

Name: Shazmeen Shaikh

Roll No: 49

Date: 17/01/2024

Experiment No. : 1

Title: Tutorial-1 Design Star and Snowflake schema.

Problem Definition: Consider a Hospital with 3 dimensions – Doctor, Patient and Time, and 2 measures, Count and Charge. Draw Star and Snowflake schema for creating Hospital DW.

Theory:

STAR SCHEMA:

The Star Schema is a data modeling technique commonly used in data warehousing where data is organized into a central fact table surrounded by denormalized dimension tables, resembling a star shape. This design simplifies querying and enables faster data retrieval, making it suitable for simpler analytical tasks. The fact table contains numerical data, such as sales figures or quantities, while the dimension tables provide context for the data, such as time, location, and product information. By storing data in this format, the Star Schema optimizes query performance by minimizing the number of table joins required to answer analytical questions, thus facilitating efficient data analysis and reporting.

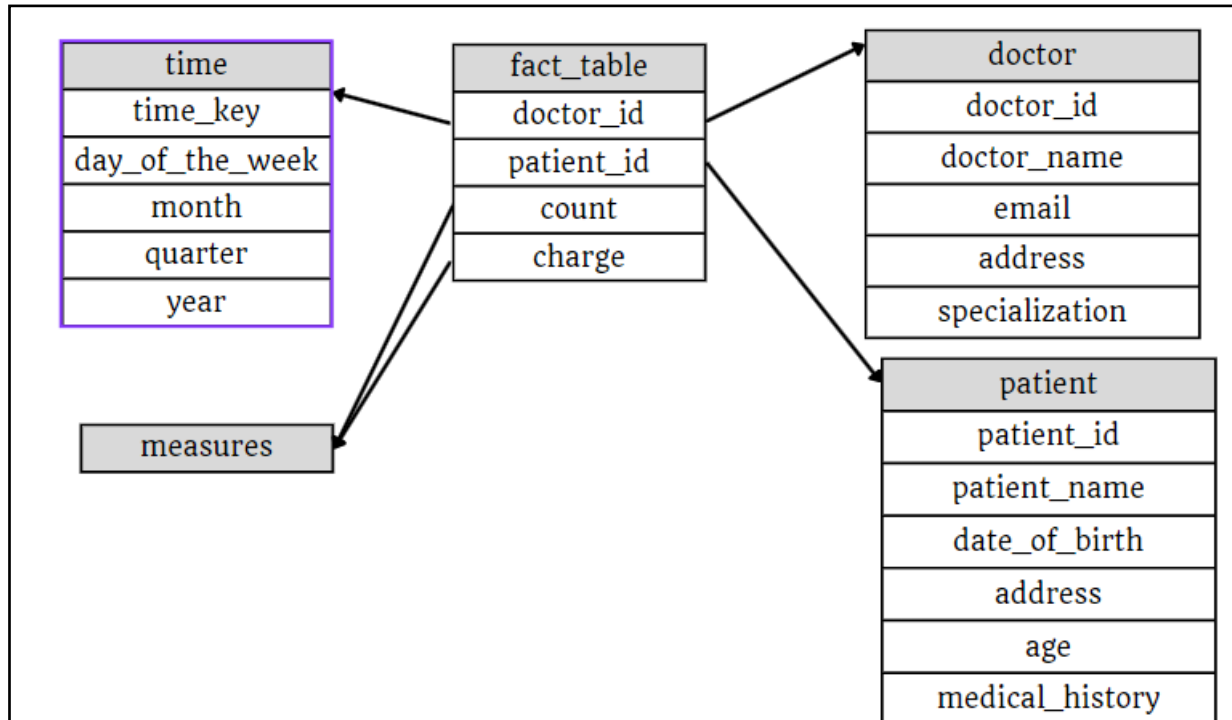
SNOWFLAKES SCHEMA:

The Snowflake Schema is an extension of the Star Schema, where dimension tables are further normalized, resulting in a more complex structure resembling a snowflake. In this schema, dimension tables are broken down into multiple related tables, reducing redundancy and improving data integrity. While this normalization conserves storage space and enhances data consistency, it can complicate query execution due to the increased number of table joins required to retrieve data. Despite its complexity, the Snowflake Schema offers advantages in scenarios where data integrity is critical and storage efficiency is paramount. Organizations must carefully evaluate their specific requirements and trade-offs between query performance and data integrity when choosing between the Star and Snowflake schemas for their data warehousing needs.

