

Experiment no.:- 1

Aim:- One Case study on building Data warehouse/Data Mart

Objectives:-

1. Case study on building Data warehouse/Data Mart.
2. Design dimensional modeling (star and snowflakes schemas) using online tools.

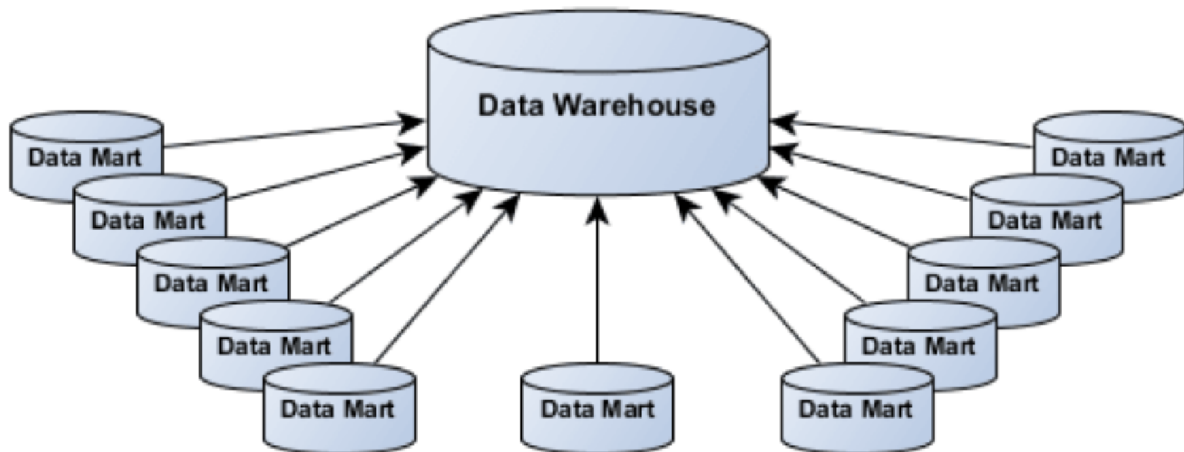
Theory:-

1. CASE STUDY ON DATA WAREHOUSE/DATA MART

DATA MART

A Data Mart is a smaller version of a data warehouse and it is meant to be used by a particular department or a group of individuals in the company.

It focuses on a single functional unit of an organization and keeps a subset of data stored in the data warehouse. It is normally controlled by a unit department in the organization. Whereas a data warehouse draws data from many sources, a Data Mart draws data from only a few sources.



TYPES OF DATA MARTS

There are 3 types of Data Marts which vary depending on their relation to the data warehouse and the data sources used to build them.

They include the following:

- **Dependent:**

These are the types of Data Marts that are created from an existing company data warehouse. It's a top-down approach that starts with storing all the company data in a single central location, then extracting a portion of data when needed for analysis. A dependent Data Mart can be a logical view or a physical view of a data warehouse:

- **Logical view:** This is a virtual table or view that is logically, rather than physically, separated from the warehouse.
- **Physical subset:** This is a data extract stored in a database that is physically separated from the data warehouse.

- **Independent:**

An independent Data Mart is created without the use of a data warehouse, meaning that it's a standalone system that focuses on one business function or subject area.

The data is extracted from internal/external data sources, processed, and loaded to the data repository where it is stored for analytics. They are easy to design and develop and they help organizations to achieve their short-term goals.

- **Hybrid:**

These are the storage structures that combine data from a data warehouse and other source systems. A hybrid Data Mart combines the speed and end-user focus of the top-down approach with the advantages of the organization-level integration of a bottom-up approach.

DATA MART vs DATA WAREHOUSE

1. Both Data Warehouses and Data Marts are crucial central data repositories, but they serve different needs within an organization.
2. A Data Warehouse is known as a system that accumulates data from multiple disparate sources into a single, central, consistent data store to support Artificial Intelligence (AI), Data Mining, and Machine Learning. This ultimately leads to the enhancement of Business Intelligence and sophisticated analytics. With the help of this strategic collection process, Data Warehouse solutions can collect data from varied sources to make it available in a single unified format.
3. A Data Mart is defined as a more focused version of a Data Warehouse. It contains a smaller subset of data integral to and needed by a single team or a select group of users within an organization. It is usually built from an existing Data Warehouse through a complex procedure that leverages various tools and technologies to construct and design a physical database, populate it with data and set up sophisticated management and access protocols.

