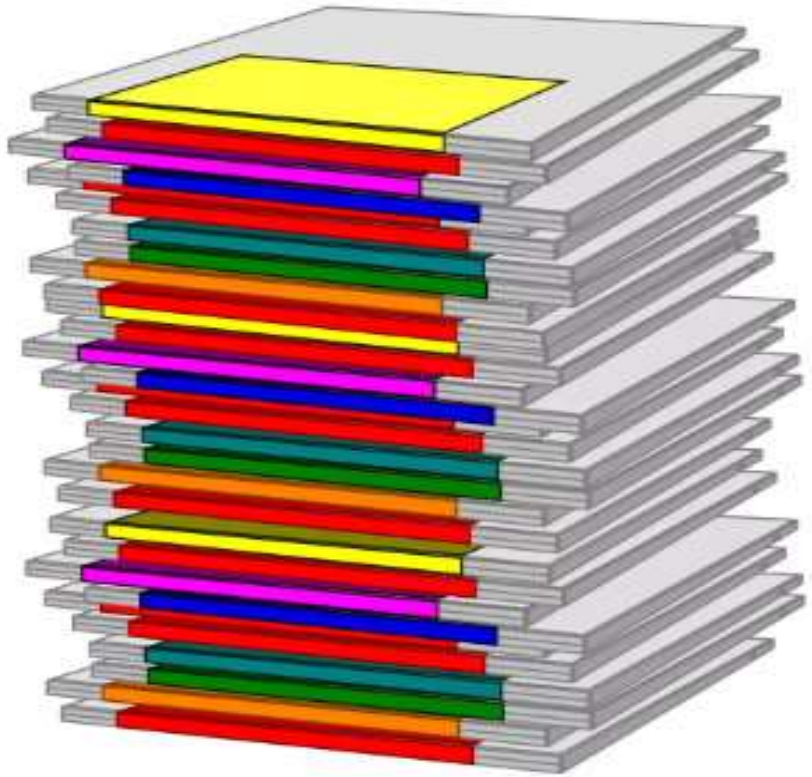
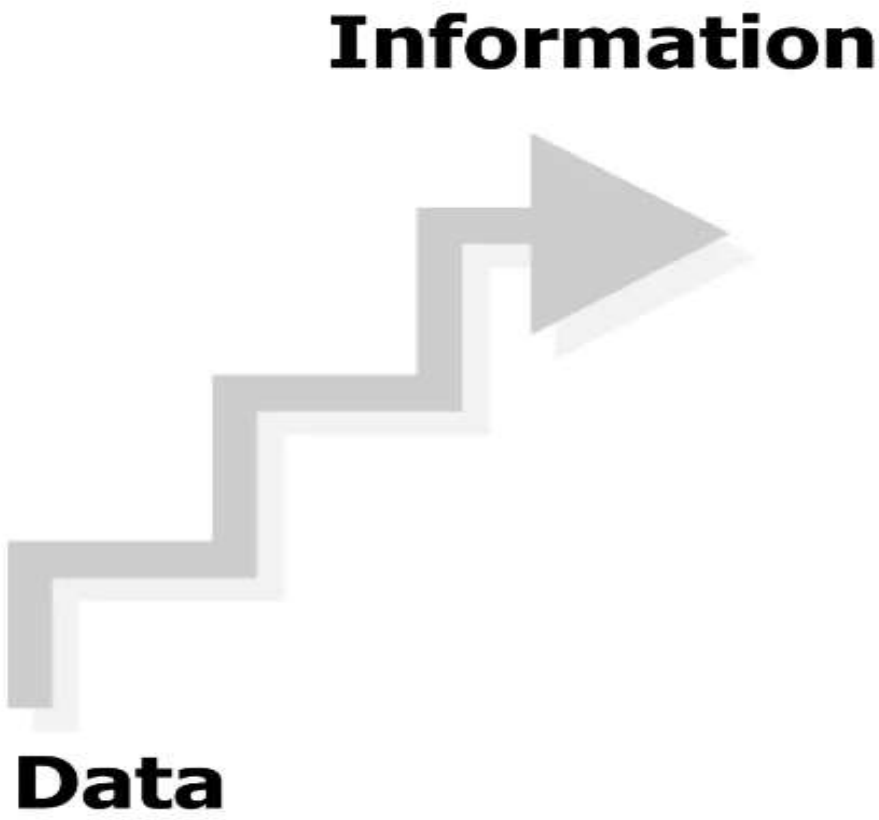
What is a Data Warehouse?

A single, complete and consistent store of data obtained from a variety of different sources made available to end users in a what they can understand and use in a business context.

