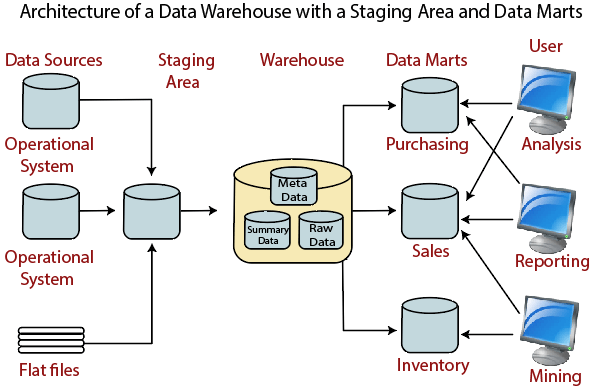
**Data Warehouse Architecture:**

* Data Warehouse Architecture is a structured design that shows how data is collected, stored, and accessed for analysis and decision-making.
* It brings together data from multiple sources, organizes it in a central storage system, and makes it easy to analyzelarge volumes of data.
* The architecture typically includes **data integration tools (ETL), centralized storage, metadata management, and data access layers**, supporting **real-time and batch processing**.
* It ensures **data is accurate, secure, and easy to use**, helping businesses **track past performance, find patterns, and make informed decisions**.

**Data Warehouse Architecture**



**Sources:**

* **Operational System:** An **operational system** is a method used in data warehousing to refer to a **system** that is used to process the day-to-day transactions of an organization.
* **Flat Files:** A **Flat file** system is a system of files in which transactional data is stored, and every file in the system must have a different name.

**Staging Area:**

ETL stands for Extract, Transform, Load and it is a process used in data warehousing to extract data from various sources, transform it into a format suitable for loading into a data warehouse, and then load it into the warehouse. The process of ETL can be broken down into the following three stages:

1. **Extract**: The first stage in the ETL process is to extract data from various sources such as transactional systems, spreadsheets, and flat files. This step involves reading data from the source systems and storing it in a staging area.
2. **Transform**: In this stage, the extracted data is transformed into a format that is suitable for loading into the data warehouse. This may involve cleaning and validating the data, converting data types, combining data from multiple sources, and creating new data fields.
3. **Load**: After the data is transformed, it is loaded into the data warehouse. This step involves creating the physical data structures and loading the data into the warehouse.

**Data Warehouse:**

* A data warehouse is a centralized repository for storing and managing large amounts of data from various sources for analysis and reporting.
* It is optimized for fast querying and analysis, enabling organizations to make informed decisions by providing a single source of truth for data.

**Data Marts:**

* A data mart is a specialized storage component focused on a specific department or function within an organization, such as Finance or Marketing. It is a subset of the data warehouse, maintained by a single authority
* A data mart is a specialized subset of a data warehouse focused on a specific functional area or department within an organization.
* Data marts are smaller in scale and scope, typically holding relevant data for a specific group of users, such as sales, marketing, or finance.

**Users:** Allows end-users to perform data analysis, create reports, and generate insights.

* Data mining is the process of discovering patterns, correlations, and trends within large datasets using statistical and machine learning techniques. (Microsoft Power BI, Tableau)
* Data analysis involves examining, processing, and interpreting stored data to extract meaningful insights and support strategic business decisions.
* Reporting involves presenting data in an organized and visually appealing format, such as dashboards, charts, graphs, and static or dynamic reports. (RapidMiner, Weka)

**Advantages of Data Warehouse Architecture:**

* Improved query performance
* Historical data analysis
* Enhanced decision-making
* Data consistency and accuracy