OBIEE

ORACLE

tutorialspoint

SIMPLYEASYLEARNING

www.tutorialspoint.com

https://www.facebook.com/tutorialspointindia

https://twitter.com/tutorialspoint

About the Tutorial

Oracle Business Intelligence Enterprise Edition (OBIEE) is a Business Intelligence (BI) tool by Oracle Corporation. Its proven architecture and common infrastructure producing and delivering enterprise reports, scorecards, dashboards, ad-hoc analysis, and OLAP analysis provides a rich end-user experience. This tutorial explains all the fundamental aspects of OBIEE.

Audience

This tutorial is designed for those who want to learn the basics of OBIEE and take advantage of its features to develop quality BI reports.

Prerequisites

Before proceeding with this tutorial, you need to have a good understanding of basic database concepts.

Disclaimer & Copyright

© Copyright 2018 by Tutorials Point (I) Pvt. Ltd.

All the content and graphics published in this e-book are the property of Tutorials Point (I) Pvt. Ltd. The user of this e-book is prohibited to reuse, retain, copy, distribute or republish any contents or a part of contents of this e-book in any manner without written consent of the publisher.

We strive to update the contents of our website and tutorials as timely and as precisely as possible, however, the contents may contain inaccuracies or errors. Tutorials Point (I) Pvt. Ltd. provides no guarantee regarding the accuracy, timeliness or completeness of our website or its contents including this tutorial. If you discover any errors on our website or in this tutorial, please notify us at contact@tutorialspoint.com.

Table of Contents

tutorialspoint
SIMPLYEASYLEARNING

tutorialspoint
SIMPLYEASYLEARNING

tutorialspoint
SIMPLYEASYLEARNING

tutorialspoint
SIMPLYEASYLEARNING

1. OBIEE – Data Warehouse

In today's competitive market, most successful companies respond quickly to market changes and opportunities. The requirement to respond quickly is by effective and efficient use of data and information. "**Data Warehouse**" is a central repository of data that is organized by category to support the organization's decision makers. Once data is stored in a data warehouse, it can be accessed for analysis.

The term "Data Warehouse" was first invented by Bill Inmon in 1990. According to him, "Data warehouse is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process."

Ralph Kimball provided a definition of data warehouse based on its functionality. He said, "Data warehouse is a copy of transaction data specifically structured for query and analysis."

Data Warehouse (DW or DWH) is a system used for analysis of data and reporting purposes. They are repositories that saves data from one or more heterogeneous data sources. They store both current and historical data and are used for creating analytical reports. DW can be used to create interactive dashboards for the senior management.

For example, analytic reports can contain data for quarterly comparisons or for annual comparison of sales report for a company.

Data in DW comes from multiple operational systems like sales, human resource, marketing, warehouse management, etc. It contains historical data from different transaction systems but it can also include data from other sources. DW is used to separate data processing and analysis workload from transaction workload and enables to consolidate the data from several data sources.

The Need for Data Warehouse

For example - You have a home loan agency, where data comes from multiple SAP/non-SAP applications such as marketing, sales, ERP, HRM, etc. This data is extracted, transformed and loaded into DW. If you have to do quarterly/annual sales comparison of a product, you cannot use an operational database as this will hang the transaction system. This is where the need for using DW arises.

Characteristics of a Data Warehouse

Some of the key characteristics of DW are:

- It is used for reporting and data analysis.
- It provides a central repository with data integrated from one or more sources.
- It stores current and historical data.

Data Warehouse vs. Transactional System

Following are few differences between Data Warehouse and Operational Database (Transaction System):

- Transactional system is designed for known workloads and transactions like updating a user record, searching a record, etc. However, DW transactions are more complex and present a general form of data.

