**ETL PROCESS:**

ETL, which stands for extract, transform and load, is a **data integration process** that combines data from multiple data sources into a single, consistent data store that is loaded into a data warehouse or other target system.

RDBMS

**RDB**

DATA WAREHOUSE

STAGING AREA

FLAT FILES

SQL SERVER

**TRANSFORMATION** **LOADING**

**EXTRACT**

**EXTRACTION:**

Data from various source systems is extracted which can be in various formats like relational databases, No SQL, XML, and flat files into the staging area.

**TRANSFORMATION:**

A set of rules or functions are applied on the extracted data to convert it into a single standard format. It may uses processes like filtering,cleaning,joining,sliptting,sorting.

**LOADING:**

The transformed data is finally loaded into the data warehouse. Sometimes the data is updated by loading into the data warehouse very frequently and sometimes it is done after longer but regular intervals. The rate and period of loading solely depends on the requirements and varies from system to system.

**ETL Tools:** Most commonly used ETL tools are Sybase, Oracle Warehouse builder, CloverETL, and MarkLogic.

* **ETL** process can also use the pipelining concept i.e. as soon as some data is extracted, it can transformed and during that period some new data can be extracted. And while the transformed data is being loaded into the data warehouse, the already extracted data can be transformed.

EXTRACT

LOAD

EXTRACT

TRANSFORM

EXTRACT

TRANSFORM

LOAD

LOAD

TRANSFORM