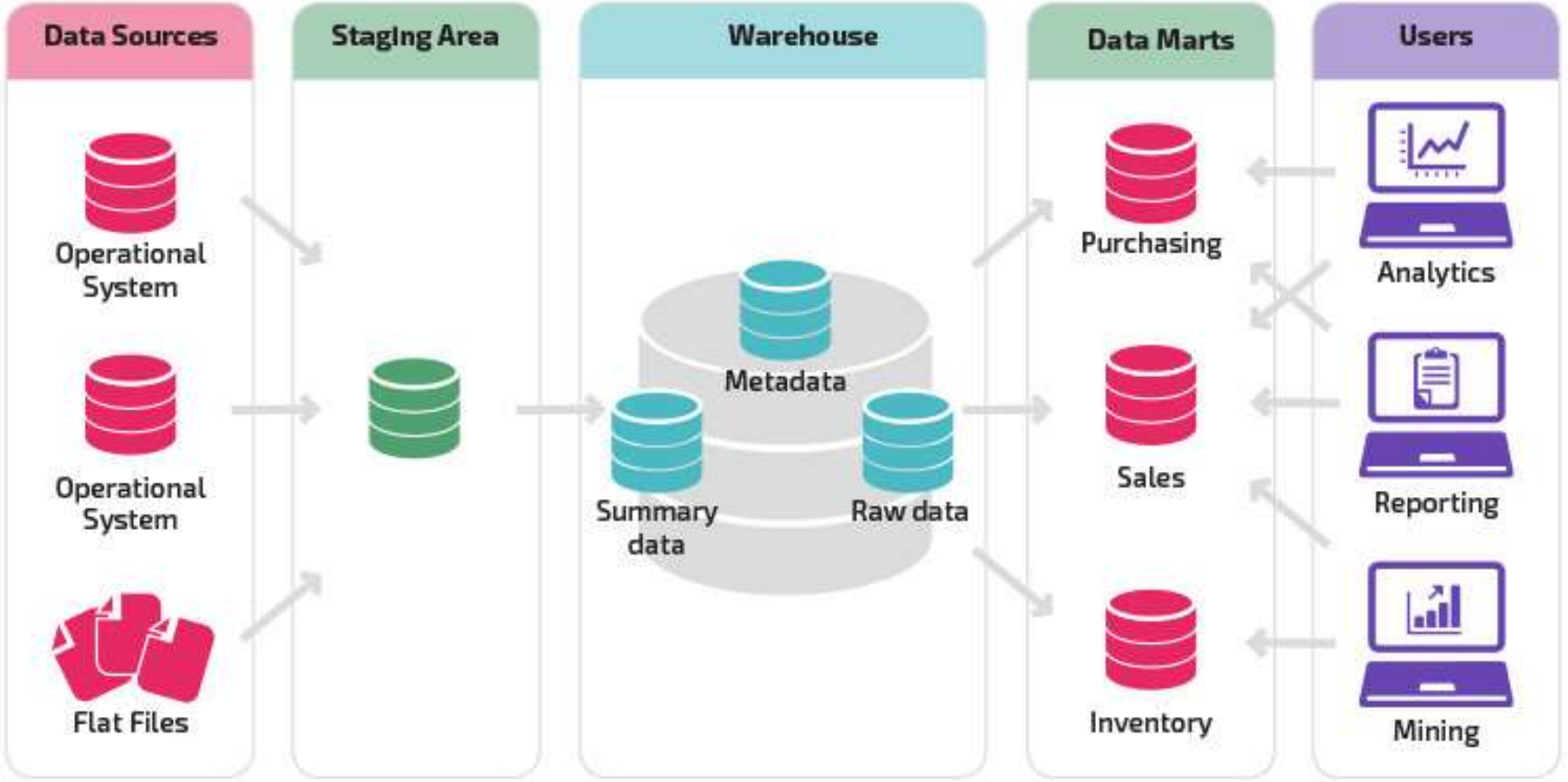


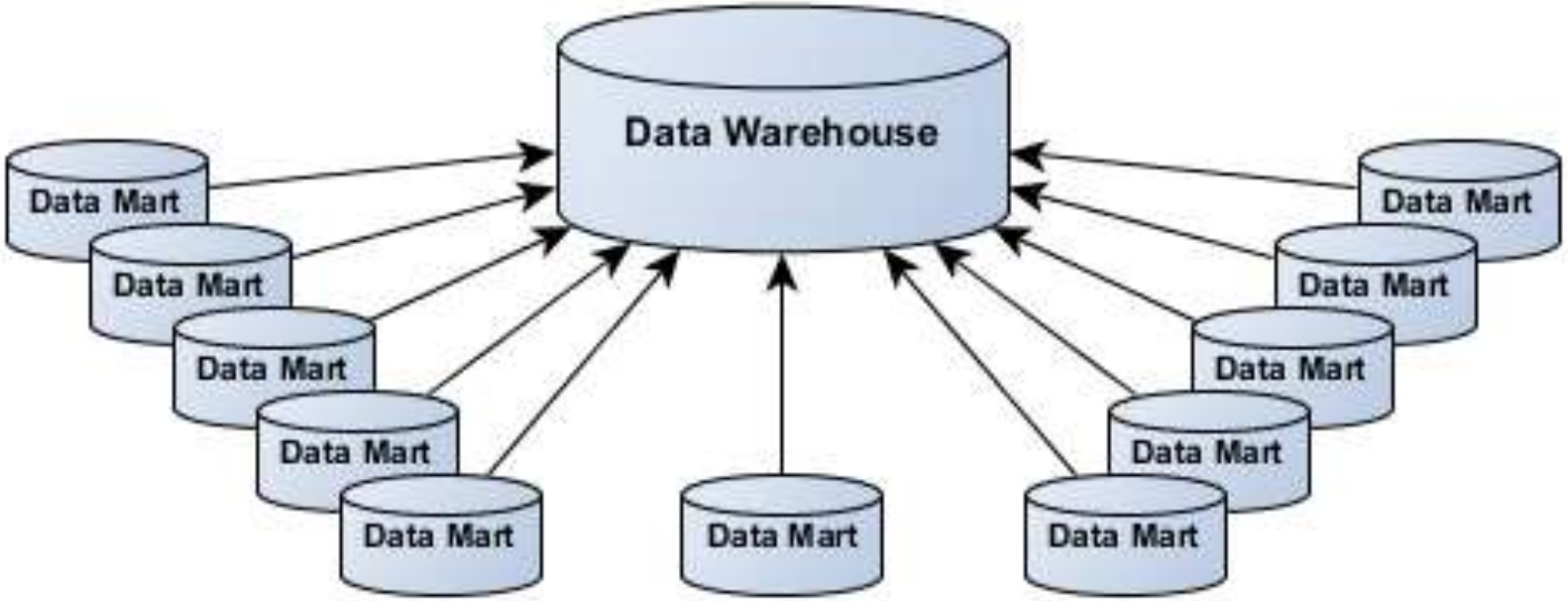
What is Data Warehousing?

A process of **transforming data into information** and making it available to users in a timely enough manner to make a difference



Data Marts


Data Sources
Staging Area
Warehouse
Data Marts
Users
Operational System
Operational System
Flat Files
Metadata
Summary data
Raw data
Purchasing
Sales
Inventory
Analytics
Reporting
Mining


Data Warehouse
Data Mart
Data Mart
Data Mart
Data Mart
Data Mart
Data Mart
Data Mart
Data Mart
Data Mart
Data Mart
Data Mart

Data Mart:

- A data mart is a subset of a data warehouse focused on a particular line of business, department, or subject area.
- Data marts make specific data available to a defined group of users, which allows those users to quickly access critical insights without wasting time searching through an entire data warehouse.

