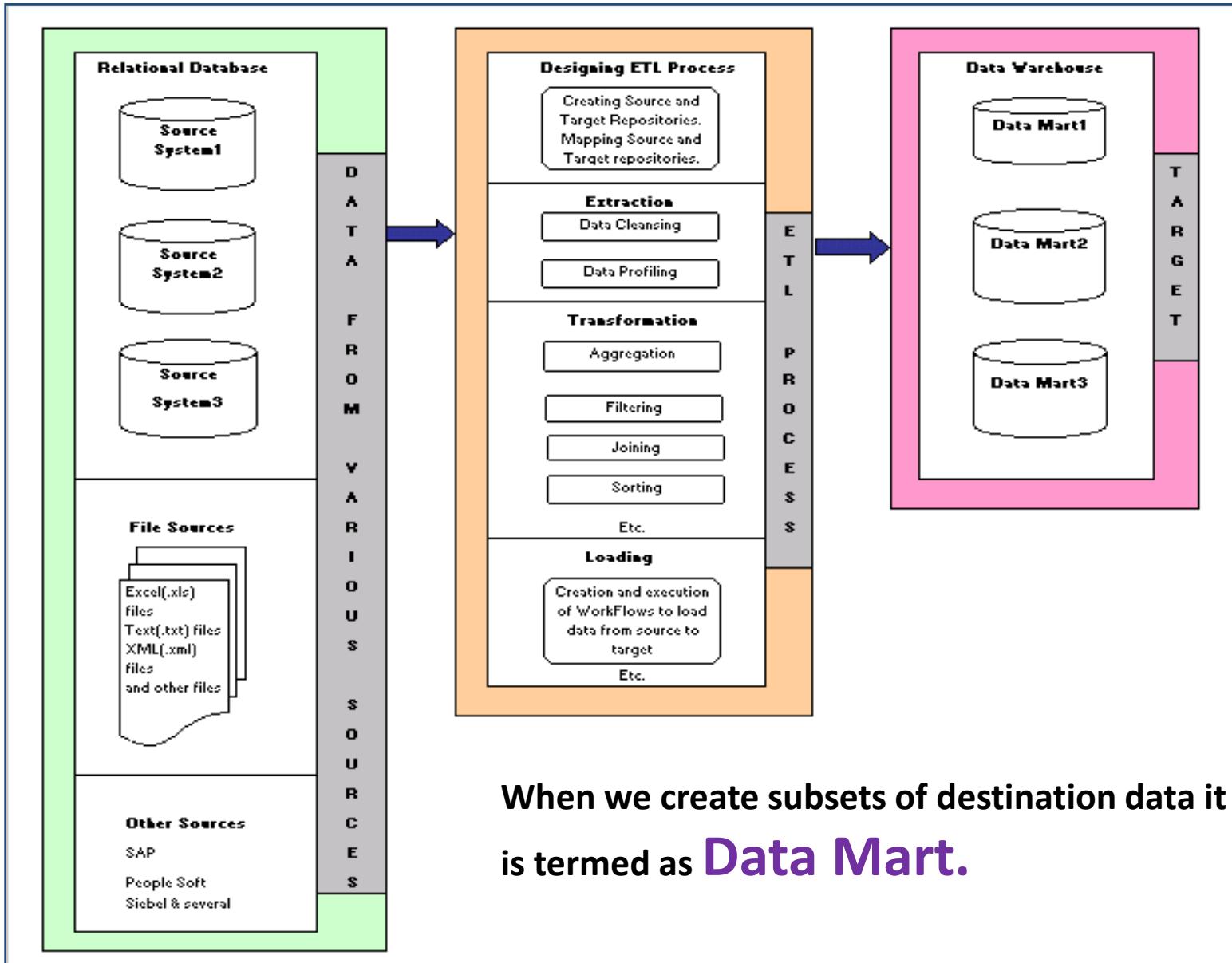


ETL

ETL

ETL Tools are meant to Extract,
Transform and Load the data into Data
Warehouse for decision making.

ETL Process



Designing ETL Process

Source & Destination

- **Creating Sources & Destinations**
 - OLE DB
 - Flat File
 - Excel
 - XML
- **Mapping Sources & Destinations**

Through Transformation Layer

 - Derived Column
 - Conditional Split
 - Aggregation

Extraction

Data Cleansing

- It is the process of amending or removing data in a database that is **incorrect, incomplete, improperly formatted, or duplicated.**

Transformation

Aggregation

- To calculate sum, max, min, avg, count from the data source for different columns and transfer summarized values to destination.

Filtering

- It is like the “**where**” clause of SQL statement.
- From Data Source as per the requirement transferring conditional records to the destination.

