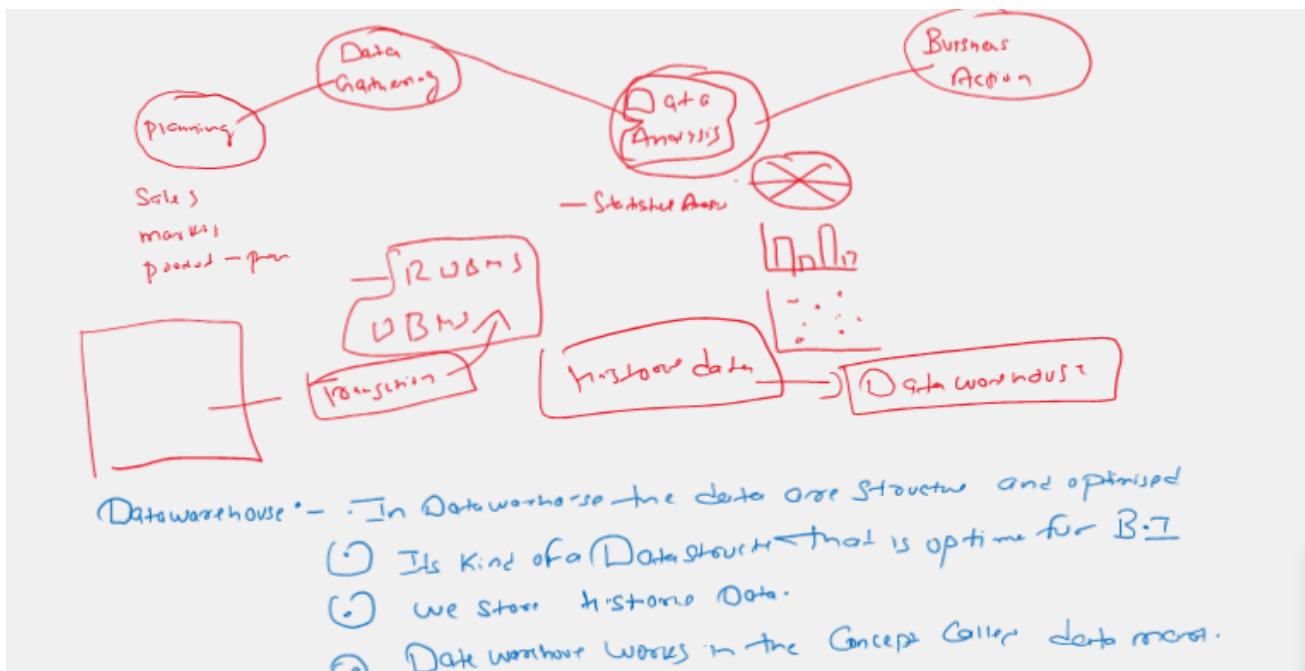


1. Data Warehousing

A **Data Warehouse (DWH)** is a **centralized data storage** system designed for **reporting and analytics**. It integrates data from multiple sources (e.g., sales, marketing) and is optimized for **querying historical data**.

💡 **Real-world example:** Amazon collects millions of transaction logs daily → loads into DWH for quarterly sales analysis.



Data Warehousing

⌚ 2. Purpose of a Data Warehouse

- Enables **business decision-making** using historic & clean data.
- Supports **data analysis, dashboards, KPIs, ML, AI models**.
- Improves **data consistency, governance, and quality**.

⌚ Data Warehousing

3. Data Warehouse Architecture

A typical DWH architecture has **3 layers**:

Layer	Role
Source Layer	Pulls raw data (CRM, ERP, APIs)
Staging Layer	Temporary processing/cleansing
Presentation Layer	Final structured format for analysis

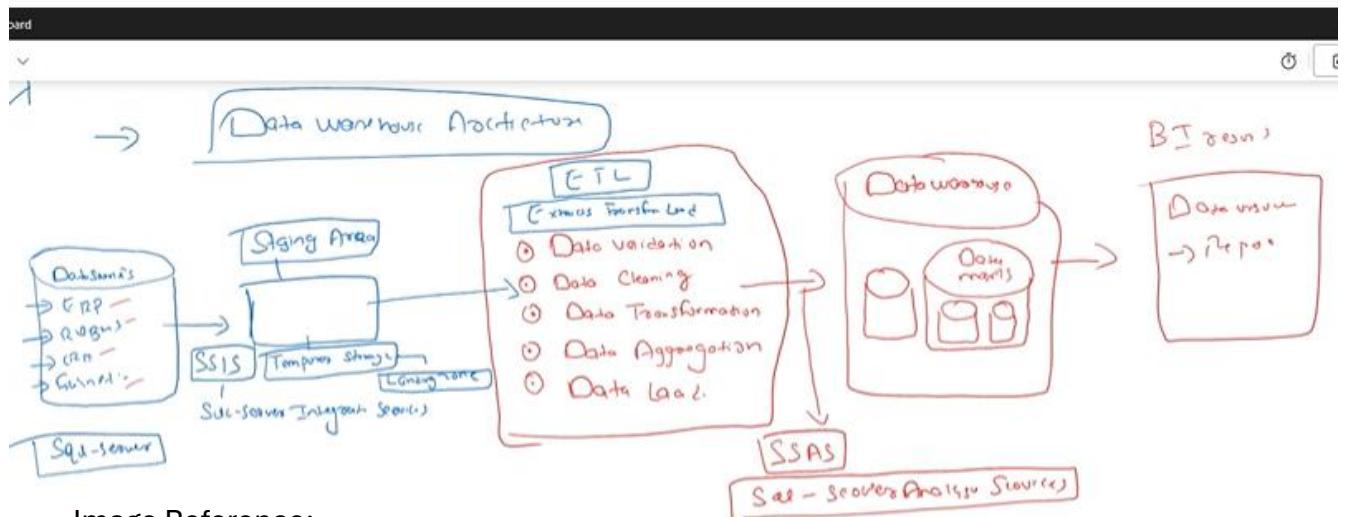


Image Reference:

 [DWH Architecture](#)

4. Operational Data Store (ODS)

 A **real-time or near real-time** temporary database used before data moves to the DWH.