STRUCTURE OF STORED DATA

Data Marts normally store transactional data in rows and columns, making it easy to access, organize, and understand data. Since they store historical data, they make it easy for data analysts to understand data trends.

They use the following types of schema:

Star Schema

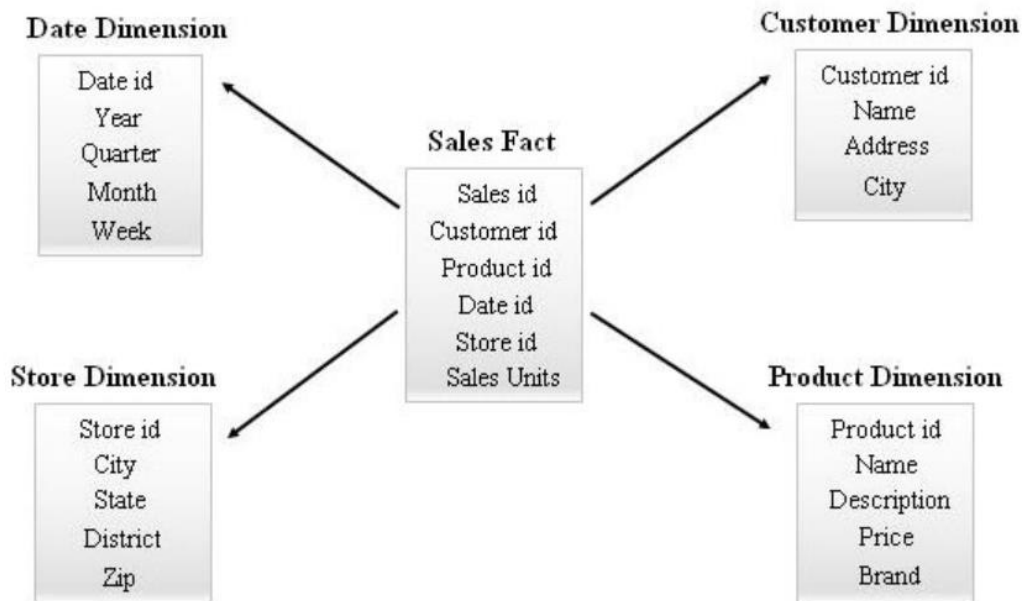
Snowflake Schema

DESIGN DIMENSIONAL MODELING (STAR AND SNOWFLAKES SCHEMAS) USING ONLINE TOOLS.

a. STAR SCHEMA

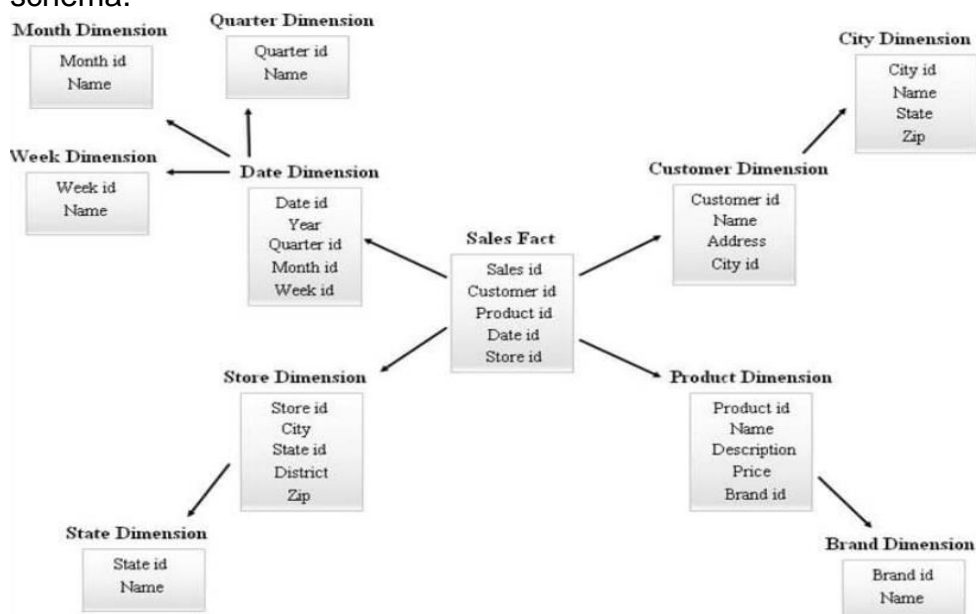
- In a star schema, the fact table will be at the center and is connected to the dimension tables.

