's design is very simple.	While it's design is complex.
5.	Queries performance is good as there is less number of joins involved.	The performance of queries is a bit less when compared to star schema as more number of joins are involved.
6.	It takes more disk space.	It takes less disk space.
7.	It has less number of foreign keys.	While it has more number of foreign keys.
8.	It has high data redundancy.	While it has low data redundancy.

Conclusion:

Studied and understood Star and Snowflake Schema and understood the limitations and advantages of both, using an example of Star and Snowflake Schema on conceptual schema model.