Title: General Concept of Data Warehouse

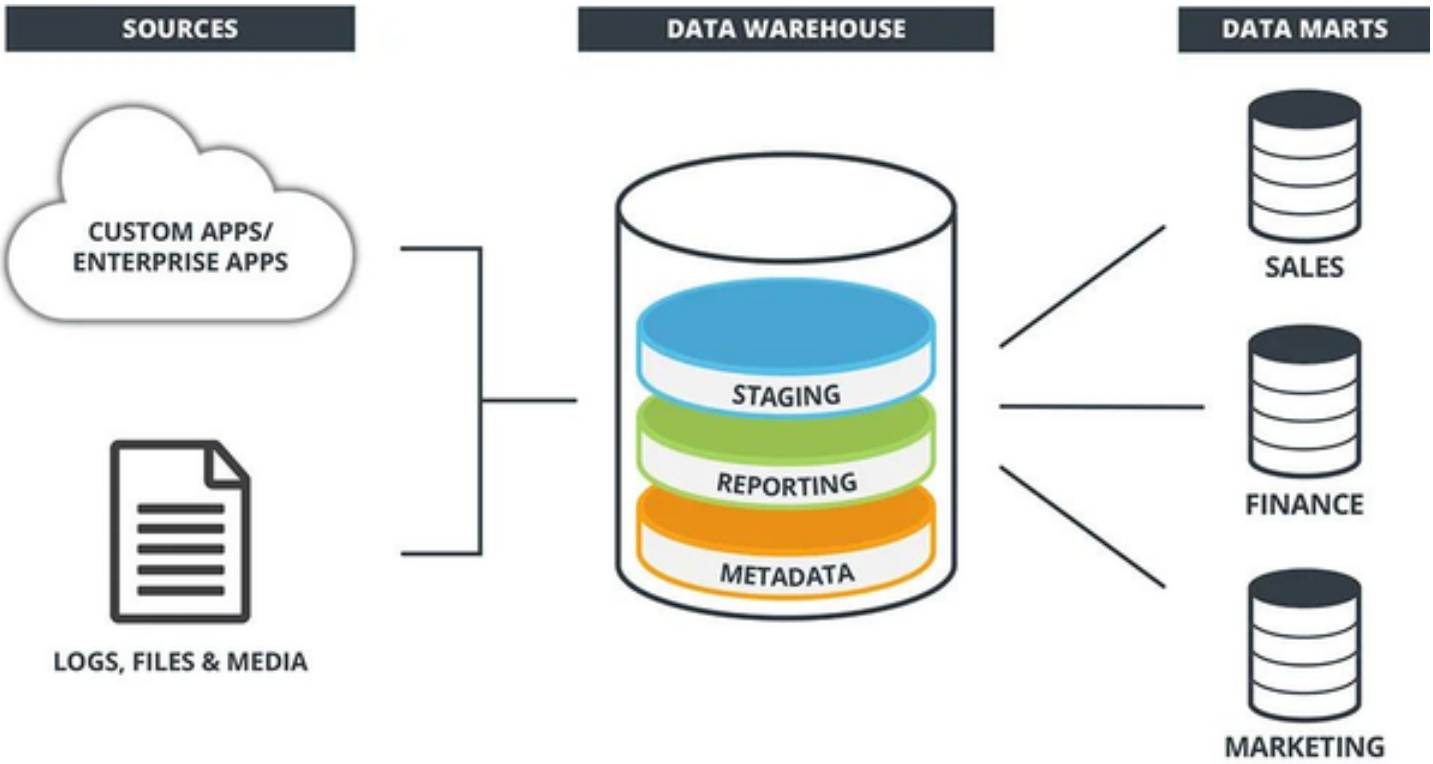
🟦 Slide 1: What is a Data Warehouse?

Description:A **Data Warehouse (DWH)** is a centralized repository that stores integrated data from multiple sources. It is used for reporting and data analysis, and supports decision-making.

Key Points:

- Optimized for read-heavy operations.
- Structured for querying and analytics (OLAP).
- Historical and subject-oriented.

Diagram:



[Operational DBs] --> [ETL Process] --> [Data Warehouse] --> [BI/Reporting Tools]

Example SQL:

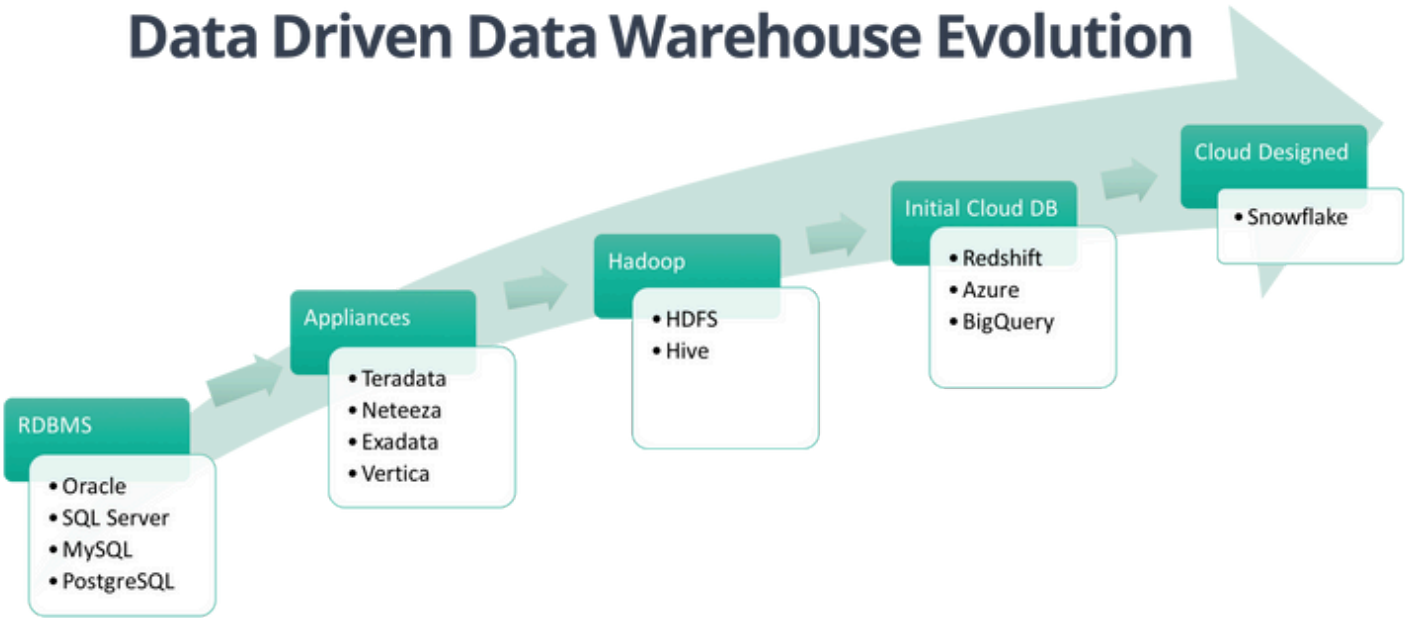
```
-- Creating a simple fact table
CREATE TABLE Sales_Fact (
    Sale_ID INT,
    Date_ID INT,
```

```
    Product_ID INT,
    Customer_ID INT,
    Amount DECIMAL(10,2)
);
```

