NAME: Vishal Shashikant Salvi   **UID: 2019230069**

DATE:04/08/2020   **BATCH:C**

Aim:

One case study given to a group of 3 /4 students, of a data mart/ data warehouse.

a. Write Detail Statement Problem and creation of dimensional modeling (creation star and snowflake schema)

b. Implementation of all dimension table and fact table

Theory:

Problem Statement:

Student Record Management System:

Student records are often viewed as paperwork produced for the education bureaucracy. However, a well-designed student record system, whether using paper documents or automated systems, yields many benefits.

A second benefit, particularly with automated systems, is efficiency in processing and exchanging student records among departments. When student records are added to an overall management information system that includes information on staff, materials, more management activities can be accomplished and efficiency will be improved.

1. File Loss:

　　When computerized system is not implemented file might get lost beacuse of human environment.

　　Due to some human error there may be a loss of data.

2. When Student Records are maintained in a file or books it is possible that it might get damaged because of

　water. Beside some natural disasters like flood or fire may also damage the files.

3. Difficult to search the record:

　　Where there is no computerized system there is always a difficulty to search for a specific record if large number of records are maintained.

4. It is also very space consuming.

5. If there is no Computerized system then adding each and every record will require papers which will increase the

cost of maintaining the records.

6. Records maintained in files or diary are not easily editable. We have to put lots of efforts to edit even a single

record.

7. Using manual system to manage the students records in the book or paper is very hectic job for staff members.
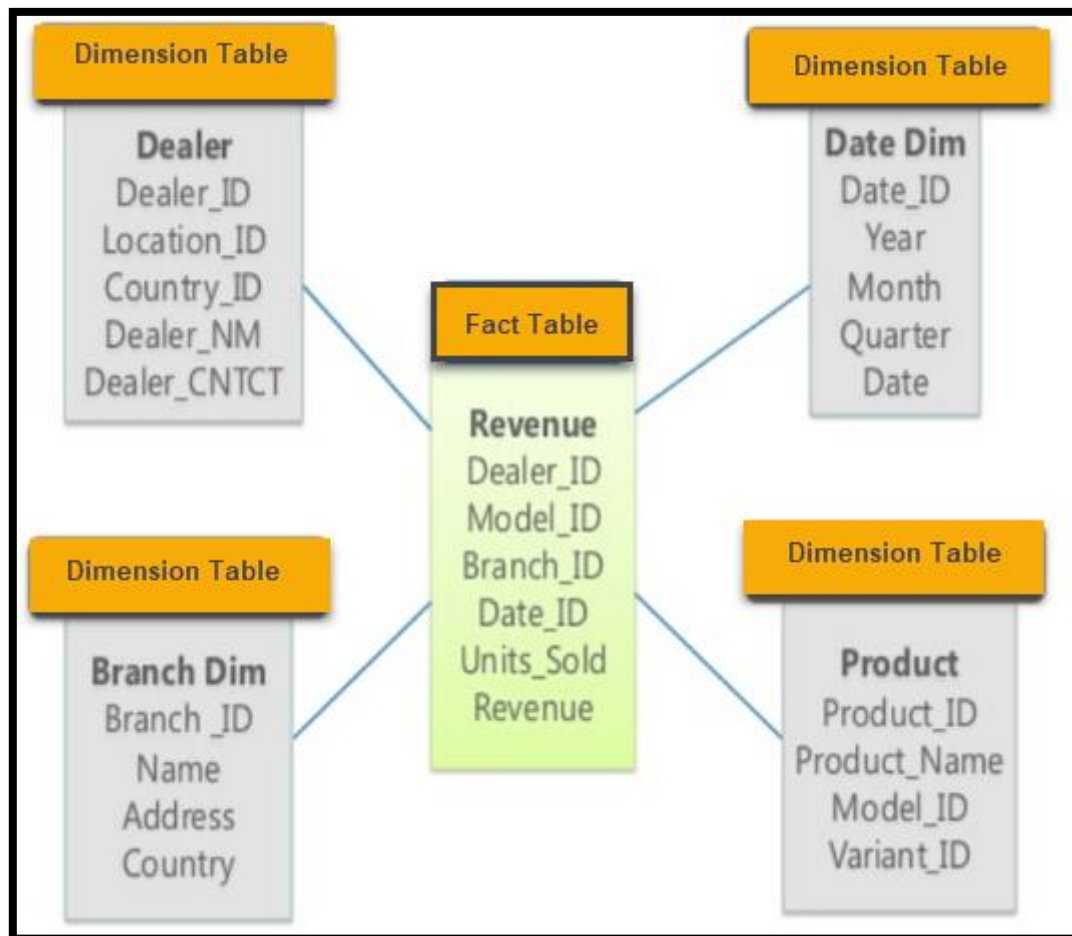

