**SQL Server 2022**

Installtion Process

1. Download(Developer Edition) Setup from below linl and istall
   1. <https://www.microsoft.com/en-us/sql-server/sql-server-downloads>
   2. Custome
   3. Install
   4. After installtion one popup page will open (SQl Server Installtion Center)
      1. Tab on Installtion
         1. Select / Tab -> New SQl Server Standalone installtion or add feacture to existing installtion
            1. Developer
            2. Uncheck Azure Extention for SQl Server => Next
            3. Check =>

Default Engine Server

SQL Server Replication

Next

* + - * 1. Chnages and add the Instance Name => Next
        2. SQL Server Broser => Enable
      1. Tab / Click On install SQL Server Management tool
         1. Install

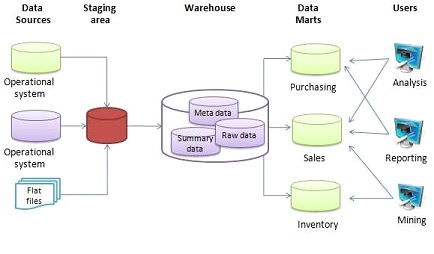
**Class: 2**

Data Warehouse:

It is colletion of huge ammount of data which are colected from different-2 sources / databases / file and that bulk of data are storeed in one place which can be queried for business benefits.

**Data warehouse is an application for storing and analyzing data and reporting. It is designed for query & analysis. It usually contains historical data derived from transtion data, and can aslo include daat from others sources**

Data Warehouse environment contains an extraction, transportation, and loading (ETL) solution, an online analytical processing (OLAP) engine, customer analysis tools, and other applications that handle the process of gathering information and delivering it to business users.



What is a Data Warehouse?

A Data Warehouse (DW) is a relational database that is designed for query and analysis rather than transaction processing. It includes historical data derived from transaction data from single and multiple sources.

A Data Warehouse provides integrated, enterprise-wide, historical data and focuses on providing support for decision-makers for data modeling and analysis.

A Data Warehouse is a group of data specific to the entire organization, not only to a particular group of users.

It is not used for daily operations and transaction processing but used for making decisions.

A Data Warehouse can be viewed as a data system with the following attributes:

* It is a database designed for investigative tasks, using data from various applications.
* It supports a relatively small number of clients with relatively long interactions.
* It includes current and historical data to provide a historical perspective of information.
* Its usage is read-intensive.
* It contains a few large tables.

"Data Warehouse is a subject-oriented, integrated, and time-variant store of information in support of management's decisions

## Characteristics of Data Warehouse

## Subject-Oriented: Used to analyze a particular subject area data.

A data warehouse target on the modeling and analysis of data for decision-makers. Therefore, data warehouses typically provide a concise and straightforward view around a particular subject, such as customer, product, or sales, instead of the global organization's ongoing operations. This is done by excluding data that are not useful concerning the subject and including all data needed by the users to understand the subject.

## Integrated : It integrate the data from multiple data sources

A data warehouse integrates various heterogeneous data sources like RDBMS, flat files, and online transaction records. It requires performing data cleaning and integration during data warehousing to ensure consistency in naming conventions, attributes types, etc., among different data sources

## Time-Variant: Historical data / accesing time / uses keep in warehouses.

Historical information is kept in a data warehouse. For example, one can retrieve files from 3 months, 6 months, 12 months, or even previous data from a data warehouse. These variations with a transactions system, where often only the most current file is kept.

## Non-Volatile: Once data in warehouse, it can not be changes.

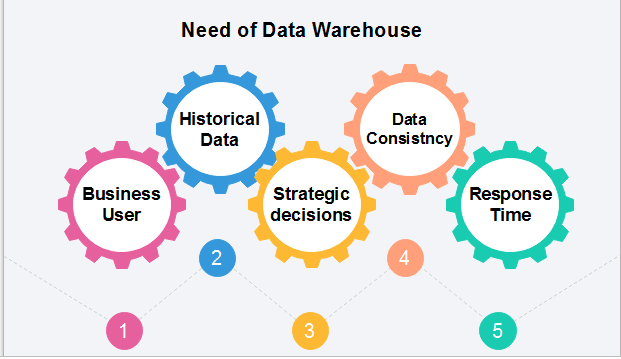
The data warehouse is a physically separate data storage, which is transformed from the source operational RDBMS. The operational updates of data do not occur in the data warehouse, i.e., update, insert, and delete operations are not performed. It usually requires only two procedures in data accessing: Initial loading of data and access to data. Therefore, the DW does not require transaction processing, recovery, and concurrency capabilities, which allows for substantial speedup of data retrieval. Non-Volatile defines that once entered into the warehouse, and data should not change.

Goals of Data Warehousing

* To help reporting as well as analysis
* Maintain the organization's historical information
* Be the foundation for decision making.

## Need for Data Warehouse

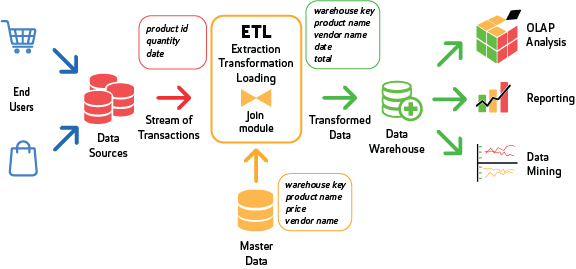
Data Warehouse is needed for the following reasons:



1. 1) **Business User:** Business users require a data warehouse to view summarized data from the past. Since these people are non-technical, the data may be presented to them in an elementary form.
2. 2) **Store historical data:** Data Warehouse is required to store the time variable data from the past. This input is made to be used for various purposes.
3. 3) **Make strategic decisions:** Some strategies may be depending upon the data in the data warehouse. So, data warehouse contributes to making strategic decisions.
4. 4) **For data consistency and quality:** Bringing the data from different sources at a commonplace, the user can effectively undertake to bring the uniformity and consistency in data.
5. 5) **High response time:** Data warehouse has to be ready for somewhat unexpected loads and types of queries, which demands a significant degree of flexibility and quick response time.

