What is Data Warehousing?

A data warehousing is defined as a technique for collecting and managing data from varied sources to provide meaningful business insights. It is a blend of technologies and components which aids the strategic use of data.

It is electronic storage of a large amount of information by a business which is designed for query and analysis instead of transaction processing. It is a process of transforming data into information and making it available to users in a timely manner to make a difference.

The decision support database (Data Warehouse) is maintained separately from the organization's operational database. However, the data warehouse is not a product but an environment. It is an architectural construct of an information system which provides users with current and historical decision support information which is difficult to access or present in the traditional operational data store.

The data warehouse is the core of the BI system which is built for data analysis and reporting.

Data warehouse system is also known by the following name:

- Decision Support System (DSS)
- Executive Information System
- Management Information System
- Business Intelligence Solution
- Analytic Application
- Data Warehouse

How Datawarehouse works?

A Data Warehouse works as a central repository where information arrives from one or more data sources. Data flows into a data warehouse from the transactional system and other relational databases.

Data may be:

1. Structured

2. Semi-structured

3. Unstructured data

The data is processed, transformed, and ingested so that users can access the processed data in the Data Warehouse through Business Intelligence tools, SQL clients, and spreadsheets. A data warehouse merges information coming from different sources into one comprehensive database. By merging all of this information in one place, an organization can analyze its customers more holistically. This helps to ensure that it has considered all the information available. Data warehousing makes data mining possible. Data mining is looking for patterns in the data that may lead to higher sales and profits.

Features of Data Warehouse

- Subject Oriented
- Integrated
- Time Variant
- Non-Volatile

Types of Data Warehouse

Three main types of Data Warehouses are:

1. **Enterprise Data Warehouse:** Enterprise Data Warehouse is a centralized warehouse. It provides decision support service across the enterprise. It offers a unified approach for organizing and representing data. It also provide the ability to classify data according to the subject and give access according to those divisions.

2**. Operational Data Store:** Operational Data Store, which is also called ODS, are nothing but data store required when neither Data warehouse nor OLTP systems support organizations reporting needs. In ODS, Data warehouse is refreshed in real time. Hence, it is widely preferred for routine activities like storing records of the Employees.

3. **Data Mart:** A data mart is a subset of the data warehouse. It specially designed for a particular line of business, such as sales, finance, sales or finance. In an independent data mart, data can collect directly from sources.

Components of Data warehouse

Four components of Data Warehouses are:

Load manager: Load manager is also called the front component. It performs with all the operations associated with the extraction and load of data into the warehouse. These operations include transformations to prepare the data for entering into the Data warehouse.

Warehouse Manager: Warehouse manager performs operations associated with the management of the data in the warehouse. It performs operations like analysis of data to ensure consistency, creation of indexes and views, generation of denormalization and aggregations, transformation and merging of source data and archiving and baking-up data.

Query Manager: Query manager is also known as backend component. It performs all the operation operations related to the management of user queries. The operations of this Data warehouse components are direct queries to the appropriate tables for scheduling the execution of queries.

End-user access tools: This is categorized into five different groups like

1. Data Reporting

2. Query Tools

3. Application development tools

4. EIS tools

5. OLAP tools and data mining tools

Who needs Data warehouse?

Data warehouse is needed for all types of users like:

- Decision makers who rely on mass amount of data Users who use customized, complex processes to obtain information from multiple data sources.
- It is also used by the people who want simple technology to access the data It also essential for those people who want a systematic approach for making decisions.

- If the user wants fast performance on a huge amount of data which is a necessity for reports, grids or charts, then Data warehouse proves useful. Data warehouse is a first step If we want to discover 'hidden patterns' of data-flows and groupings.

What Is a Data Warehouse Used For?

Here, are most common sectors where Data warehouse is used:

Airline: In the Airline system, it is used for operation purpose like crew assignment, analyses of route profitability, frequent flyer program promotions, etc.

Banking: It is widely used in the banking sector to manage the resources available on desk effectively. Few banks also used for the market research, performance analysis of the product and operations.

Healthcare: Healthcare sector also used Data warehouse to strategize and predict outcomes, generate patient's treatment reports, share data with tie-in insurance companies, medical aid services, etc.

Public sector: In the public sector, data warehouse is used for intelligence gathering. It helps government agencies to maintain and analyze tax records, health policy records, for every individual.

Investment and Insurance sector: In this sector, the warehouses are primarily used to analyze data patterns, customer trends, and to track market movements. Retain chain: In retail chains, Data warehouse is widely used for distribution and marketing. It also helps to track items, customer buying pattern, promotions and also used for determining pricing policy.

Telecommunication: A data warehouse is used in this sector for product promotions, sales decisions and to make distribution decisions.

Hospitality Industry: This Industry utilizes warehouse services to design as well as estimate their advertising and promotion campaigns where they want to target clients based on their feedback and travel patterns.

Best practices to implement a Data Warehouse

- Decide a plan to test the consistency, accuracy, and integrity of the data.
- The data warehouse must be well integrated, well defined and time stamped.
- While designing Datawarehouse make sure you use right tool, stick to life cycle, take care about data conflicts and ready to learn you're your mistakes.
- Never replace operational systems and reports
- Don't spend too much time on extracting, cleaning and loading data. Ensure to involve all stakeholders including business personnel in Datawarehouse implementation process. Establish that Data warehousing is a joint/ team project. You don't want to create Data warehouse that is not useful to the end users.
- Prepare a training plan for the end users.