Example:

- Many companies may have a data mart that aligns with a specific department in the business, such as **finance**, **sales**, or **marketing**.

Define Data Mart:

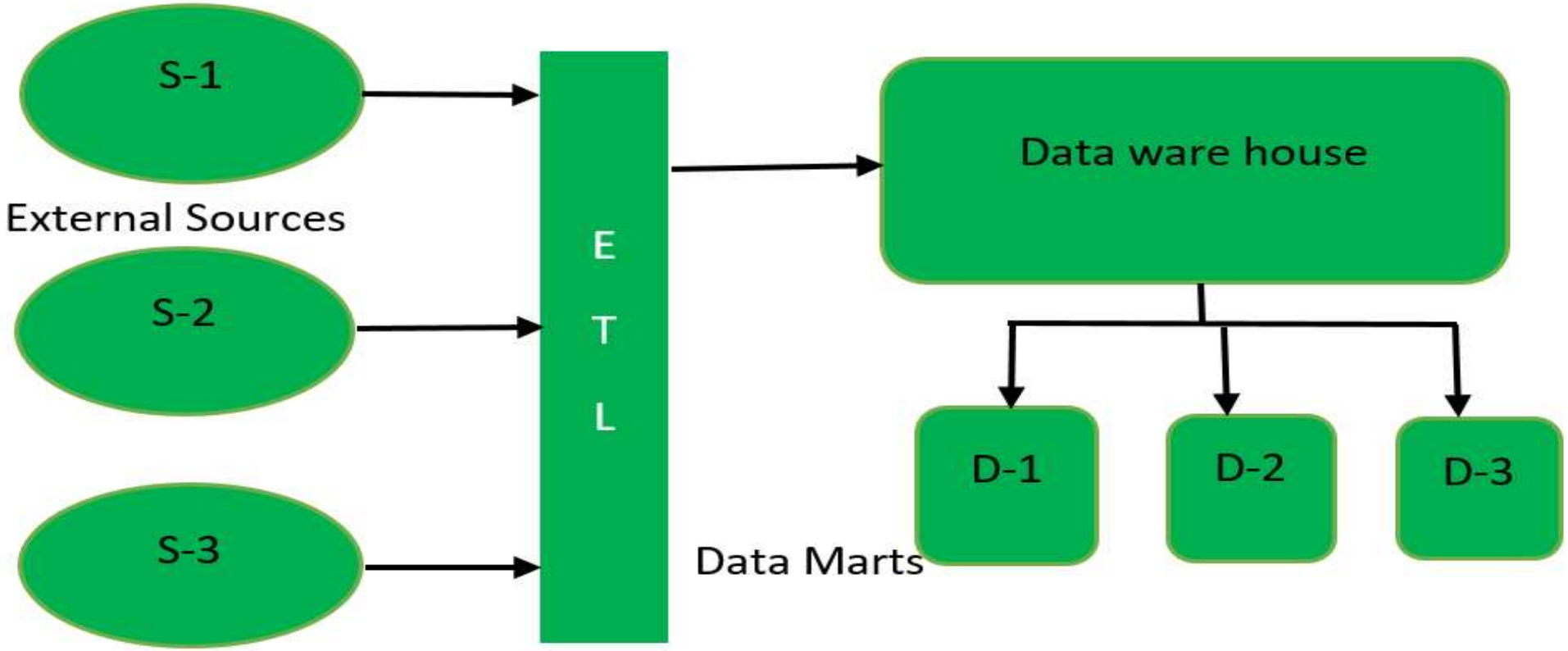
- Data mart is a **storage** component which is concerned on a **specific** department of an organization.
- It is a **subset** of the data stored in the data-warehouse.
- Data mart is focused only on particular function of an organization and it is maintained by single authority only, e.g. finance, Marketing.
- Data Marts are **small** in size and are **flexible**.

Types of Data Mart:

There are three types of data marts:

1. Dependent Data Mart
2. Independent Data Mart
3. Hybrid Data Mart

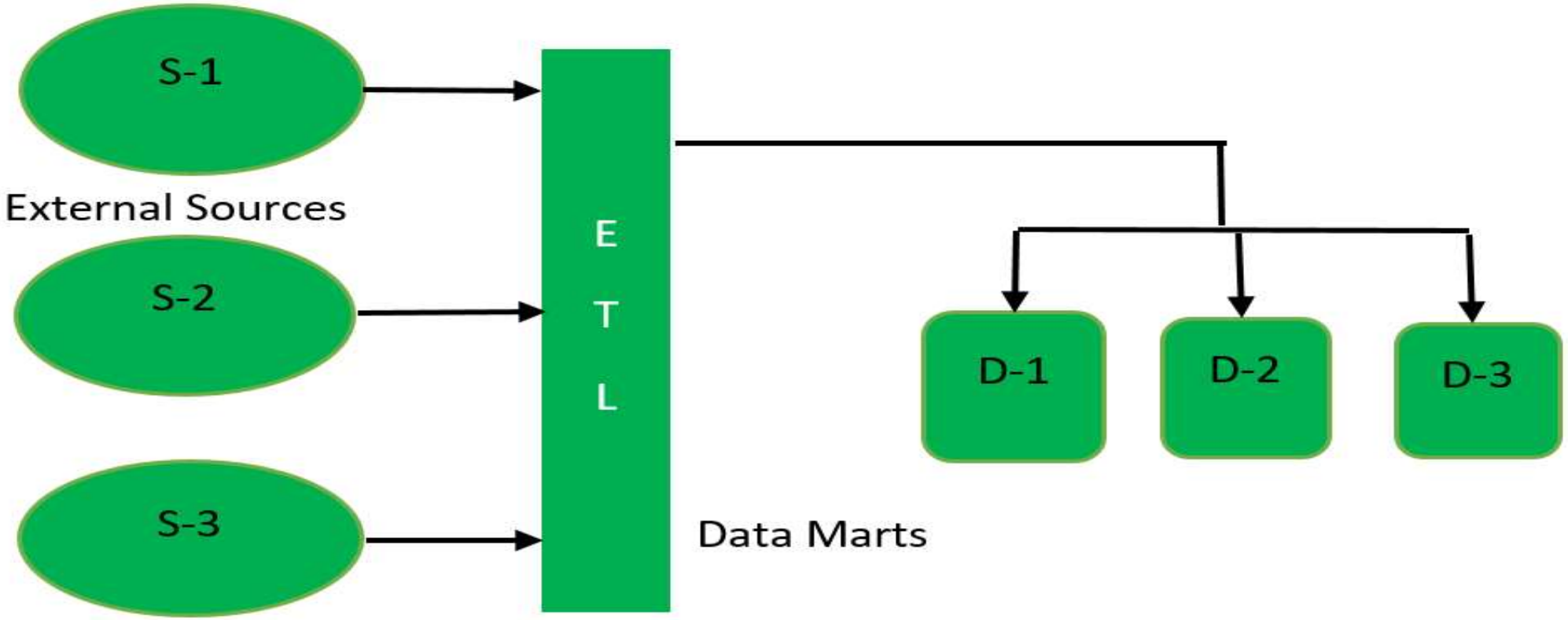
1. Dependent Data Mart :