Joining

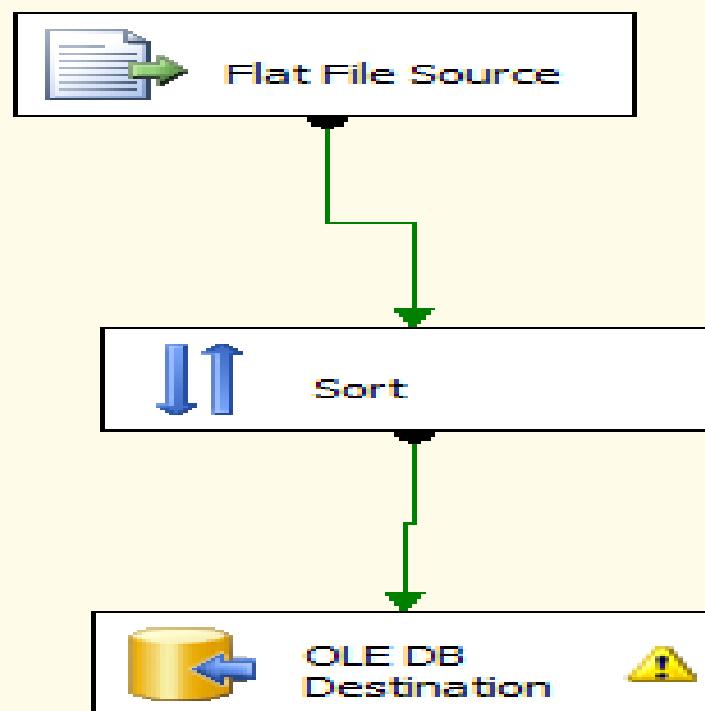
- Combining data from multiple sources and transferring into a destination.

Sorting

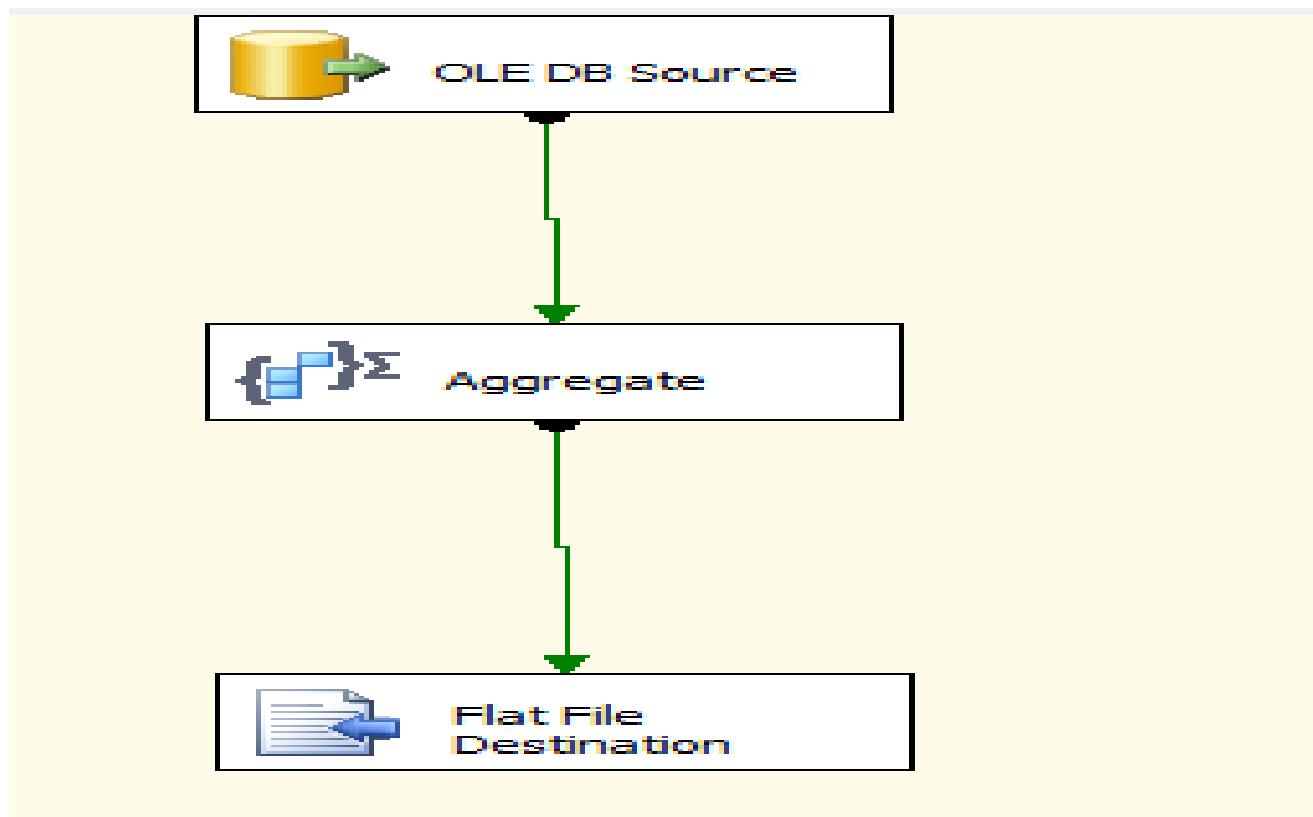
Rearranging rows from the source (ascending or descending) as sending to the destination.

Some examples of ETL

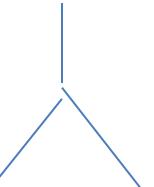
E.g. 1 Transformation from text files to Databases



E.g. 2 Transformation from databases to text files



Designs of ETL

1. I design → One Source and One Destination
2. Y Design → Multiple sources and One Destination.
3.  Design → One Source & Multiple Destinations
4. X Design → Multiple sources & Multiple Destinations

Summary

The process of taking data from different sources, converting it into expected format and sending to the Data Warehouse location is ETL

ETL Tools

Tool Name	Company Name
Ab Initio	Ab Initio Software Corporation
Data Stage	IBM
Informatica	Informatica Corporation
SQL Server Integration Services (SSIS)	Microsoft
Oracle Data Integrator	Oracle Corporation
Alteryx	Alteryx
Azure Data Factory	Microsoft

And there are many more.....