Answer:

# Explain data warehousing.

data warehousing is a process of collecting, storing, and managing large volumes of data from different sources in a centralized repository. The primary objective of data warehousing is to create a unified view of an organization's data from various departments and business processes to enable effective decision making.

A data warehouse is a database that is designed to support business intelligence activities such as data mining, reporting, and analytics. Data warehousing involves several tasks, such as data extraction, data transformation, data loading, and data modeling.

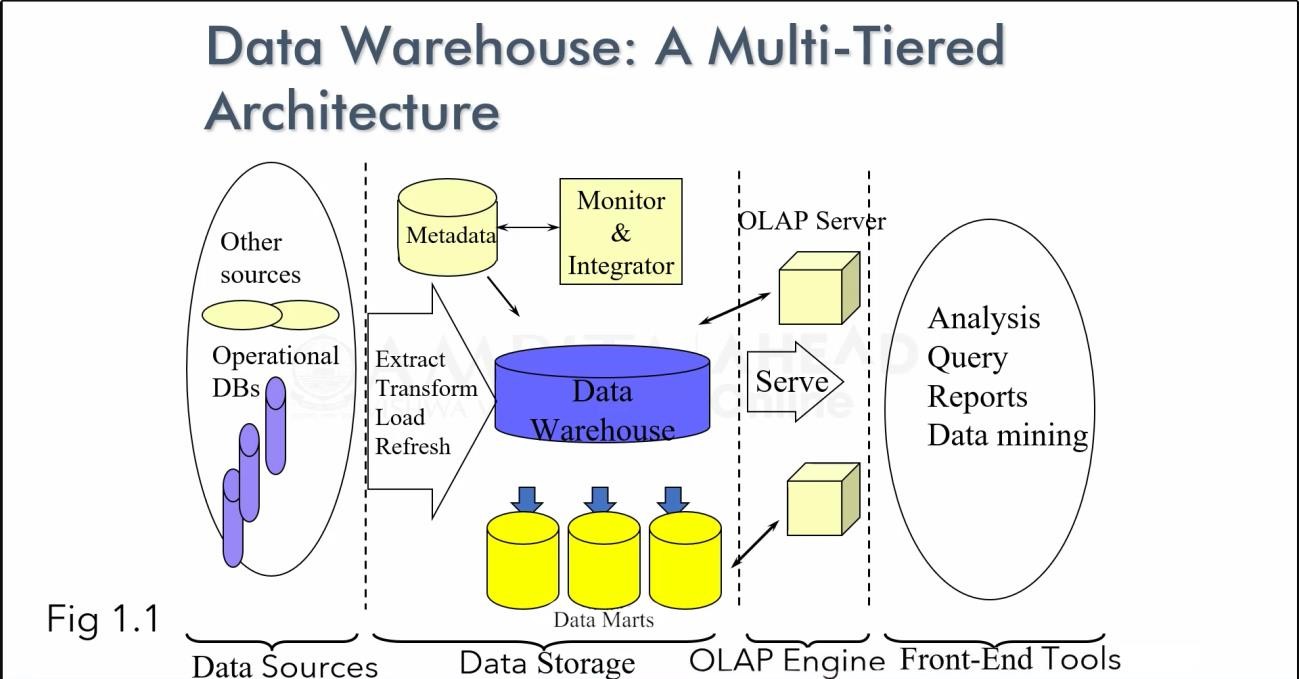
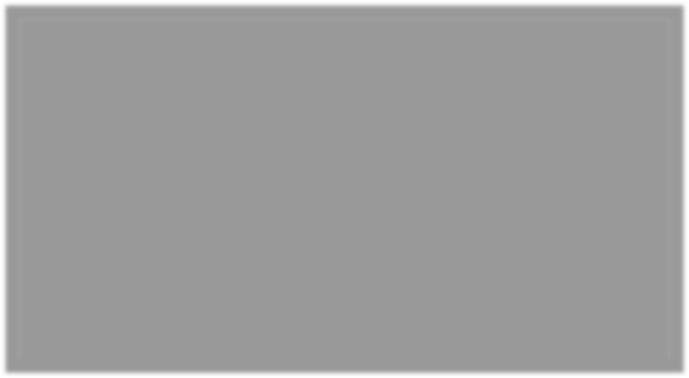
Data extraction involves pulling data from various sources, such as transactional databases, web services, or external files, and integrating them into a common format that can be used in a data warehouse. Data transformation includes cleaning, filtering, and organizing the extracted data to improve its quality and make it more suitable for analysis.

The next step is data loading, where the transformed data is loaded into the data warehouse. This step includes creating indexes, building tables, and loading data into the warehouse. The data is organized in a manner that facilitates faster and easier access to the data.

Finally, data modeling involves organizing the data in the data warehouse in a way that makes it easy to query and analyze. This involves creating a dimensional model that includes facts, which are the numerical data points, and dimensions, which are the characteristics that describe the facts.

Data warehousing can provide several benefits to organizations, including a unified view of data, improved data quality, improved decision-making capabilities, and more efficient use of resources. By enabling organizations to access and analyze data from multiple sources in a centralized location, data warehousing can help organizations gain valuable insights into their operations and make more informed decisions.

# Multi-Tiered Architecture



**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.

# Types of Data Warehouse

## Three main types of Data Warehouses (DWH) are:

1. **Enterprise Data Warehouse (EDW):**

Enterprise Data Warehouse (EDW) 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.

## 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.

## 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.

# General stages of Data Warehouse

Earlier, organizations started relatively simple use of data warehousing. However, over time, more sophisticated use of data warehousing begun.

The following are general stages of use of the data warehouse (DWH):

## Offline Operational Database:

In this stage, data is just copied from an operational system to another server. In this way, loading, processing, and reporting of the copied data do not impact the operational system’s performance.

## Offline Data Warehouse:

Data in the Datawarehouse is regularly updated from the Operational Database. The data in Datawarehouse is mapped and transformed to meet the Datawarehouse objectives.

## Real time Data Warehouse:

In this stage, Data warehouses are updated whenever any transaction takes place in operational database. For example, Airline or railway booking system.

## Integrated Data Warehouse:

In this stage, Data Warehouses are updated continuously when the operational system performs a transaction. The Datawarehouse then generates transactions which are passed back to the operational system.

# 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.

# 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.