Dependent Data Mart:

- Dependent Data Mart is created by **extracting** the data from central repository, Datawarehouse.
- First data warehouse is created by extracting data (through ETL tool) from external sources and then data mart is created from data warehouse.
- Dependent data mart is created in top-down approach of data-warehouse architecture.
- This model of data mart is used by big organizations.

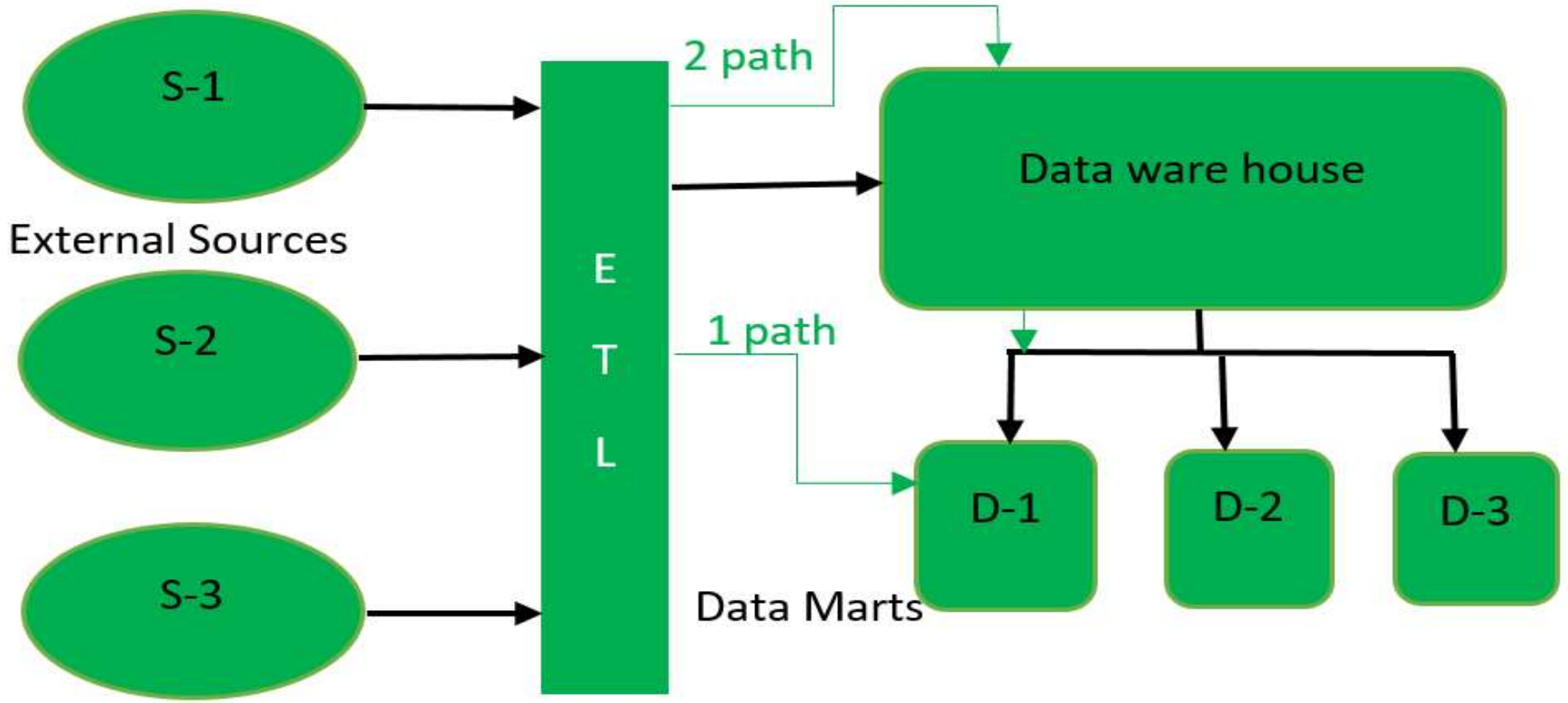
2. Independent Data Mart:



Independent Data Mart:

- Independent Data Mart is created directly from external sources instead of data warehouse.
- First data mart is created by extracting data from external sources and then data-warehouse is created from the data present in data mart.
- Independent data mart is designed in bottom-up approach of data-warehouse architecture.
- This model of data mart is used by small organizations and is cost effective comparatively.

3. Hybrid Data Mart:



Hybrid Data Mart:

- This type of Data Mart is created by extracting data from operational source or from data warehouse.
- **1Path** reflects accessing data directly from external sources and
- **2Path** reflects dependent data model of data mart.