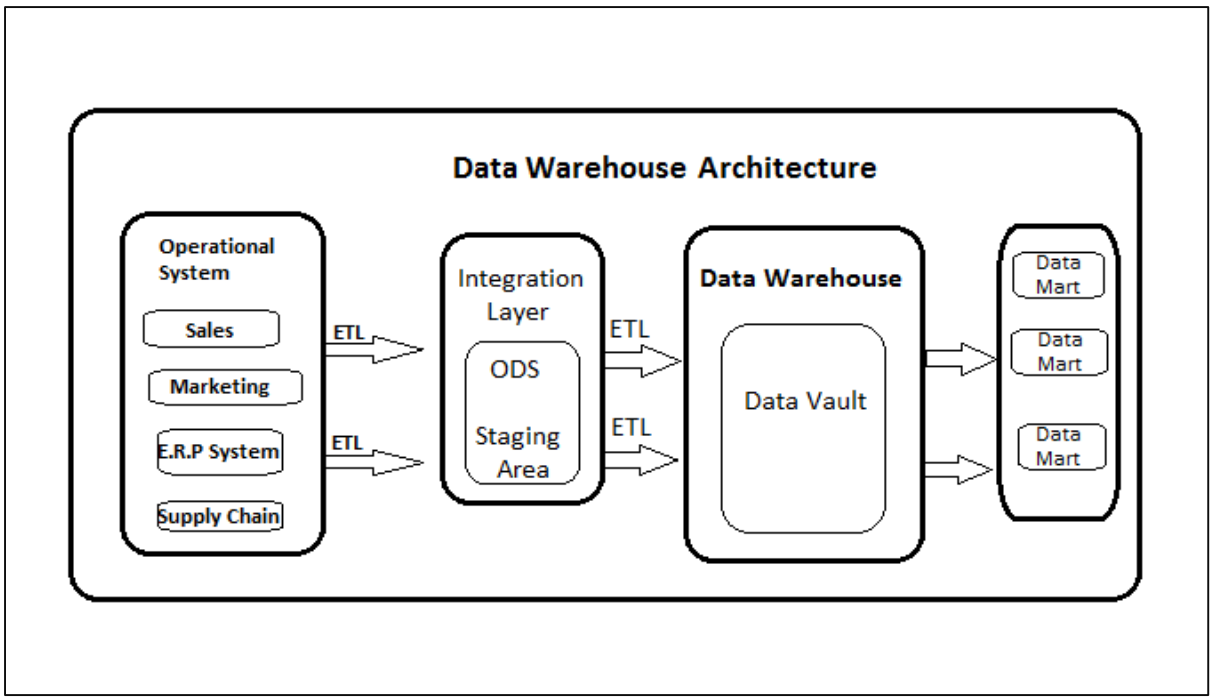
- Transactional system contains the current data of an organization whereas DW normally contains historical data.

- Transactional system supports parallel processing of multiple transactions. Concurrency control and recovery mechanisms are required to maintain consistency of the database.

- Operational database query allows to read and modify operations (delete and update), while an OLAP query needs only read-only access of stored data (select statement).

- DW involves data cleaning, data integration, and data consolidations.

DW has a three-layer architecture: Data Source Layer, Integration Layer, and Presentation Layer. The following diagram shows the common architecture of a Data Warehouse system.