- The tables are completely in a denormalized structure.
- SQL queries performance is good as there is less number of joins involved.
- Data redundancy is high and occupies more disk space.



b. SNOWFLAKE SCHEMA

- A snowflake schema is an extension of a star schema where the dimension tables are connected to one or more dimensions.
- The tables are partially denormalized in structure.
- The performance of SQL queries is a bit less when compared to star schema as more joins are involved.
- Data redundancy is low and occupies less disk space when compared to star schema.

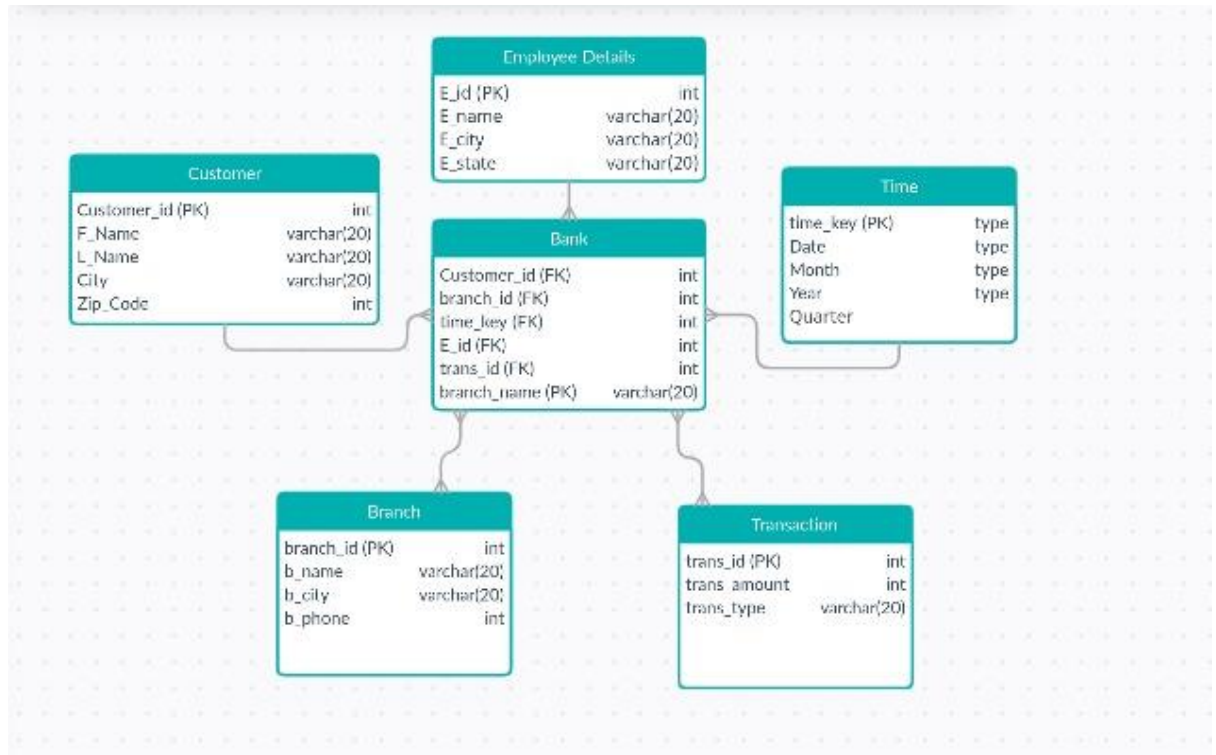


PROCEDURE:

Using some online tools and editor like <https://creatly.com/> prepare star schema diagram and snowflakes schema diagram for BANK MANAGEMENT SYSTEM.