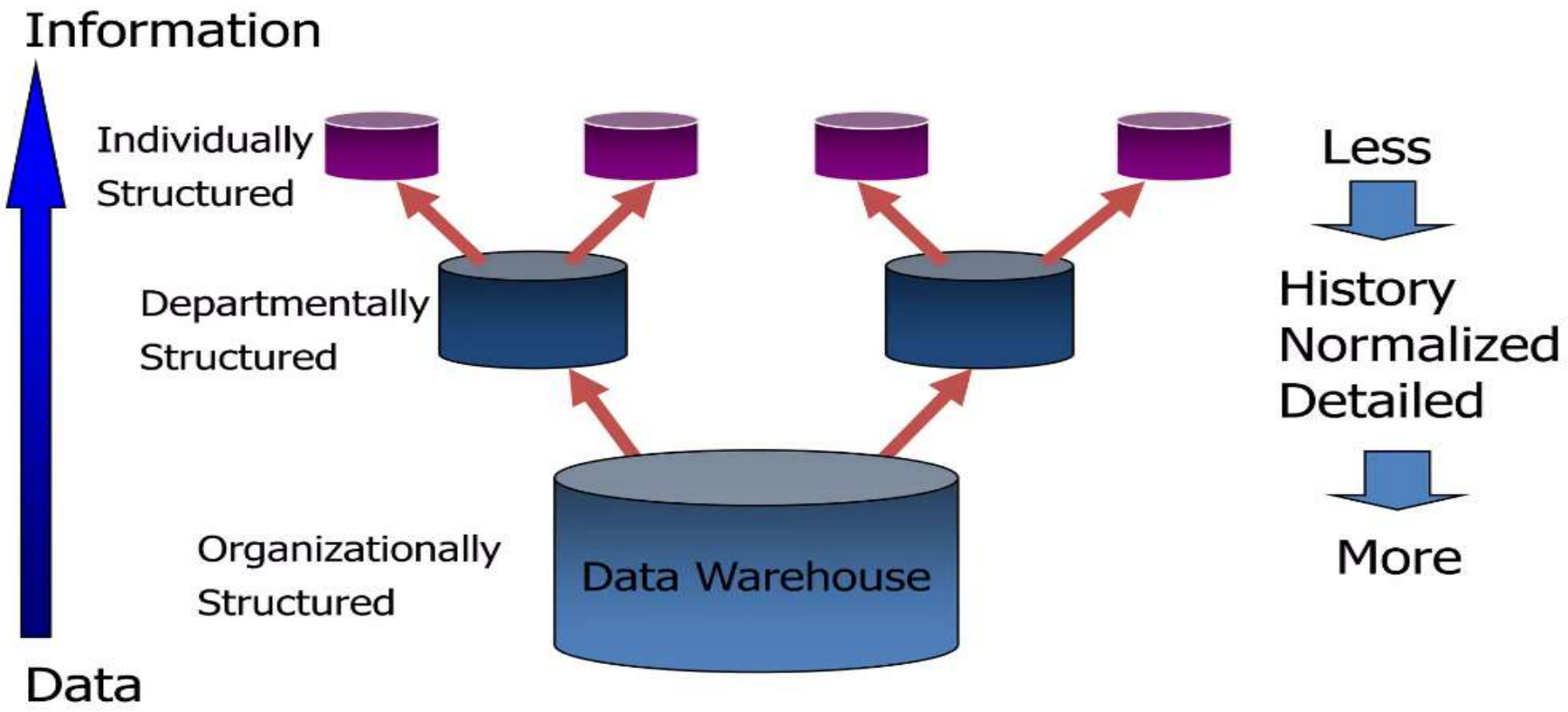
Need Of Data Mart:

- Data Mart focuses only on functioning of particular department of an organization.
- It is maintained by single authority of an organization.
- Since, it stores the data related to specific part of an organisation, data retrieval from it is very quick.
- Designing and maintenance of data mart is found to be quite cinch as compared to data warehouse.
- It reduces the response time of user as it stores small volume of data.
- It is small in size due to which accessing data from it very fast.
- This Storage unit is used by most of organizations for the smooth running of their departments.

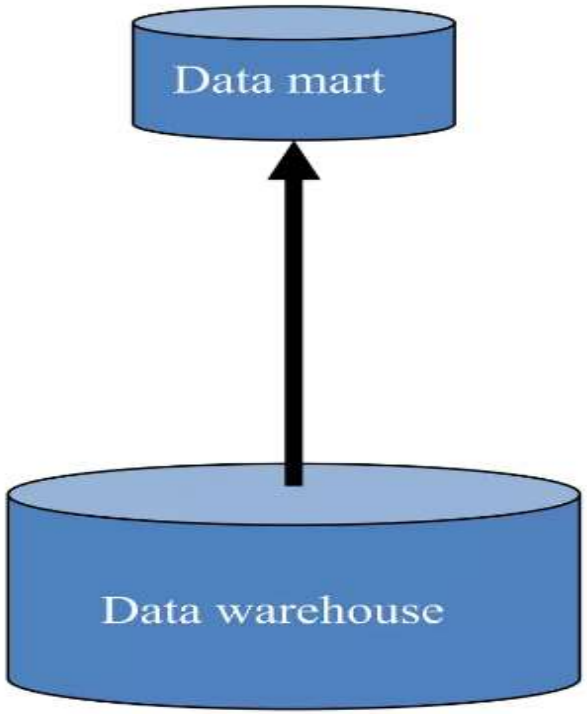
Advantages of Data Mart:

- Implementation of data mart needs less time as compared to implementation of data-warehouse as data mart is designed for a particular department of an organization.
- Organizations are provided with choices to choose model of data mart depending upon cost and their business.
- Data can be easily accessed from data mart.
- It contains frequently accessed queries, so enable to analyse business trend.