Types of Data Warehouse System

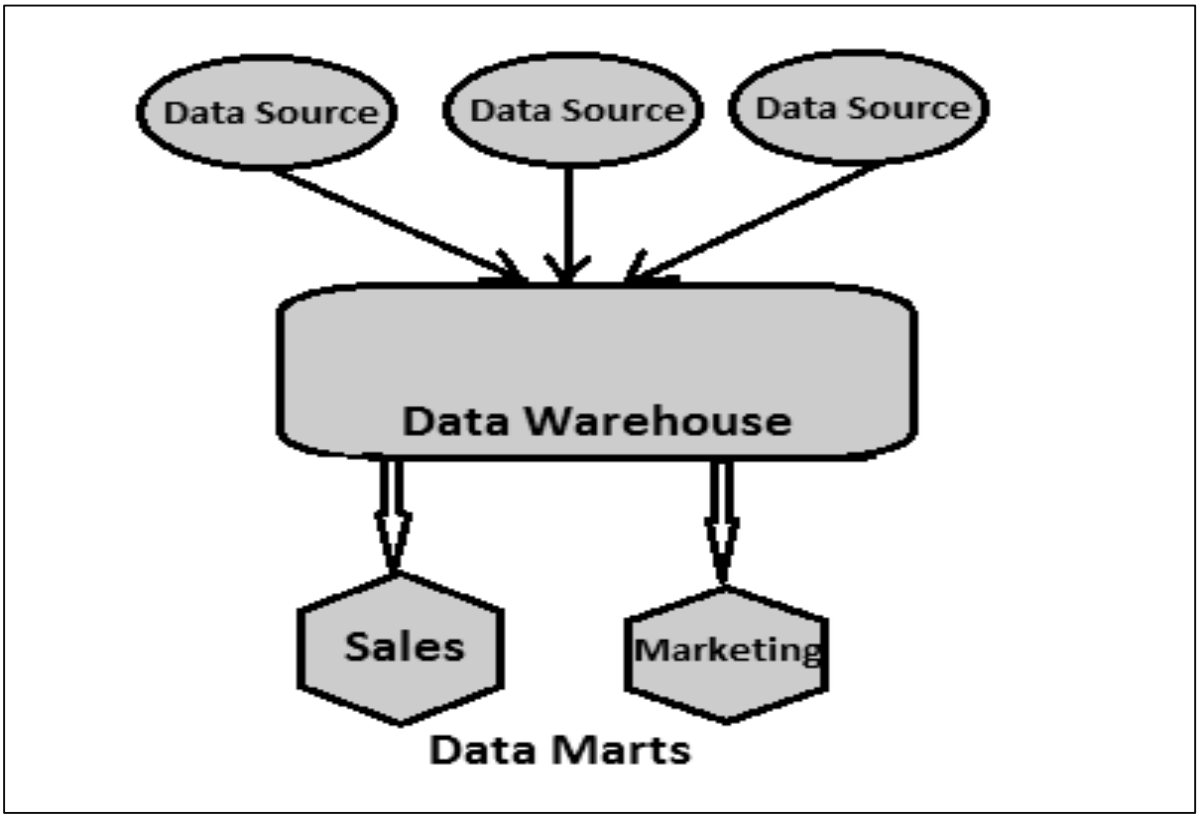
Following are the types of DW system:

- Data Mart
- Online Analytical Processing (OLAP)
- Online Transaction Processing (OLTP)
- Predictive Analysis

Data Mart

Data Mart is the simplest form of DW and it normally focuses on a single functional area, such as sales, finance or marketing. Hence, data mart usually gets data only from few data sources.

Sources could be an internal transaction system, a central data warehouse, or an external data source application. De-normalization is the norm for data modeling techniques in this system.



tutorialspoint
SIMPLYEASYLEARNING

Online Analytical Processing (OLAP)

An OLAP system contains less number of transactions but involves complex calculations like use of Aggregations - Sum, Count, Average, etc.

What is Aggregation?

We save tables with aggregated data like yearly (1 row), quarterly (4 rows), monthly (12 rows) and now we want to compare data, like Yearly only 1 row will be processed. However, in an un-aggregated data, all the rows will be processed.

OLAP system normally stores data in multidimensional schemas like Star Schema, Galaxy schemas (with Fact and Dimensional tables are joined in logical manner).

In an OLAP system, response time to execute a query is an effectiveness measure. OLAP applications are widely used by Data Mining techniques to get data from OLAP systems. OLAP databases store aggregated historical data in multi-dimensional schemas. OLAP systems have data latency of a few hours as compared to Data Marts where latency is normally closer to few days.

Online Transaction Processing (OLTP)