🟦 Slide 2: History of Data Warehousing

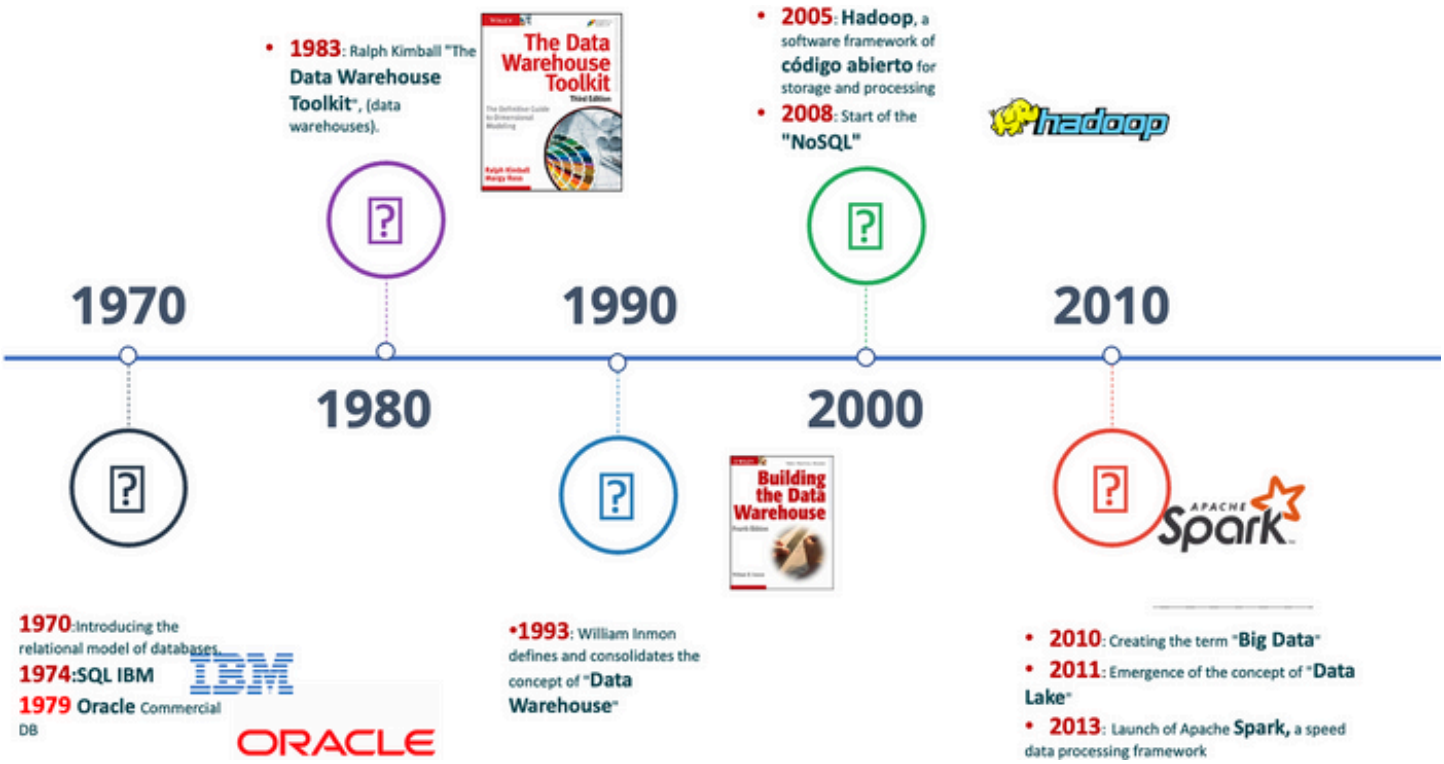
Description:

- Coined in **1988** by **Paul Murphy** and later popularized by **Bill Inmon**.
- Initially developed to separate analytical processing from transactional systems.
- Shifted from **mainframes** → **RDBMS** → **Cloud DWs** (like Redshift, BigQuery, Snowflake).



Timeline:

- 1980s: Initial concept by IBM.
- 1990s: Relational OLAP systems.
- 2000s+: Modern DWs with real-time and big data integration.