From the Data Warehouse to Data Marts

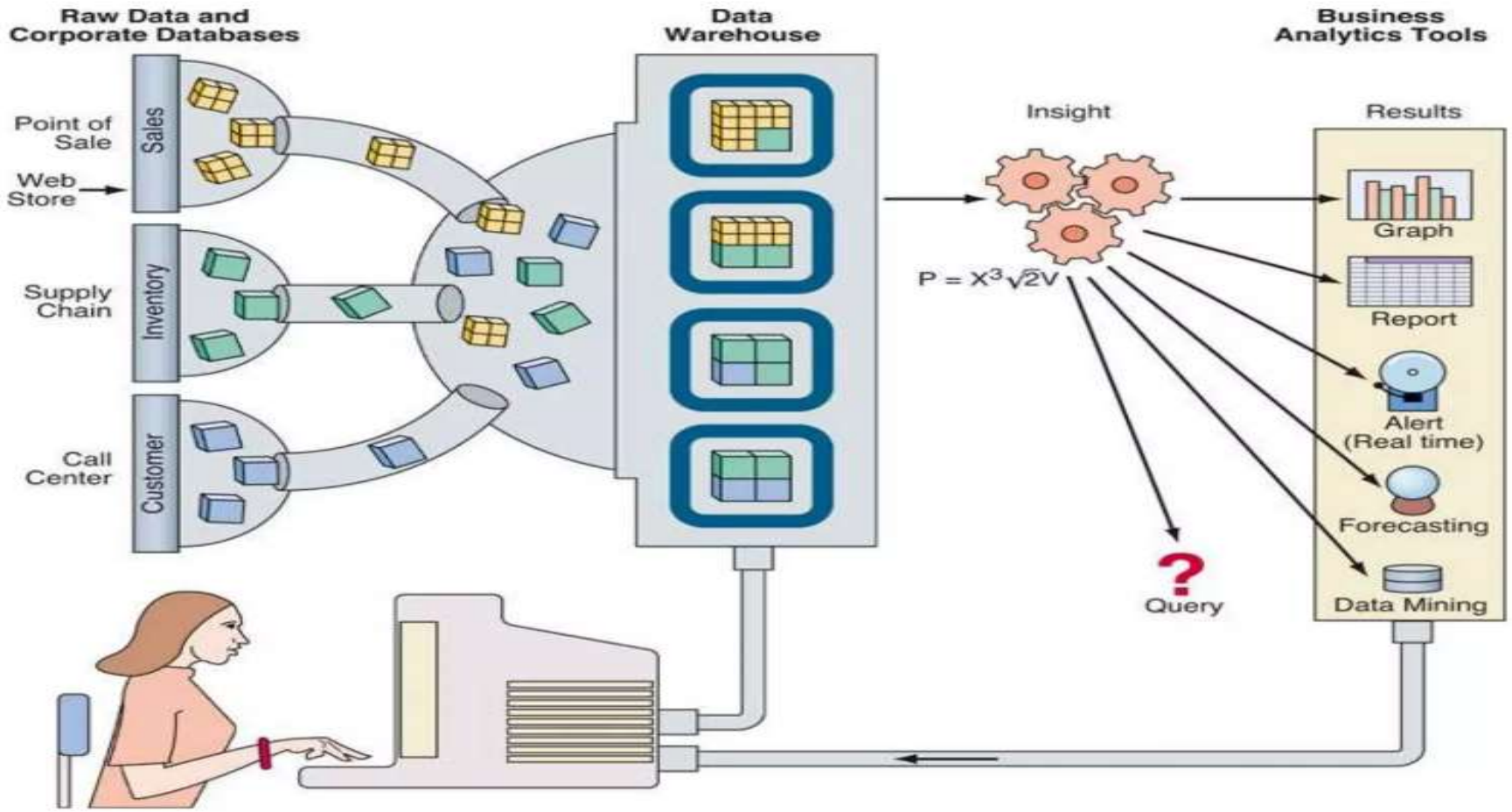


Characteristics of the Departmental Data Mart

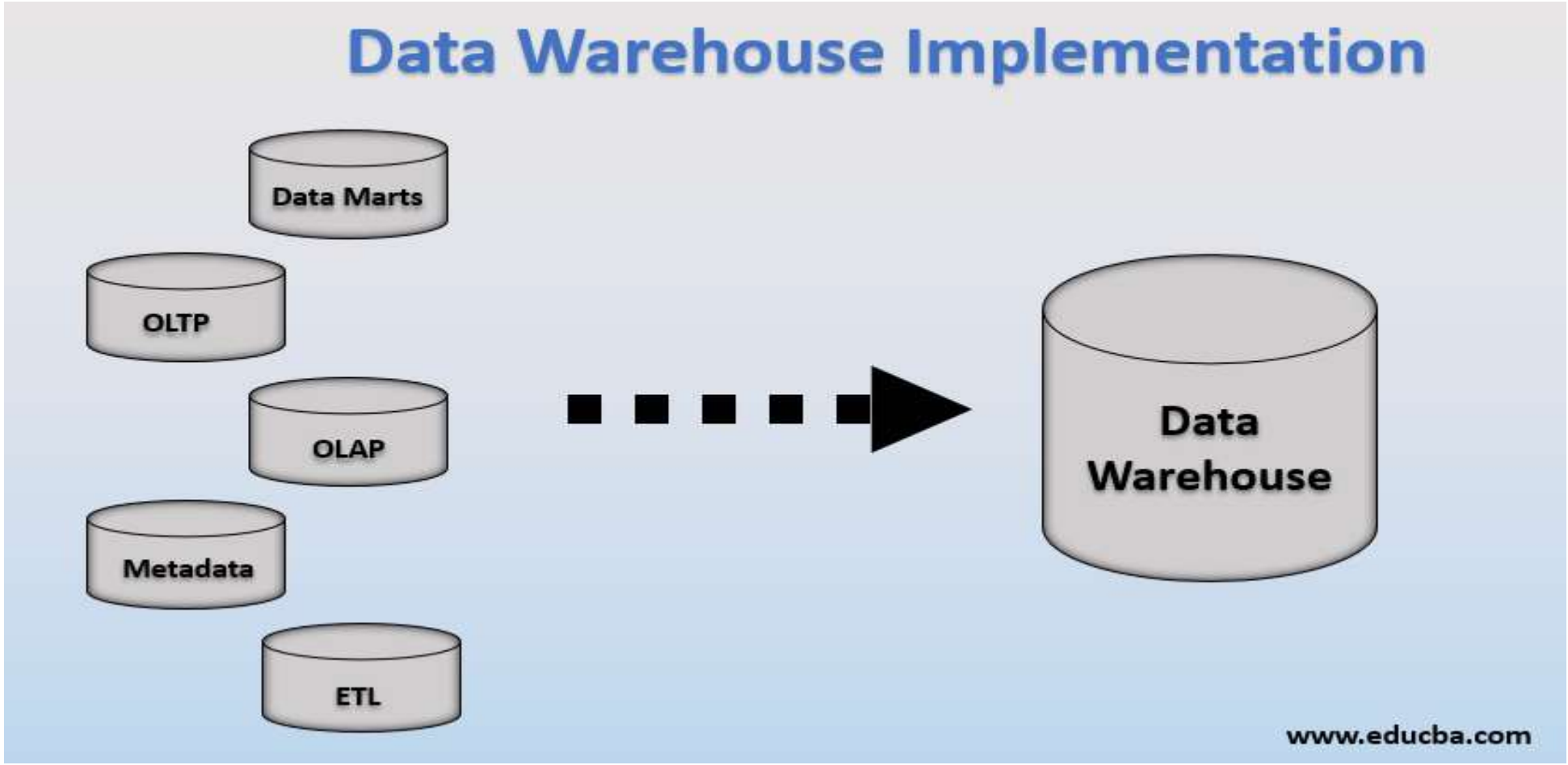


- Small
- Flexible
- Customized by Department
- OLAP
- Source is departmentally structured data warehouse

Business intelligence and data warehousing