Star Schema:

In data warehousing and business intelligence (BI), a star schema is the simplest form of a dimensional model, in which data is organized into facts and dimensions. A fact is an event that is counted or measured, such as a sale or login. A dimension contains reference information about the fact, such as date, product, or customer. A star schema is diagramed by surrounding each fact with its associated dimensions. The resulting diagram resembles a star. Star schemas are optimized for querying large data sets and are used in data warehouses and data marts to support OLAP cubes, business intelligence and analytic applications

Characteristics of Star Schema:

> ➢ Every dimension in a star schema is represented with the only one-dimension table.
> ➢ The dimension table should contain the set of attributes.
> ➢ The dimension table is joined to the fact table using a foreign key
> ➢ The dimension table are not joined to each other
> ➢ Fact table would contain key and measure
> ➢ The Star schema is easy to understand and provides optimal disk usage.
> ➢ The dimension tables are not normalized. For instance, in the above figure, Country_ID does not have Country lookup table as an OLTP design would have.
> ➢ The schema is widely supported by BI Tools

In the **STAR Schema**, the center of the star can have one fact table and a number of associated dimension tables. It is known as star schema as its structure resembles a star. The star schema is the simplest type of Data Warehouse schema. It is also known as Star Join Schema and is optimized for querying large data sets.

In the following example, the fact table is at the center which contains keys to every dimension table like Dealer_ID, Model ID, Date_ID, Product_ID, Branch_ID & other attributes like Units sold and revenue.
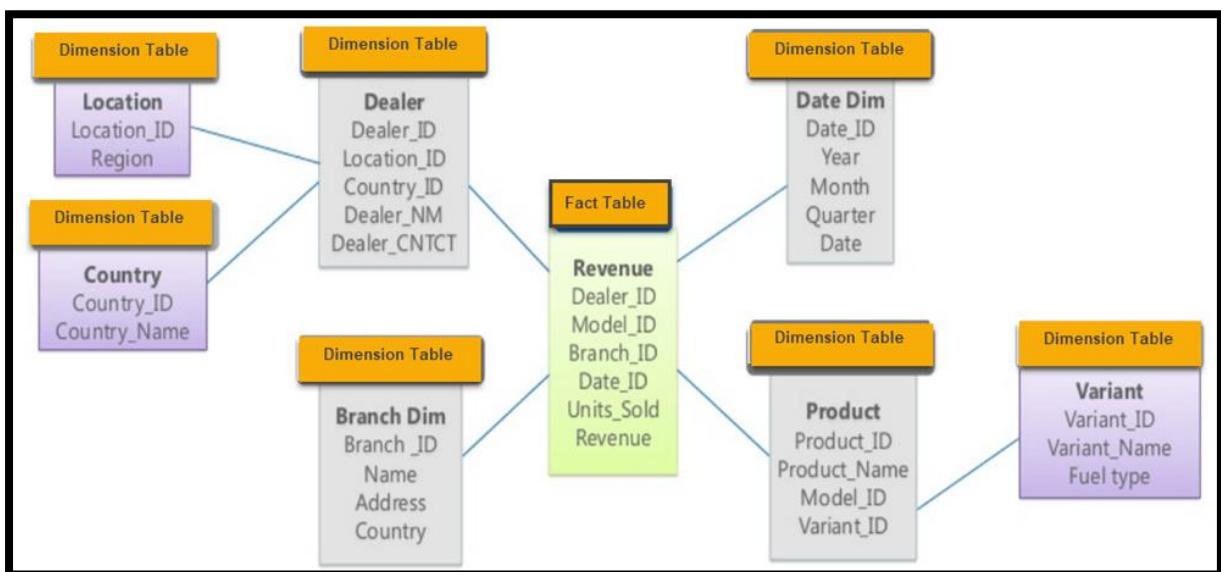

Example of Star Schema

Snowflake Schema:

In computing, a snowflake schema is a logical arrangement of tables in a multidimensional database such that the entity relationship diagram resembles a snowflake shape. The snowflake schema is represented by centralized fact tables which are connected to multiple dimensions. [citation needed]. "Snowflaking" is a method of normalizing the dimension tables in a star schema. When it is completely normalized along all the dimension tables, the resultant structure resembles a snowflake with the fact table in the middle.