🟦 Slide 3: Need for Data Warehouse

Description:

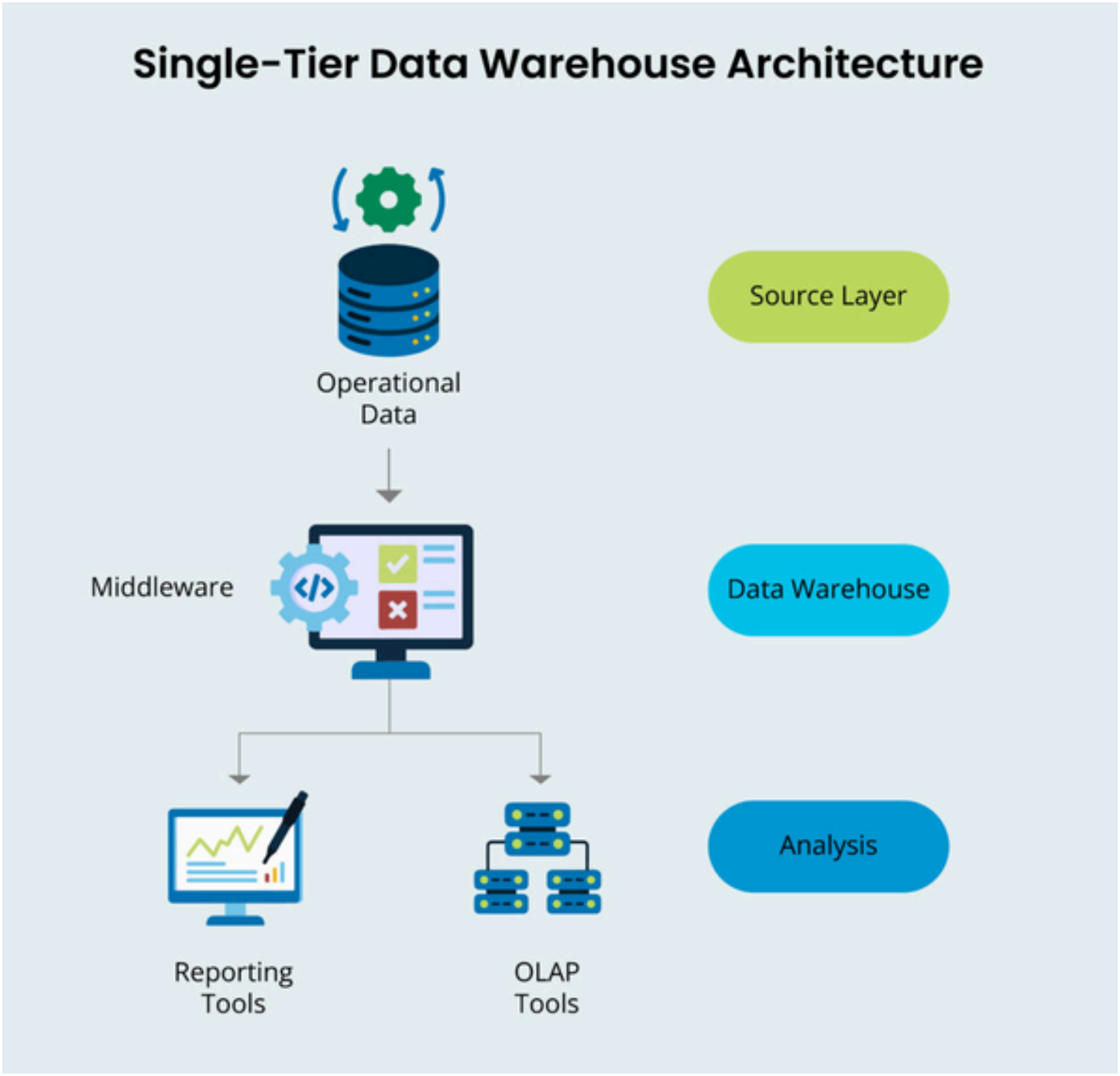
- **Consolidates** disparate data sources.
- Enables **historical analysis** and trend spotting.
- Enhances **data quality** and governance.
- Supports **informed decision-making**.

Use Case:Retail companies use DWH to analyze sales trends across stores and time.

