DATA WAREHOUSE

**What is Data Warehouse?**

Data warehouse is a system that accumulates a collection of information and data from different sources. The data in data warehouse is usually corporate information received from their operational system or external data sources. Data warehouse consists of data stores and physical, logical, and conceptual model that a corporate can utilize to support their business goals. The common term used when talking about data warehouse is data warehousing.

**Data Warehousing**

Date warehouse is a system while data warehousing is the system operation. Thus, data warehousing can be defined as the process of collecting, transforming, and loading of corporate data into a data warehouse.

**How Data Warehousing Works**

Data usually flows into the data warehouse from sources like transactional systems and other relational databases. Example of transactional systems are ERP (Enterprise Resource Planning), HRM (Human Resources Management), and CRM (Customer Relationship Management). The data that comes from external data sources can be structured, semi-structured, and unstructured.

Before the data get ingested into the warehouse, it has to be processed and transformed first. The processed data in the warehouse can then be extracted and analysed for making insightful business decision. The way the data in the warehouse can be retrieved is through BI (Business Intelligence) tools, SQL clients, and spreadsheets.

Data warehouse functions on the basis of OLAP (Online Analytical Processing). It helps to make a complicated analysis of multidimensional data. Unlike OLTP (Online Transaction Processing) system in which data are stored in rows and columns (e.g., spreadsheet), the data in OLAP are viewed in three-dimensional view.

**Data Warehousing VS Data Mining**

Data Mining concept is similar to data warehousing, but not the same. During data warehousing operation, data are compiled into a database. On the other hand, data mining is the operation that involves the retrieval or extraction of meaningful data from that database. So, data mining operation comes after the data has been retrieved from data sources, processed, transformed and ingested into the data warehouse.

**Why Data Warehousing**

A corporate that encourages data warehousing can reap its benefits when it comes to designing new business model. The information from data warehouse can be used to predict marketing and pricing strategies, product development and customer satisfaction. It can provide historical analysis and be used for forecasting.

**Types of Data Warehouse**

1. **Enterprise Data Warehouse**
2. **Data Marts or Data Lakes**
3. **Operational Data Store**

**Enterprise Data Warehouse**

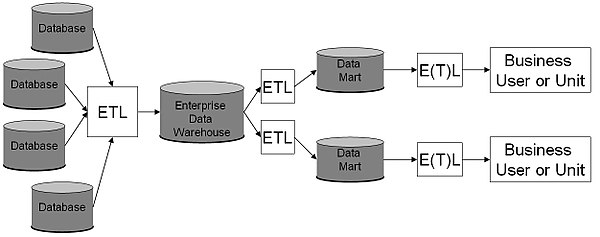
Enterprise Data Warehouse can be thought of as a warehouse where a unified information and data of an organisation can be found.

There three data storage layers in this type of warehouse:

1. **Raw Staging Layer:** The raw staging layer is the stage at which all the data, including raw data information are collected from different data sources.
2. **Atomic Layer:** In the atomic layer, the data from raw staging layer are processed, transformed and linked together.
3. **Dimensional Layer:** The dimensional layer provides the data that the organisation needs the most. This is where the most useful and needed data to the organisation is stored.

**Data Marts or Data Lakes**

Data Mart is a subset of data warehouse where part of the data in the data warehouse is stored. A data warehouse can have many data marts or data lakes that hold the data of a particular subject, e.g Sales data. End users can only access the data marts where the particular data they need are stored.



**Operation Data Store**

Operational Data Store or ODS is used to store the current data of a particular organisation. The data in ODS come from different sources. Unlike data warehouse, ODS does not keep historical data but a current data, e.g., the information about current employees of a company. ODS can feed the data to the data warehouse.

**Sectors Where Data Warehouse is Used**

* Public Sector
* Health Care
* Banking
* Airline
* Telecommunication
* Retail Chain
* Hospitality Industry
* Investment and Insurance