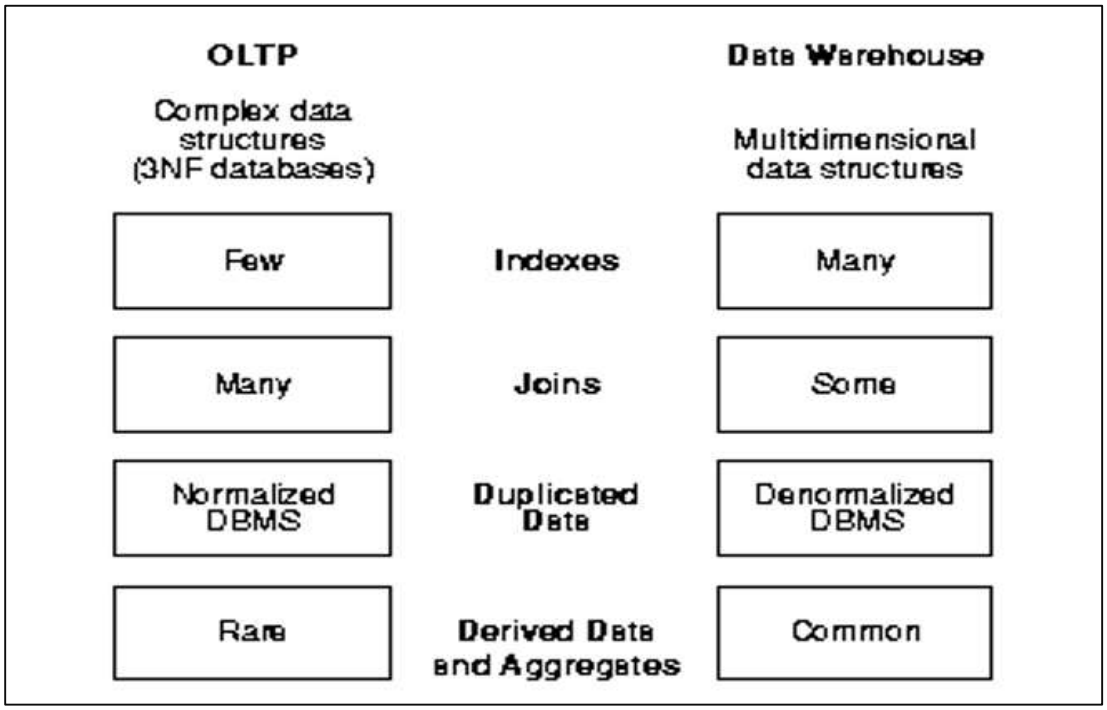
An OLTP system is known for large number of short online transactions like insert, update, delete, etc. OLTP systems provide fast query processing and also responsible to provide data integrity in multi-access environment.

For an OLTP systems, effectiveness is measured by the number of transactions processed per second. OLTP systems normally contain only current data. The schema used to store transactional databases is the entity model. Normalization is used for data modeling techniques in OLTP system.

tutorialspoint
SIMPLYEASYLEARNING

OLTP vs OLAP

The following illustration shows the key differences between an OLTP and OLAP system.



Indexes: In an OLTP system, there are only few indexes while in an OLAP system there are many indexes for performance optimization.

Joins: In an OLTP system, large number of joins and data is normalized; however, in an OLAP system there are less joins and de-normalized.

Aggregation: In an OLTP system, data is not aggregated while in an OLAP database more aggregations are used.

tutorialspoint
SIMPLYEASYLEARNING

2. OBIEE – Dimensional Modeling

Dimensional modeling provides set of methods and concepts that are used in DW design. According to DW consultant, Ralph Kimball, dimensional modeling is a design technique for databases intended to support end-user queries in a data warehouse. It is oriented around understandability and performance. According to him, although transaction-oriented ER is very useful for the transaction capture, it should be avoided for end-user delivery.

Dimensional modeling always uses facts and dimension tables. Facts are numerical values which can be aggregated and analyzed on the fact values. Dimensions define hierarchies and description on fact values.

Dimension Table

Dimension table stores the attributes that describe objects in a Fact table. A Dimension table has a primary key that uniquely identifies each dimension row. This key is used to associate the Dimension table to a Fact table.

Dimension tables are normally de-normalized as they are not created to execute transactions and only used to analyze data in detail.

Example

In the following dimension table, the customer dimension normally includes the name of customers, address, customer id, gender, income group, education levels, etc.

Customer ID	Name	Gender	Income	Education	Religion
1	Brian Edge	M	2	3	4
2	Fred Smith	M	3	5	1
3	Sally Jones	F	1	7	3

Fact Tables

Fact table contains numeric values that are known as measurements. A Fact table has two types of columns: facts and foreign key to dimension tables.

Measures in Fact table are of three types:

- **Additive** – Measures that can be added across any dimension.
- **Non-Additive** – Measures that cannot be added across any dimension.
- **Semi-Additive** – Measures that can be added across some dimensions.

tutorialspoint
SIMPLYEASYLEARNING

Example

Time ID	Product ID	Customer ID	Unit Sold
4	17	2	1
8	21	3	2
8	4	1	1

This fact tables contains foreign keys for time dimension, product dimension, customer dimension and measurement value unit sold.