- Warehouse Implementation



Data Warehouse Implementation

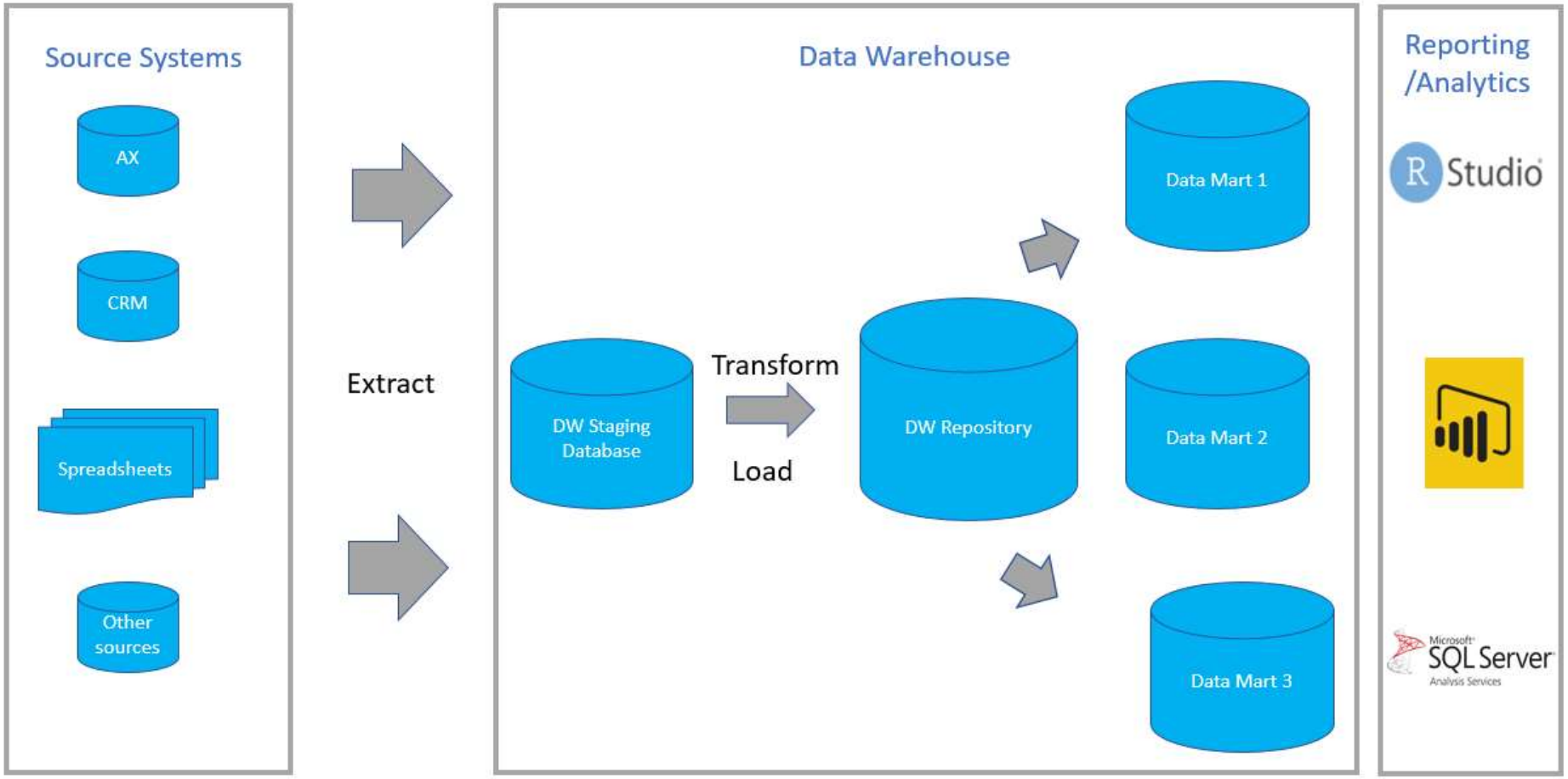
- Data Warehouse Implementation is a series of activities that are essential to create a fully functioning Data Warehouse, after classifying, analyzing and designing the Data Warehouse with respect to the requirements provided by the client.