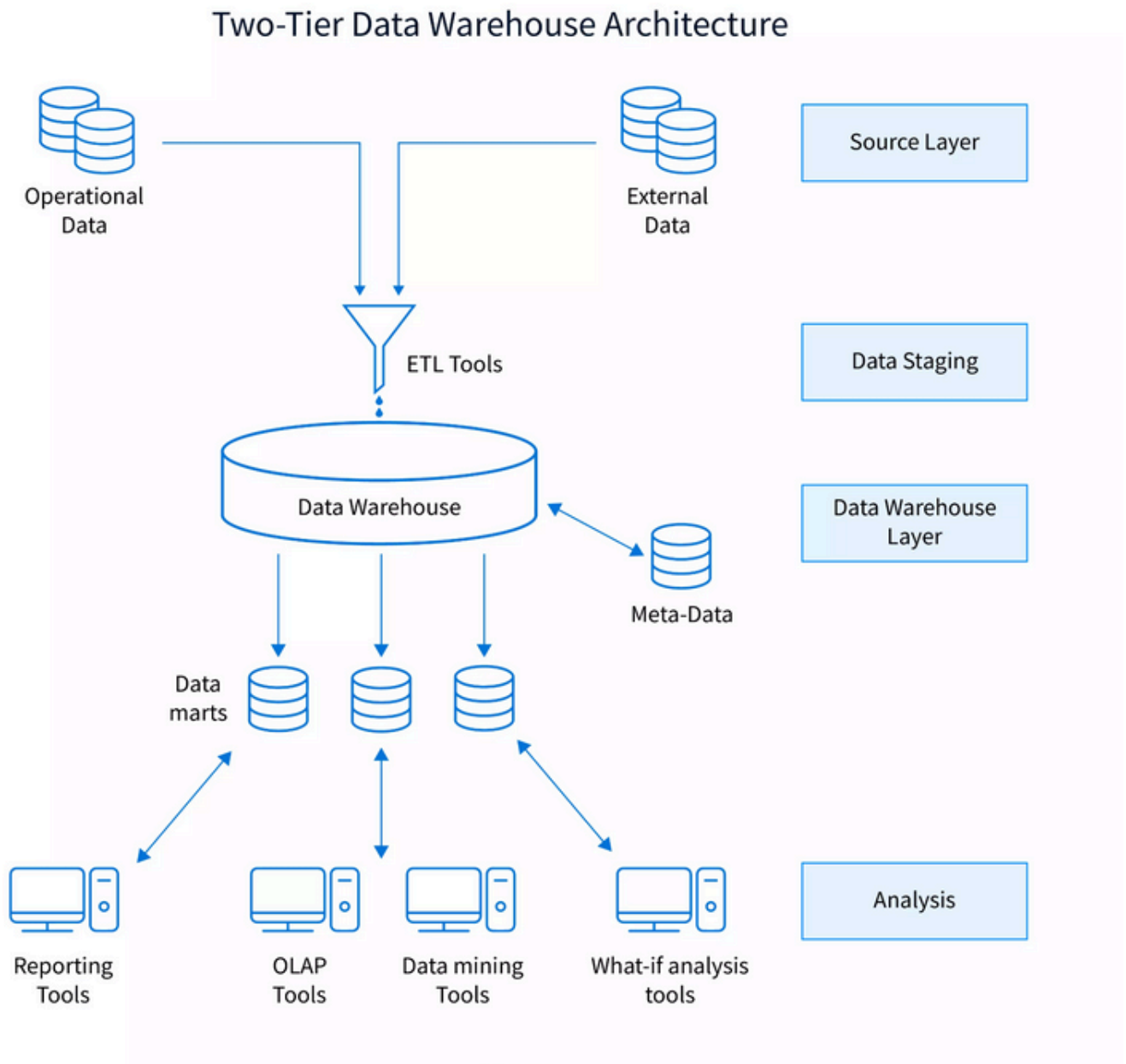
🟦 Slide 4: Data Warehouse Architecture

Description:Three main types:

- **Single-Tier**: Rare, simplified.



- **Two-Tier**: Data source → warehouse → client.



- **Three-Tier**:
  - **Bottom**: Data Sources
  - **Middle**: ETL tools and DW
  - **Top**: BI tools, Data Marts

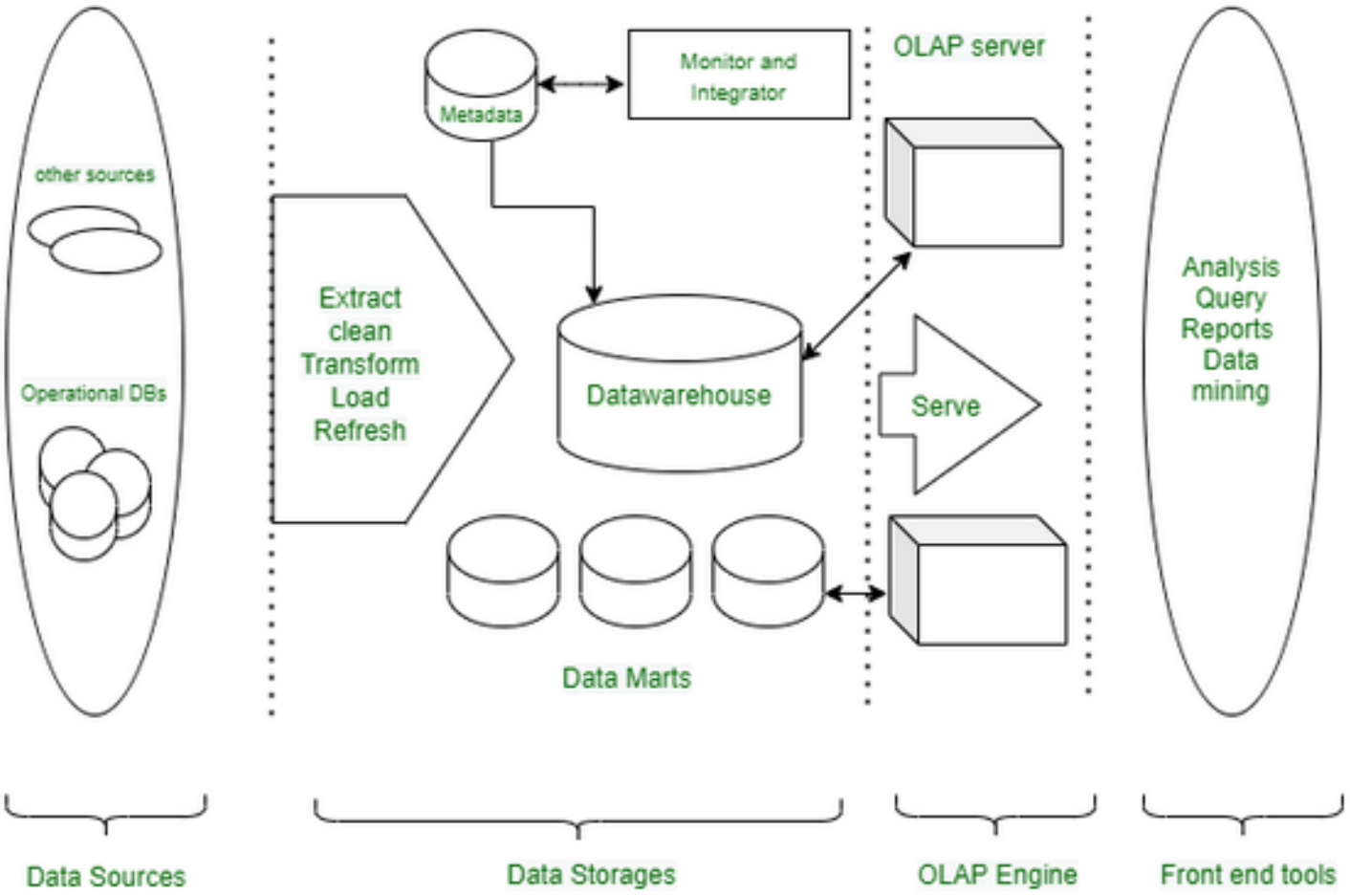
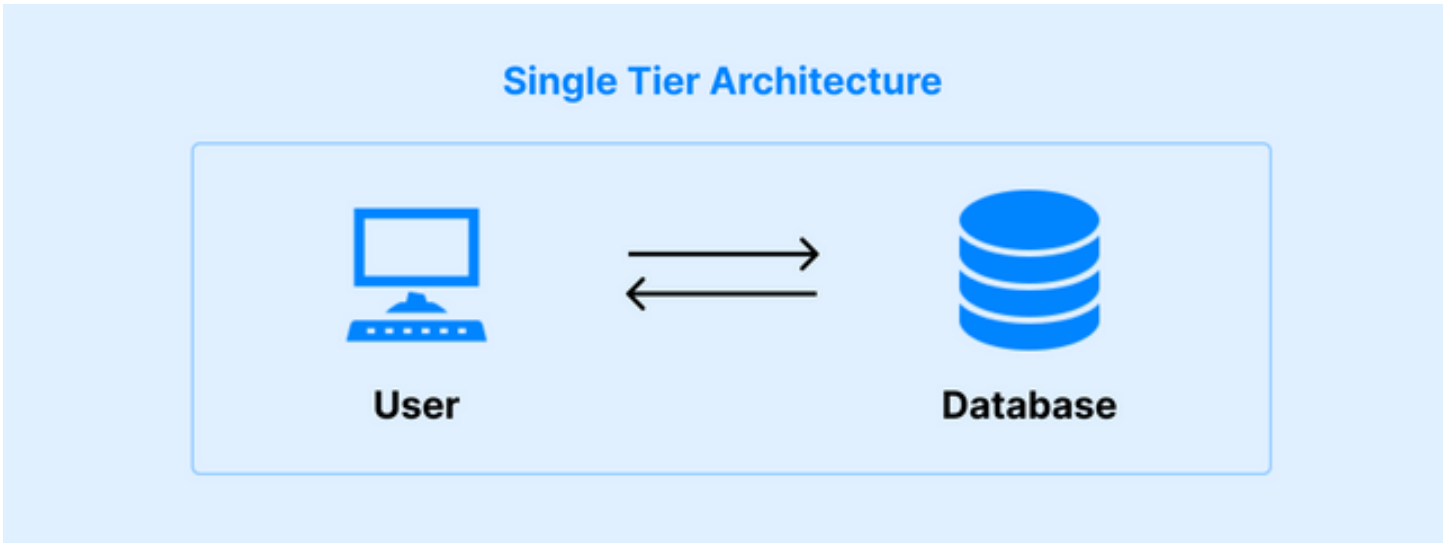
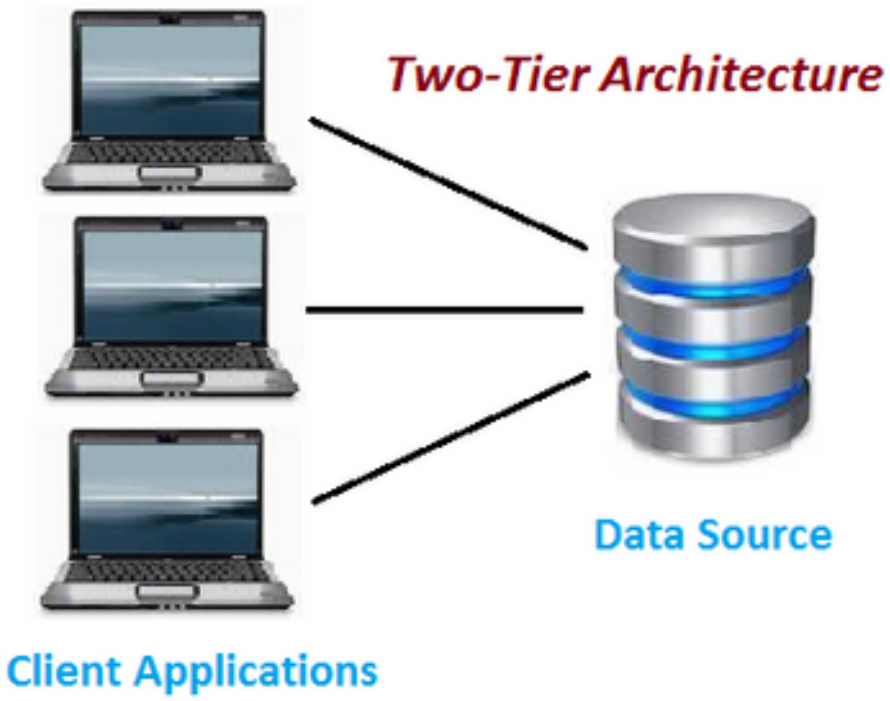