Suppose a company sells products to customers. Every sale is a fact that happens within the company, and the fact table is used to record these facts.

Common facts are: number of unit sold, margin, sales revenue, etc. The dimension table list factors like customer, time, product, etc. by which we want to analyze the data.

Now if we consider the above Fact table and Customer dimension then there will also be a Product and time dimension. Given this fact table and these three dimension tables, we can ask questions like: How many watches were sold to male customers in 2010?

Difference between Dimension and Fact Table

The functional difference between dimension tables and fact tables is that fact tables hold the data we want to analyze and dimension tables hold the information required to allow us to query it.

Aggregate Table

Aggregate table contains aggregated data which can be calculated by using different aggregate functions.

An **aggregate function** is a function where the values of multiple rows are grouped together as input on certain criteria to form a single value of more significant meaning or measurement.

Common aggregate functions include:

- Average()
- Count()
- Maximum()
- Median()
- Minimum()
- Mode()
- Sum()

tutorialspoint
SIMPLYEASYLEARNING

These aggregate tables are used for performance optimization to run complex queries in a data warehouse.

Example

You save tables with aggregated data like yearly (1 row), quarterly (4 rows), monthly (12 rows) and now you have to do comparison of data, like Yearly only 1 row will be processed. However in an un-aggregated table, all the rows will be processed.

MIN	Returns the smallest value in a given column
MAX	Returns the largest value in a given column
SUM	Returns the sum of the numeric values in a given column
AVG	Returns the average value of a given column
COUNT	Returns the total number of values in a given column
COUNT (*)	Returns the number of rows in a table

Select Avg (salary) from employee where title = ‘developer’. This statement will return the average salary for all employees whose title is equal to 'Developer'.

Aggregations can be applied at database level. You can create aggregates and save them in aggregate tables in the database or you can apply aggregate on the fly at the report level.

Note: If you save aggregates at the database level it saves time and provides performance optimization.

