**OLTP vs OLAP:**

|  |  |
| --- | --- |
| OLTP | OLAP |
| Recourses used:  Gets data from more than source such as  **Database MS** in Normalized Form for Query optimization and fast response  **Cloud** | **Data wareh**ouse in Denormalized Form for analysis |
| Purpose:  It has to be Optimized for data large number of daily transactions, Enable multi-user access to the same data | Quickly query, report on and analyze multidimensional data |
| Usage :  Is used for daily transaction at the system such as ATM Machines and processing database | Is used by data analyist and data engineers to analysis data |

**SLICING & DICING**

In DWH, A cube in multidimensional data model to store transformed, summeriazed and aggregated data , to be used in OLAP tools.

Let’s start by

**Slicing** it is used to select a slice of one of cube dimensions to produce a new cube

**Dicing:** it selects multiple dimension of data cube to produce another cube.

**What are data marts?**

Data mart is a part of data Ware house or simple form of DWH, is used for analytical propose if we have a large data warehouse and we want to focus in particular subject ,makes it easy to access the data without spending much time in searching in complex DWH.

**Difference between snowflake schema, star schema, galaxy schema**

Let’s start by defining Schema:

Schema shows Architecture of how the data will be stored and organized how it is relates to other data in DWH

DWH schema constructs of Fact Table and Dimension Table

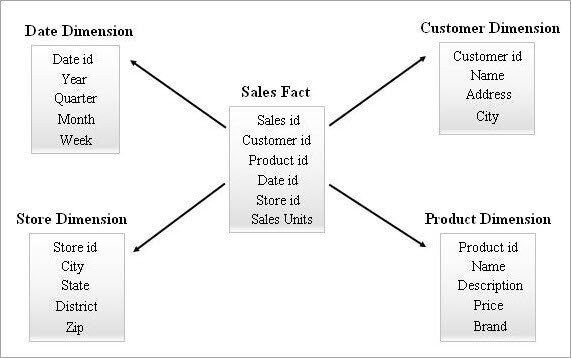
* Fact Table: contains measurements and metrics about the business and stores the primary key of dimensional table to show how the data relates in DWH
* Dimensional Table: are denormalized tables used to store data Attributes and dimension , dimension tables doesn’t join together.

Let’s start by simplest one of them.

**Start schema**

It consist of one Fact Table at the center with dimensional tables in denormalized form which enables quick query response.

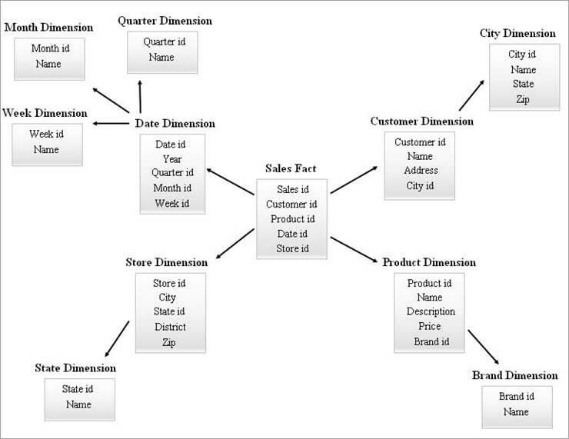
Dimensional tables are not connected directly, the fact table has One-to-many Relationships with all dimensional tables thus the navigation between table and retrieving data become easy.



**Snowflake Schema**

It designed by using start schema as an input, this done by normalizing dimensional table from start schema thus the query performance increased and data redundancy gets removed .

Each row of the fact table relates with several rows in the dimension tables using its foreign key reference.



**Galaxy Schema**

It contains more than fact tables are able to share dimensional tables between them.

It is used for propose of getting sophisticated requirements for the aggregated fact tables

Because of higher level of complexity in this model makes it more difficult to maintain than the others.