Illustration (text-based):

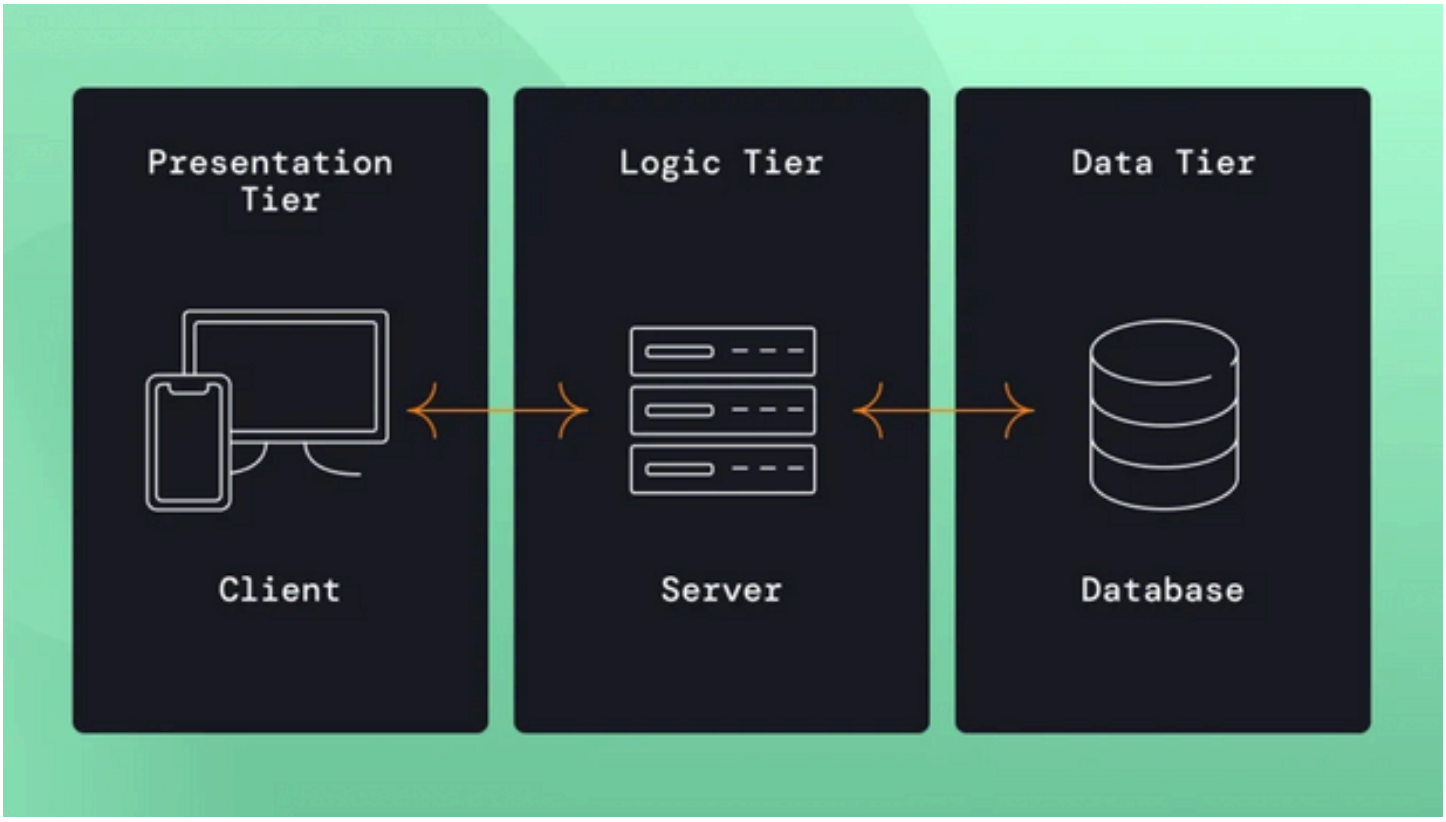
Single Tier



Two Tier



Three Tier



[External Sources]

↓

[ETL Layer]

↓

[Central Data Warehouse]

↓

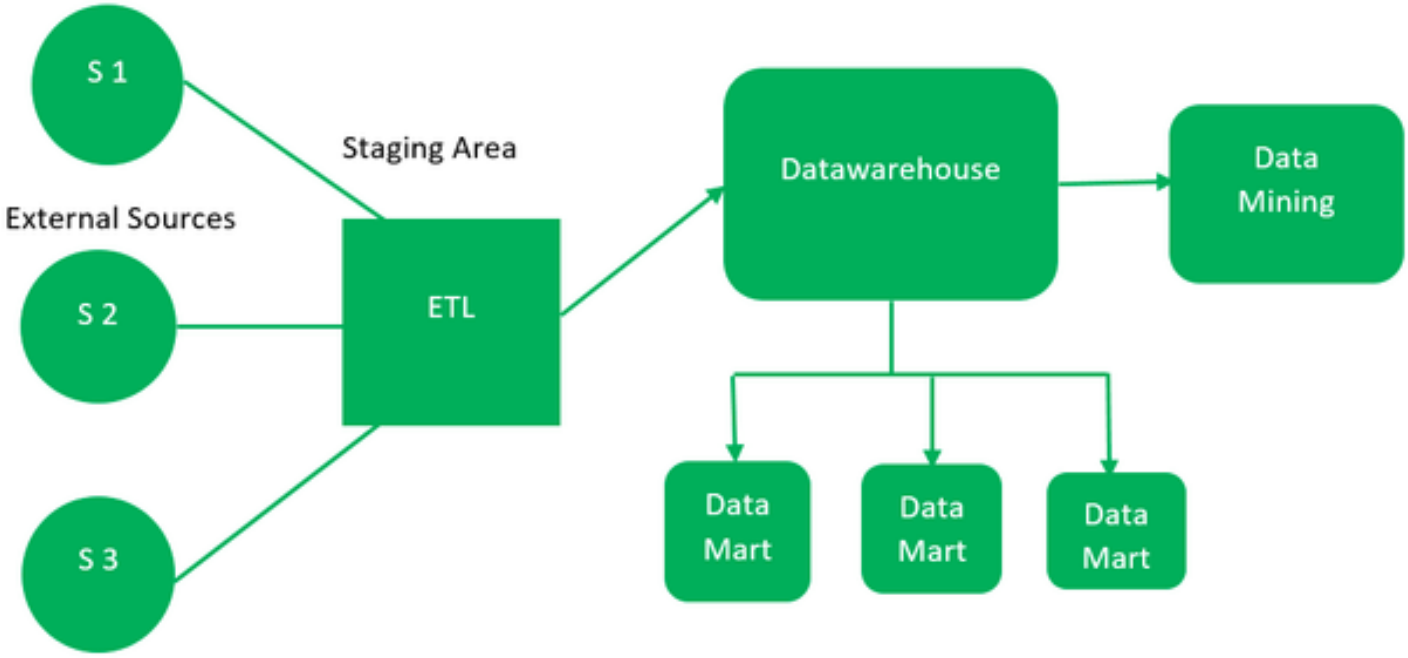
[Data Marts / OLAP Tools / Reports]

🟦 Slide 5: How Data Mining Works with DWH

Description:Data mining is used to discover patterns and relationships from large datasets stored in the warehouse.

Techniques:

- Classification
- Clustering
- Association Rules



Code Example (Pseudocode):

```
-- Find top 3 customer segments by total purchase
SELECT Segment, SUM(Amount) AS Total
FROM Customer_Sales
GROUP BY Segment
ORDER BY Total DESC
LIMIT 3;
```