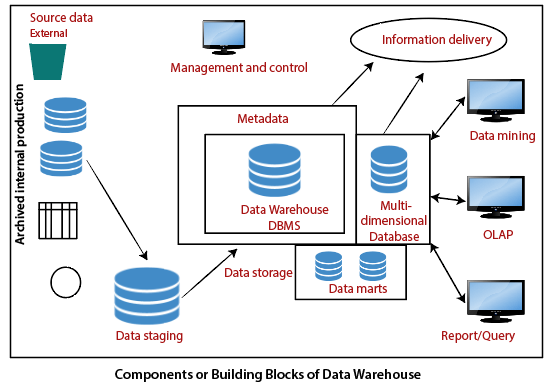
Benefits of Data Warehouse

1. Understand business trends and make better forecasting decisions.
2. Data Warehouses are designed to perform well enormous amounts of data.
3. The structure of data warehouses is more accessible for end-users to navigate, understand, and query.
4. Queries that would be complex in many normalized databases could be easier to build and maintain in data warehouses.
5. Data warehousing is an efficient method to manage demand for lots of information from lots of users.
6. Data warehousing provide the capabilities to analyze a large amount of historical data.



# Components or Building Blocks of Data Warehouse

Architecture is the proper arrangement of the elements. We build a data warehouse with software and hardware components. To suit the requirements of our organizations, we arrange these building we may want to boost up another part with extra tools and services. All of these depends on our circumstances.



The figure shows the essential elements of a typical warehouse. We see the Source Data component shows on the left. The Data staging element serves as the next building block. In the middle, we see the Data Storage component that handles the data warehouses data. This element not only stores and manages the data; it also keeps track of data using the metadata repository. The Information Delivery component shows on the right consists of all the different ways of making the information from the data warehouses available to the users.

## Source Data Component

Source data coming into the data warehouses may be grouped into four broad categories:

**Production Data:** This type of data comes from the different operating systems of the enterprise. Based on the data requirements in the data warehouse, we choose segments of the data from the various operational modes.

**Internal Data:** In each organization, the client keeps their "**private**" spreadsheets, reports, customer profiles, and sometimes even department databases. This is the internal data, part of which could be useful in a data warehouse.

**Archived Data:** Operational systems are mainly intended to run the current business. In every operational system, we periodically take the old data and store it in achieved files.

**External Data:** Most executives depend on information from external sources for a large percentage of the information they use. They use statistics associating to their industry produced by the external department

## Data Staging Component

**1) Data Extraction:** This method has to deal with numerous data sources. We have to employ the appropriate techniques for each data source.

**2) Data Transformation:** As we know, data for a data warehouse comes from many different sources. If data extraction for a data warehouse posture big challenges, data transformation present even significant challenges. We perform several individual tasks as part of data transformation.

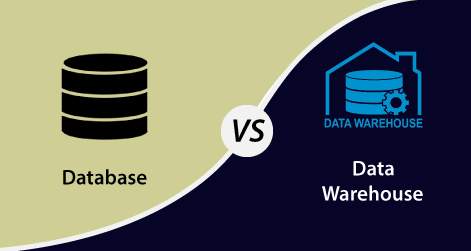
First, we clean the data extracted from each source. Cleaning may be the correction of misspellings or may deal with providing default values for missing data elements, or elimination of duplicates when we bring in the same data from various source systems.

Standardization of data components forms a large part of data transformation. Data transformation contains many forms of combining pieces of data from different sources. We combine data from single source record or related data parts from many source records.

On the other hand, data transformation also contains purging source data that is not useful and separating outsource records into new combinations. Sorting and merging of data take place on a large scale in the data staging area. When the data transformation function ends, we have a collection of integrated data that is cleaned, standardized, and summarized.

**3) Data Loading:** Two distinct categories of tasks form data loading functions. When we complete the structure and construction of the data warehouse and go live for the first time, we do the initial loading of the information into the data warehouse storage. The initial load moves high volumes of data using up a substantial amount of time.

Difference between Database and Data Warehouse



|  |  |
| --- | --- |
| **Database** | **Data Warehouse** |
| 1. It is used for Online Transactional Processing (OLTP) but can be used for other objectives such as Data Warehousing. This records the data from the clients for history. | 1. It is used for Online Analytical Processing (OLAP). This reads the historical information for the customers for business decisions. |
| 2. The tables and joins are complicated since they are normalized for RDBMS. This is done to reduce redundant files and to save storage space. | 2. The tables and joins are accessible since they are de-normalized. This is done to minimize the response time for analytical queries. |
| 3. Data is dynamic | 3. Data is largely static |
| 4. **Entity:** Relational modeling procedures are used for RDBMS database design. | 4. **Data:** Modeling approach are used for the Data Warehouse design. |
| 5. Optimized for write operations. | 5. Optimized for read operations. |
| 6. Performance is low for analysis queries. | 6. High performance for analytical queries. |
| 7. The database is the place where the data is taken as a base and managed to get available fast and efficient access. | 7. Data Warehouse is the place where the application data is handled for analysis and reporting objectives. |