

NAME: Aniket Biyani

DATA ENGINEERING- BATCH 1

DAY 1 ASSIGNMENT

EXPLAIN DATA WAREHOUSING

->> A data warehouse is a collection of corporate information, derived directly from operational system and some external data sources.

->> It is a specialized type of database designed for the storage, retrieval, and analysis of large volumes of data.

->> The primary goal of a data warehouse is to support decision-making processes by providing a comprehensive and historical perspective on the organization's data.

->> Ex:- Retail Data Warehouse

FEATURES ->>

1. Data Sources:

- The retail company operates multiple physical stores and an online e-commerce platform.
- Data is generated from various sources, including point-of-sale (POS) systems at stores, online transactions, customer interactions, inventory systems, and marketing campaigns.

2. Data Integration:

- The data from these diverse sources is extracted, transformed, and loaded (ETL) into the data warehouse. For example, sales data from the POS systems is combined with online transaction data, and customer data is integrated from various touchpoints.

3. Subject-Oriented:

- The data warehouse is organized around subjects relevant to the retail business, such as sales, customer relationships, inventory, and

marketing. Each subject area is defined and structured for effective analysis.

4. **Time-Variant:**

- Historical data is stored in the data warehouse, allowing the company to track sales performance, customer behavior, and inventory levels over time. This historical perspective is crucial for identifying trends, seasonality, and making informed decisions.

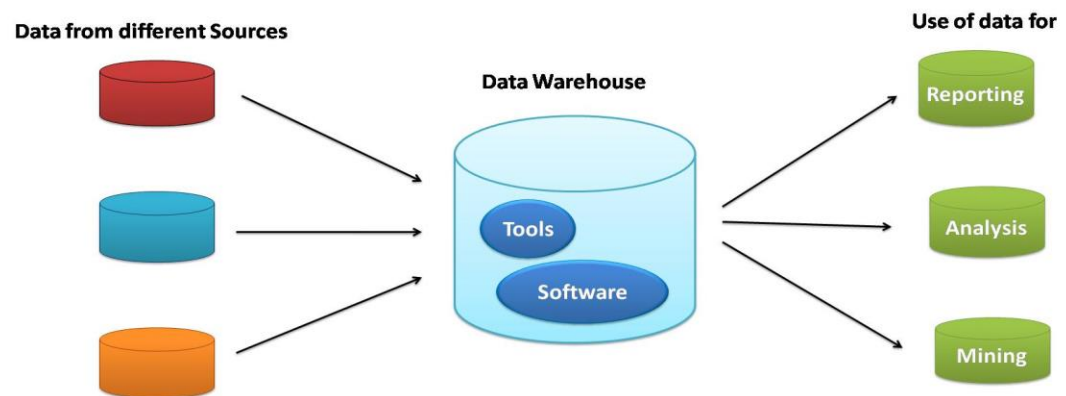


FIGURE OF DATA WAREHOUSE

->> Data warehouses often employ OLAP for interactive data exploration, dimensional modeling for effective organization, and metadata management for comprehensive understanding.

->> In essence it also plays a crucial role in helping organizations derive valuable insights, identify trends, and make informed decisions based on their accumulated and consolidated data.