🟦 Slide 6: Features of Data Warehouse

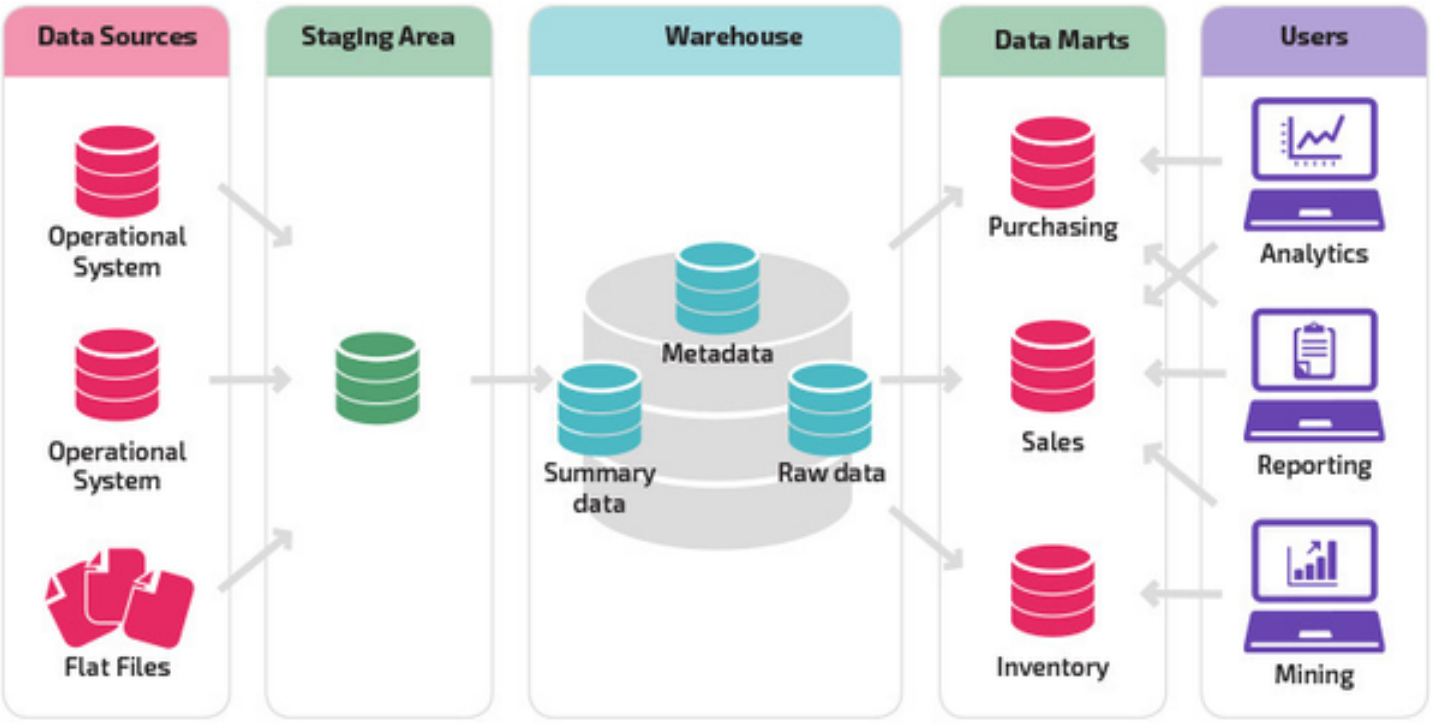
Key Features:

1. **Subject-Oriented** – Focused on high-level entities.
2. **Integrated** – Combines data from various sources.
3. **Non-volatile** – Data is stable and doesn't change.
4. **Time-variant** – Stores historical data.

Real-world Example:A DWH retains 10 years of sales records for analysis.

🟦 Slide 7: Data Mart

Description:A **Data Mart** is a subset of a data warehouse, tailored for a specific department or function.



Types:

- **Dependent**: Created from central DWH.
- **Independent**: Built without a central DWH.

Use Case:Finance team uses a financial data mart; Sales team uses a sales mart.

SQL Example:

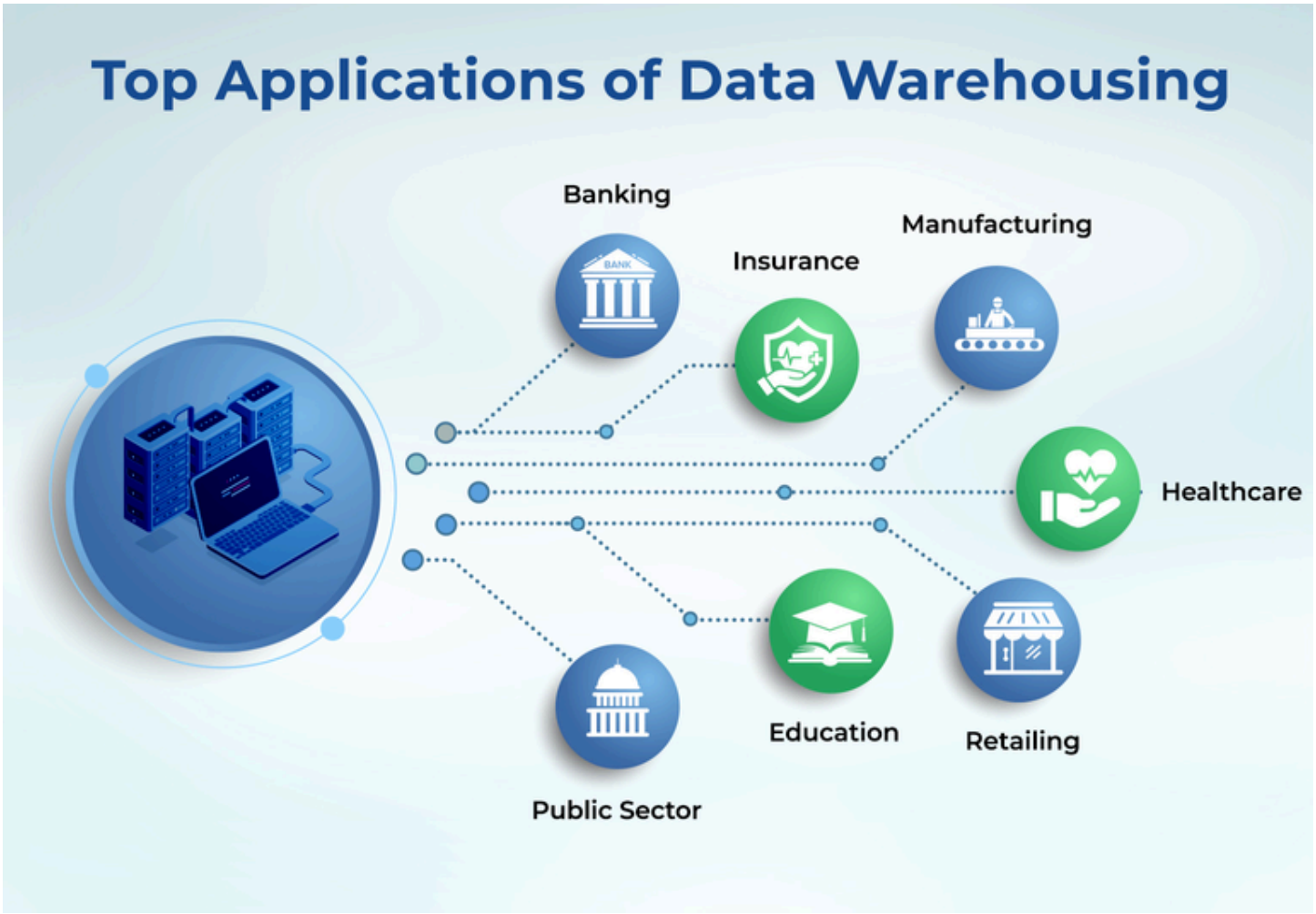
```
-- Data mart for Finance
CREATE VIEW Finance_Mart AS
SELECT Date_ID, Customer_ID, Amount
FROM Sales_Fact
WHERE Department = 'Finance';
```