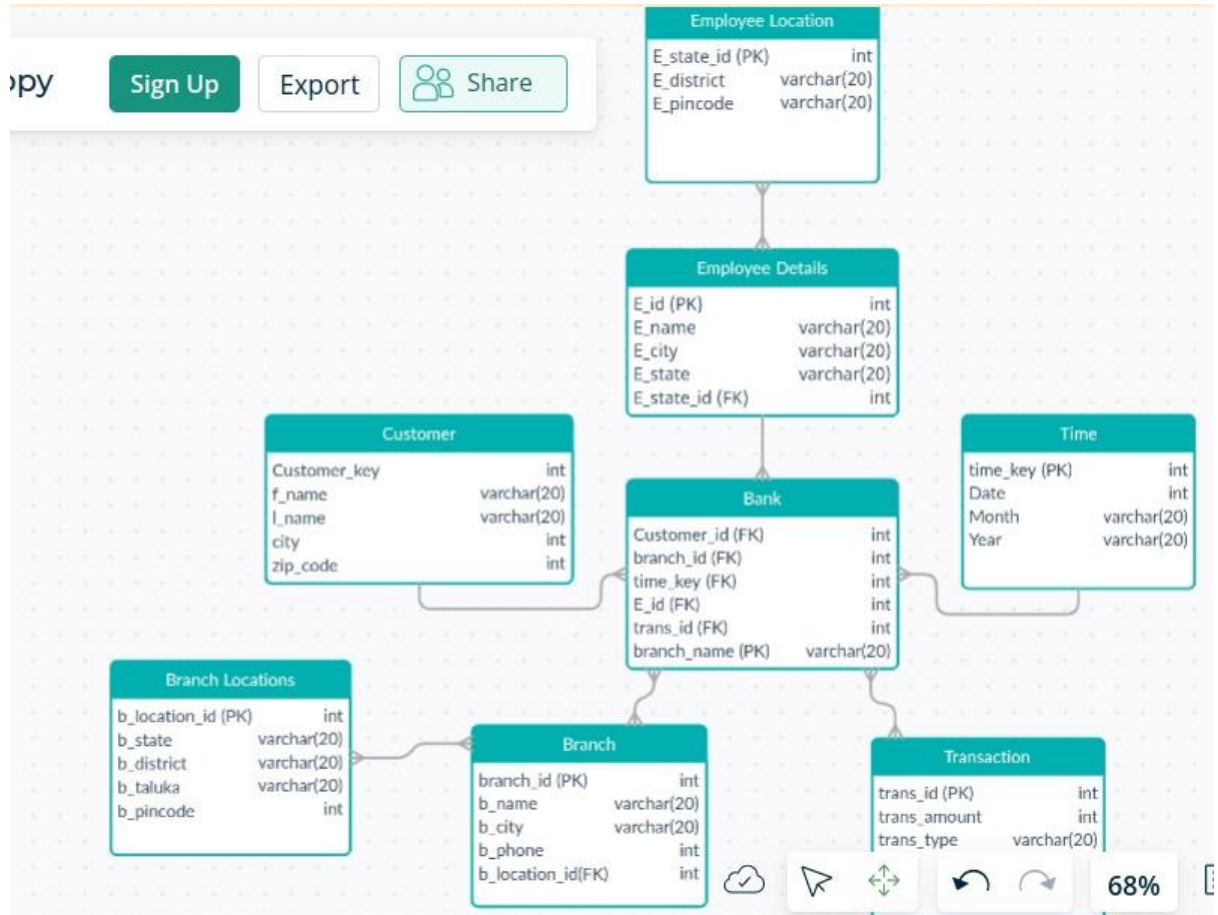
1. STAR

SCHEMA



2. SNOWFLAKES

SCHEMA



Conclusion :

1. On the basis of the research we had successfully made a case study on Data Warehouse/Data Mart.
2. Studied about Star Schema and Snowflakes schema and successfully designed the diagrams.