 Ideal for **operational reports** that need fresh data (e.g., today's sales, current order status).

 Think of it as a “buffer zone” before final analysis.

5. OLTP vs OLAP (Warehouse Applications)

Feature	OLTP (Online Transaction Processing)	OLAP (Online Analytical Processing)
Focus	Daily operations	Analysis, reporting

Feature	OLTP (Online Transaction Processing)	OLAP (Online Analytical Processing)
Example	ATM Withdrawal, Shopping Cart	Quarterly Sales Analysis
Data Volume	Small transactions	Large, historical datasets
Tables	Normalized	Denormalized
Limit	100 MB-2GB	100GB-1 TB
Relationship	ER	Schema – snowflake, Star, etc.

OLTP vs OLAP

6. Data Marts

A **data mart** is a subset of the DWH designed for a **specific department or purpose**.

- ◆ *Example:* Sales Mart, HR Mart, Finance Mart.
 - 📦 Instead of querying entire DWH, teams access **custom-cut views**.
-

7. Data Marts vs Data Warehouses

Feature	Data Mart	Data Warehouse
Scope	Department-specific	Enterprise-wide
Size	Smaller	Larger
Development Time	Short	Long
Data Source	One or few	Many

 *Think of DWH as a university and Data Marts as individual departments.*

Data models:

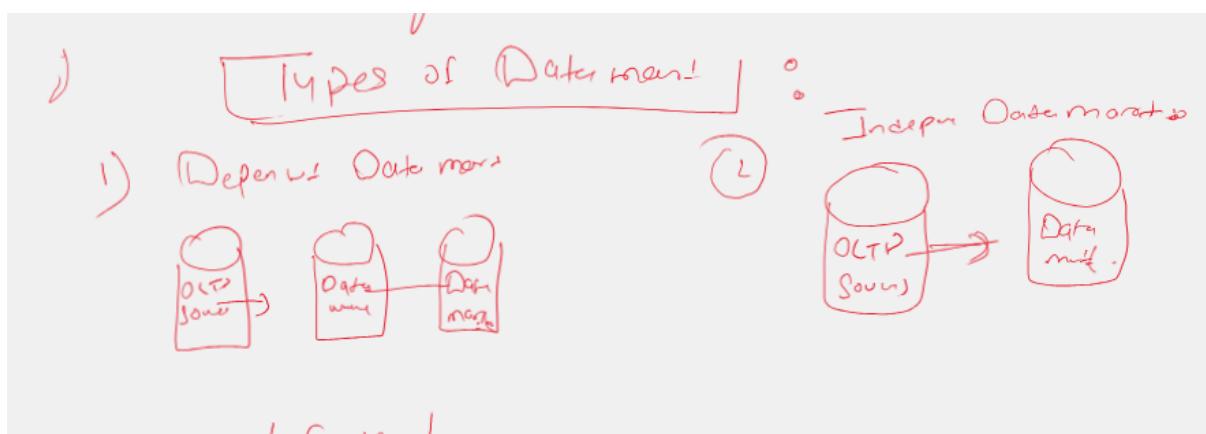
- Star - Snowflake Analysis Sources
- Data marts are the Smaller version of Data warehouse.
- Data marts is being focused in one area.
- It uses less time to create and can be used less memory.

Data warehouse

- ① Enterprise level data
- ② Multiple data source
- ③ Large memory

Data model:

- ① Reproduce specific departmental data
- ② Single source
- ③ Lesser memory



8. Data Warehouse Life Cycle

1. **Requirement Analysis** – Define business goals
2. **Data Modeling** – Star/Snowflake schema
3. **ETL Development** – Extract, Transform, Load
4. **Testing & QA**
5. **Deployment**
6. **Monitoring & Maintenance**

Lifecycle

Dimensionality in Data Warehousing

Dimensionality refers to the structure of data in a **dimensional model** used in Data Warehouses, primarily consisting of **Facts** and **Dimensions**.

It helps in **organizing and analyzing** data across different perspectives (time, geography, product, etc.)

Key Components of Dimensional Modeling

Component Description

Fact Table	Contains measurable, quantitative data (e.g., sales, profit, revenue)
Dimension Table	Contains descriptive attributes (e.g., product name, customer name, date) that provide context to facts

Data warehouse - Dimension modeling

- Delivers data in more understandable format.
- Faster query performance.
- **Facts** → 
(What we want to measure)
Fact contains the measuring or the action of the dimension table. Sales, profit, no of transactions
- **Dimension** :- Dimension is something by which fact table is measured as it contains descriptive information.
 - Dimension contains the attribute on the dimension table (or) the metric.

Schema :- Logical description of entire database

① Star Schema & Snowflake Schema

fact construction scheme.

Bonus Quiz:

Try this interactive data warehouse quiz:

 [Data Warehousing Quiz](#)

Summary

Concept Summary

DWH Storage for analytics-ready data

Purpose Business intelligence, KPIs

Architecture 3-layer: Source → Staging → Presentation

ODS Temporary real-time store

OLTP vs OLAP Transactional vs Analytical

Data Marts Department-level mini-DWH

DWH vs Mart Enterprise-wide vs Specific

Lifecycle Plan → Design → ETL → Report