🟦 Slide 8: Application Areas

Sectors using DWH:

- **Retail**: Inventory & customer behavior analysis
- **Healthcare**: Patient data aggregation
- **Banking**: Fraud detection and trend analysis

- **Telecom**: Call pattern and network usage
- **Manufacturing**: Supply chain optimization

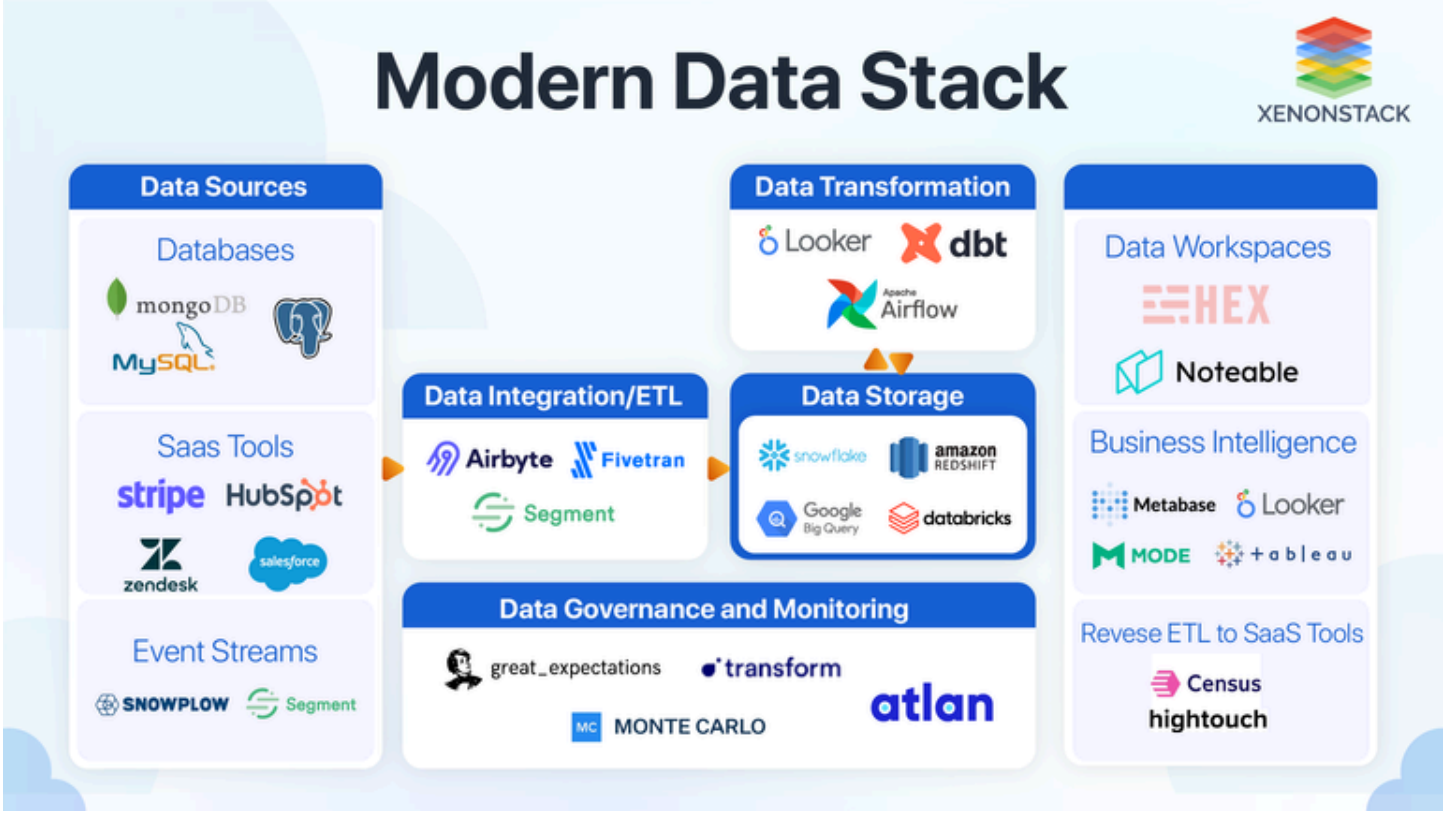


BI Tool Integration:Power BI, Tableau, Looker, Excel dashboards.

🟦 Slide 9: Additional Section – Modern Data Warehousing

New Trends:

- **Cloud DWHs** (Snowflake, BigQuery)
- **Real-Time Analytics**
- **AI & ML integration**
- **Data Lakehouse** (combining lake + warehouse)



Best Practices:

- Design for scalability
- Automate ETL pipelines
- Implement data governance

🟦 Slide 10: Sample Excel Dashboard Ideas

Sheets:

- Sales Overview
- Product Performance
- Customer Analysis

Charts:

- Time-series graphs
- Pie chart by region
- Bar chart by product category