tutorialspoint
SIMPLYEASYLEARNING

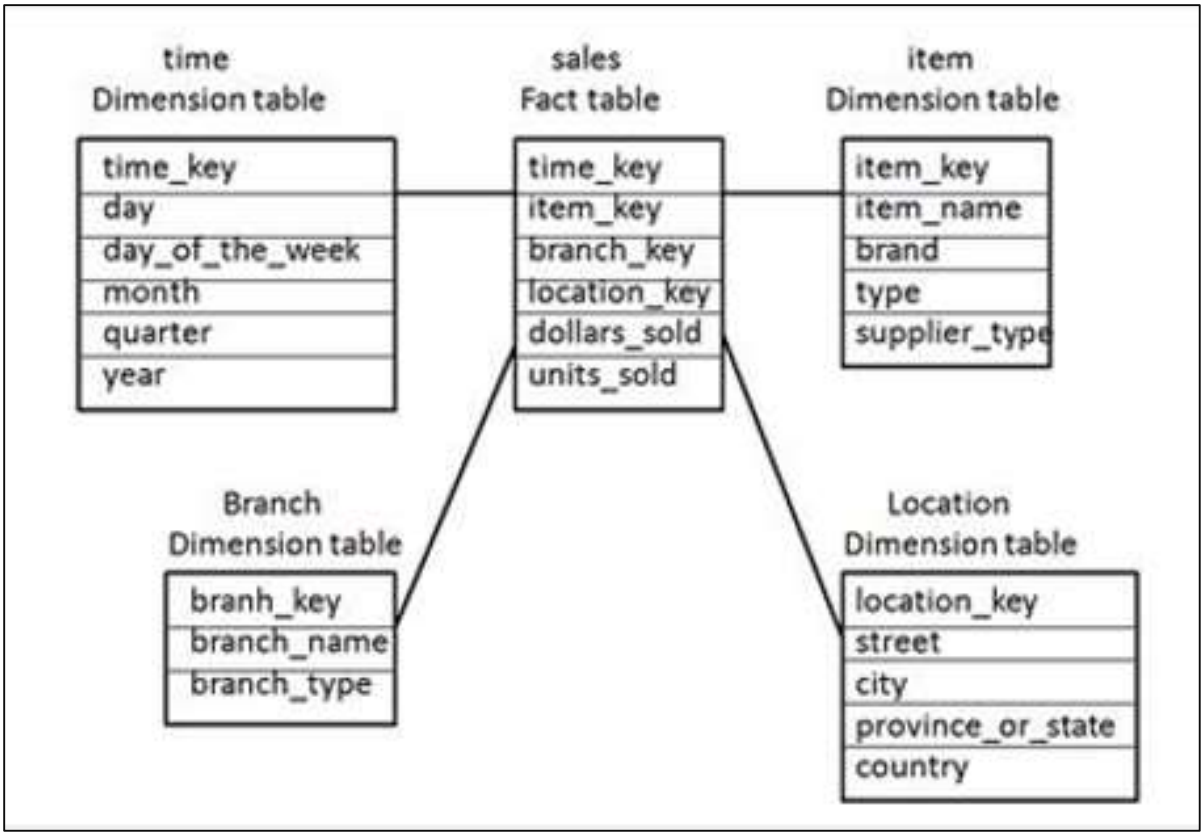
3. OBIEE – Schema

Schema is a logical description of the entire database. It includes the name and description of records of all types including all associated data-items and aggregates. Much like a database, DW also requires to maintain a schema. Database uses relational model, while DW uses Star, Snowflake, and Fact Constellation schema (Galaxy schema).

Star Schema

In a Star Schema, there are multiple dimension tables in de-normalized form that are joined to only one fact table. These tables are joined in a logical manner to meet some business requirement for analysis purpose. These schemas are multidimensional structures which are used to create reports using BI reporting tools.

Dimensions in Star schemas contain a set of attributes and Fact tables contain foreign keys for all dimensions and measurement values.



tutorialspoint
SIMPLYEASYLEARNING

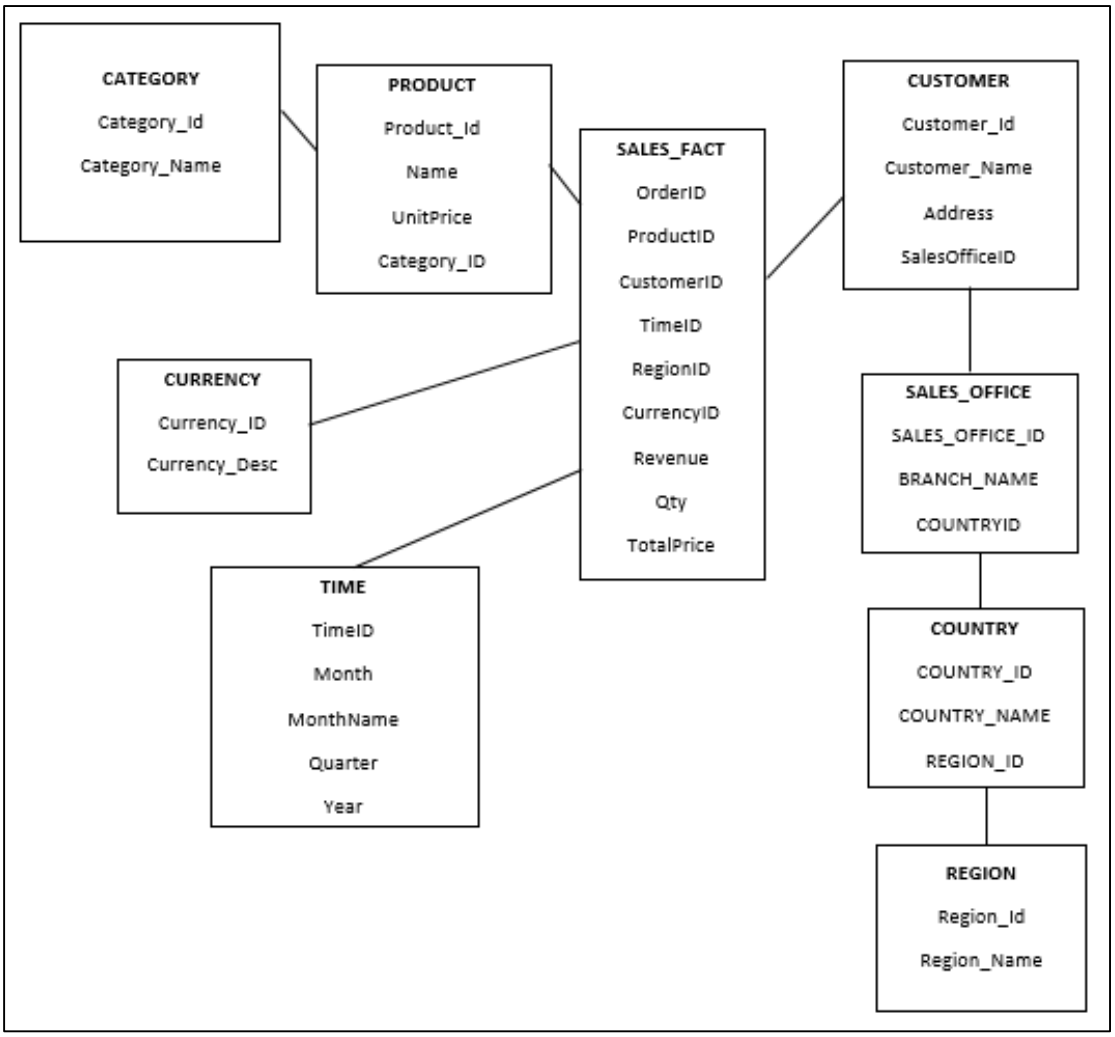
In the above Star Schema, there is a fact table "Sales Fact" at the center and is joined to 4 dimension tables using primary keys. Dimension tables are not further normalized and this joining of tables is known as Star Schema in DW.