Characteristics of Snowflake Schema:

- ➢ The main benefit of the snowflake schema it uses smaller disk space.
- ➢ Easier to implement a dimension is added to the Schema
- ➢ Due to multiple tables query performance is reduced
- ➢ The primary challenge that you will face while using the snowflake Schema is that you need to perform more maintenance efforts because of the more lookup tables.

SNOWFLAKE SCHEMA is a logical arrangement of tables in a multidimensional database such that the ER diagram resembles a snowflake shape. A Snowflake Schema is an extension of a Star Schema, and it adds additional dimensions. The dimension tables are **normalized** which splits data into additional tables.
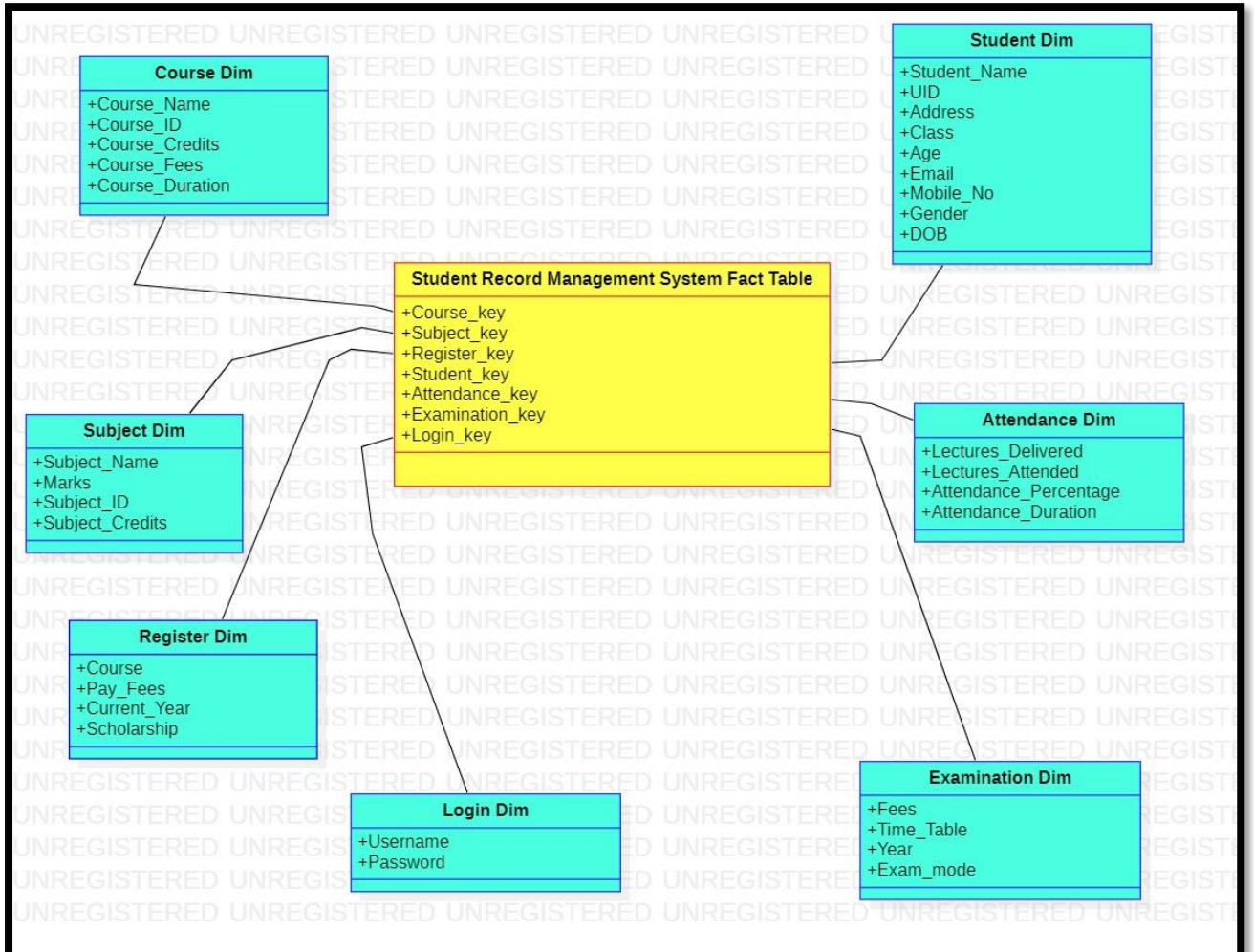
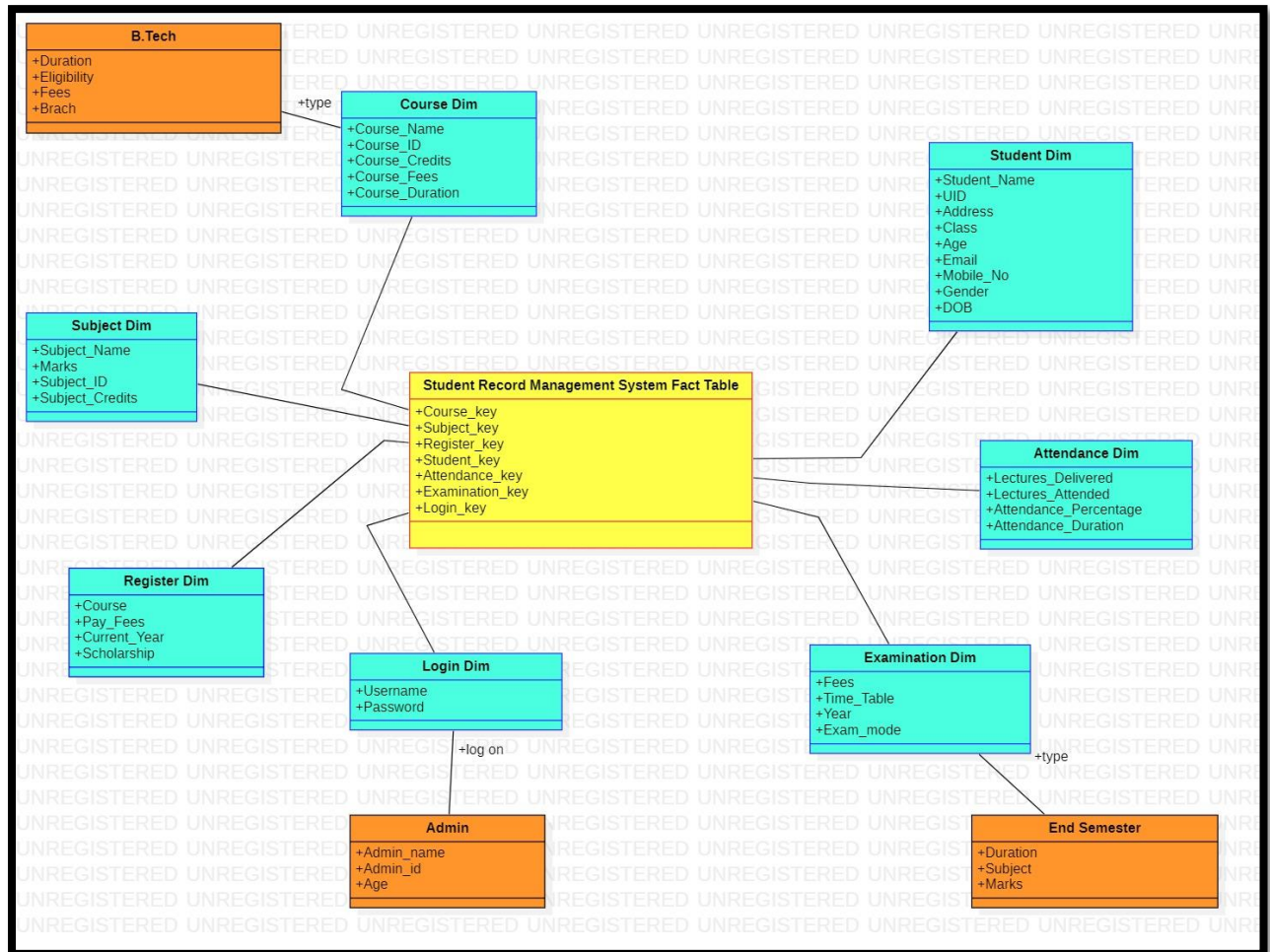In the following example, Country is further normalized into an individual table.

Example of Snowflake Schema

Implementation of all dimension table and fact table:

Star Schema:

Snowflake Schema:

Conclusion:

So, Multidimensional schema is especially designed to model data warehouse systems. We study that the Star Schema is the simplest form of a dimensional model used in data warehousing whereas the snowflake schema uses smaller disk space. Also we implement all dimension table as well as fact table.