The various **phases of Data Warehouse Implementation** are

- 'Planning',
- 'Data Gathering',
- 'Data Analysis' and
- 'Business Actions'.

Every Data Warehouse needs a **few important components**, that needs to be defined while designing the implementation of the system, such as

- Data Marts,
- OLTP/ OLAP,
- ETL,
- Metadata, etc.


Source Systems
AX
CRM
Spreadsheets
Other sources
Extract
Data Warehouse
DW Staging Database
Transform
Load
DW Repository
Data Mart 1
Data Mart 2
Data Mart 3
Reporting /Analytics
R Studio
Microsoft SQL Server Analysis Services

Data warehouse implementation

- The process of establishing and implementing a data warehouse system in an organization is known as data warehouse implementation.
- Data warehousing is one of the most important components of the business intelligence process for an organization.
- The data warehousing implementation process requires a series of steps that need to be followed in a very effective manner.

- The processes are

1. Planning

- Planning is one of the most important steps of a process.
- It helps in getting a pathway or the road map that we have to follow to achieve our described goals and objectives.
- It is the cornerstone of every successful project that is implemented in organizations.
- In case of the absence of sound planning, then there are high chances of failure of the project.

2. Data gathering

- As data is available everywhere, but all the data available is not helpful for an organization.
- Data gathering is a process that involves the collection of data from various sources that can be used for data analysis and reporting.
- It involves a wide range of steps and it is a time-consuming process.
- We need to first identify the data that is going to be helpful for organization.

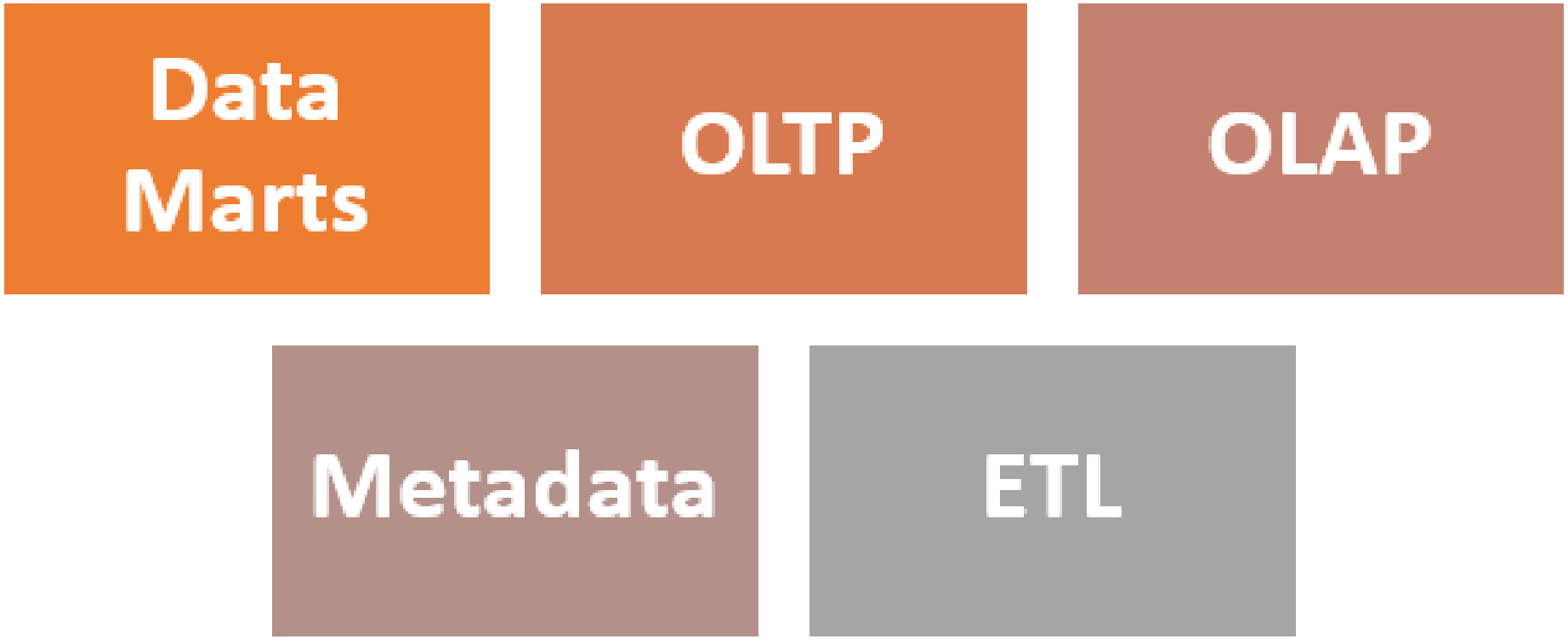
3. Data analysis

- Once the data is collected, the next step which comes into the picture is data analysis.
- The process of generating and getting meaningful insights out of the day together is known as data analysis.

4. Business actions

- The insights and information attained from data analysis are further used for making decisions for the organization.
- Higher the level of insights higher would be the efficiency of the business decisions and these decisions are going to decide the future of the organization.

Components of Data Warehouse Implementation



1. Data Marts

- A data mart is an important component of data warehousing.
- It can be said as the subset of a data warehouse that is focused on a particular Business line like sales, marketing, human resource, etc.

2. OLTP

- The OLTP layer deals with the processing of transactional data on the task associated with an organization.
- It stands for online transactional processing.
- It deals with transactional data which is frequently changing in nature.

3. OLAP

- OLAP layer helps in processing and analyzing the data stored in the database.
- It stands for the online analytical process.
- This layer deals with the master data which is not frequently changing in nature.

4. ETL

- The ETL process helps in fetching the data from different sources into a single data warehouse.
- The process of extraction transformation and loading is used for data warehousing.

5. Metadata

- The data of data is known as metadata.
- It helps in getting granularity of data.
- It helps in getting the information about the data.
- For example, if we have country data, then state data, city data, and the area level can be called the metadata of the data.