Fact table also contains measure values: dollar_sold and units_sold.

Snowflakes Schema

In a Snowflakes Schema, there are multiple dimension tables in normalized form that are joined to only one fact table. These tables are joined in a logical manner to meet some business requirement for analysis purpose.

Only difference between a Star and Snowflakes schema is that dimension tables are further normalized. The normalization splits up the data into additional tables. Due to normalization in the Snowflake schema, the data redundancy is reduced without losing any information and



therefore it becomes easy to maintain and saves storage space.

tutorialspoint
SIMPLYEASYLEARNING

In above Snowflakes Schema example, Product and Customer table are further normalized to save storage space. Sometimes, it also provides performance optimization when you execute a query that requires processing of rows directly in normalized table so it doesn't process rows in primary Dimension table and comes directly to Normalized table in Schema.

Granularity

Granularity in a table represents the level of information stored in the table. High granularity of data means that data is at or near the transaction level, which has more detail. Low granularity means that data has low level of information.

A fact table is usually designed at a low level of granularity. This means that we need to find the lowest level of information that can be stored in a fact table. In date dimension, the granularity level could be year, month, quarter, period, week, and day.

The process of defining granularity consists of two steps:

- Determining the dimensions that are to be included.
- Determining the location to place the hierarchy of each dimension of information.

Slowly Changing Dimensions

Slowly changing dimensions refer to changing value of an attribute over time. It is one of the common concepts in DW.

Example

Andy is an employee of XYZ Inc. He was first located in New York City in July 2015. Original entry in the employee lookup table has the following record:

Employee ID	10001
Name	Andy
Location	New York

At a later date, he has relocated to LA, California. How should XYZ Inc. now modify its employee table to reflect this change?

This is known as "Slowly Changing Dimension" concept.

There are three ways to solve this type of problem:

Solution 1

The new record replaces the original record. No trace of the old record exists.

Slowly Changing Dimension, the new information simply overwrites the original information. In other words, no history is kept.

Employee ID	10001
Name	Andy
Location	LA, California

- **Benefit:** This is the easiest way to handle the Slowly Changing Dimension problem as there is no need to keep track of the old information.

- **Disadvantage:** All historical information is lost.

- **Use:** Solution 1 should be used when it is not required for DW to keep track of historical information.

tutorialspoint
SIMPLYEASYLEARNING

End of ebook preview
If you liked what you saw…
Buy it from our store @ **https://store.tutorialspoint.com**

tutorialspoint
SIMPLYEASYLEARNING