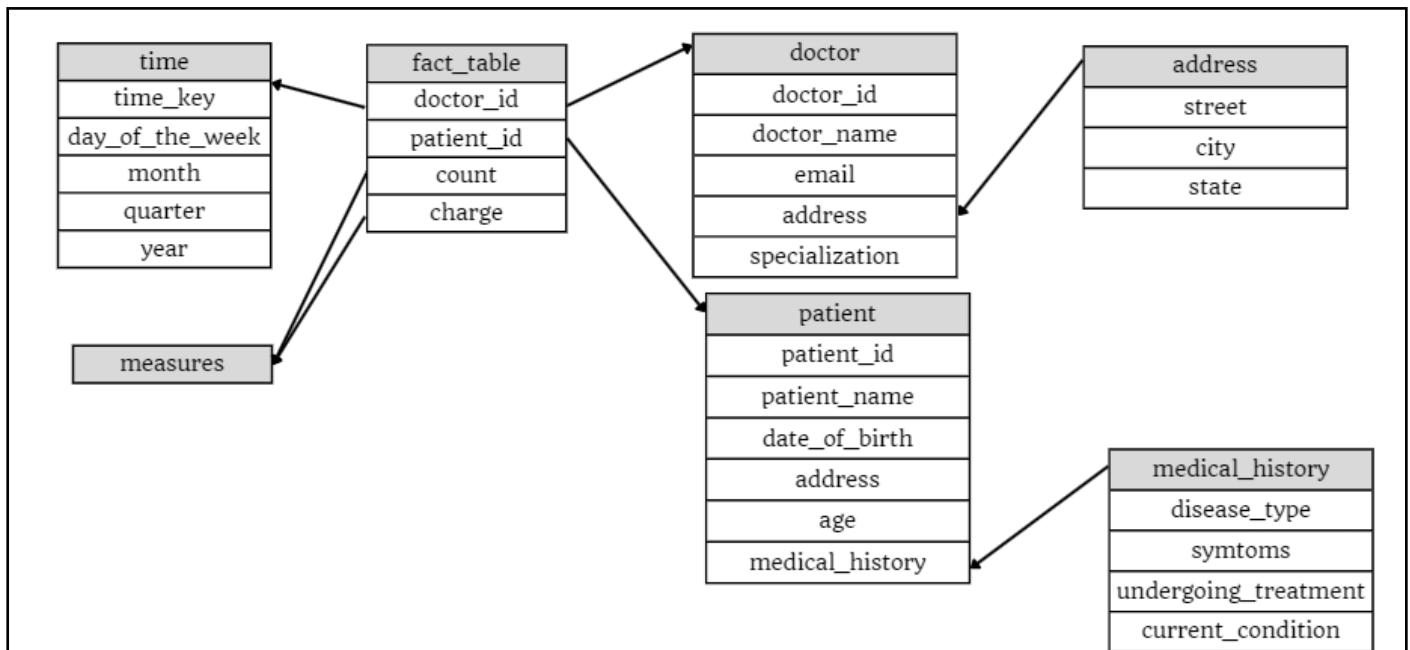
COMPARISON:

Feature	Snowflake Schema	Star Schema
Structure	Normalized structure resembling a snowflake	Denormalized structure resembling a star
Dimension Tables	Dimension tables are further normalized	Dimension tables are denormalized
Table Joins	Requires more table joins due to normalization	Requires fewer table joins for query execution
Data Integrity	Enhanced data integrity due to normalization	May sacrifice some data integrity for performance
Storage Efficiency	Optimized storage space due to normalization	May require more storage space due to denormalization
Query Complexity	Increased query complexity due to additional joins	Reduced query complexity due to fewer joins
Performance	Potentially slower query performance due to additional joins	Generally faster query performance due to fewer joins
Suitable For	Complex data structures where data integrity is paramount	Simplified analytical tasks with emphasis on query performance

Results:



STAR SCHEMA



SNOWFLAKES SCHEMA

References :

<https://www.geeksforgeeks.org/star-schema-in-data-warehouse-modeling/>
<https://www.geeksforgeeks.org/snowflake-schema-in-data-warehouse-model/>