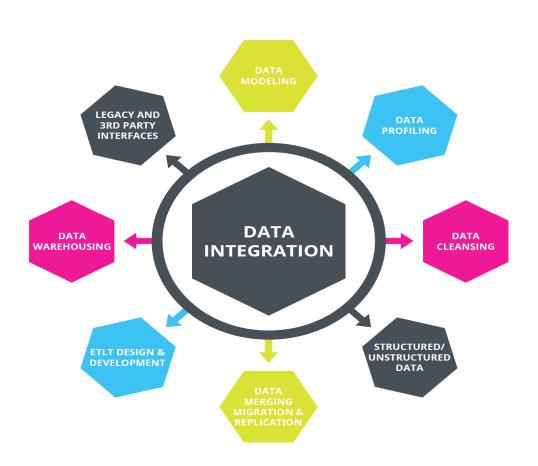
- Every company these days have to process large sets of data from varied sources.
- This data needs to be processed to give meaningful information for making business decisions



• Data Integration is the process in which heterogeneous data is retrieved and combined as an incorporated form and structure



Data Modeling

 First create a model and then perform operations on data

Data Profiling

 Take sample data and check for inconsistency / errors / some variations in data.

Data Cleansing

It is the process of detecting and correcting (or removing) corrupt or inaccurate records from a record set, table, or **database** and refers to identifying incomplete, incorrect, inaccurate or irrelevant parts of the **data** and then replacing, modifying, or deleting the dirty or coarse **data**.

• ETL

Extract, Transform and Load

Data Warehousing

Data warehousing is the process of constructing and using a data warehouse. A data warehouse is constructed by integrating data from multiple heterogeneous sources that support analytical reporting, structured and/or ad hoc queries, and decision making.



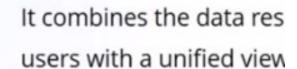
Some of the well known integration tools are:





→ Data can be loaded in parallel to many varie ightarrow SSIS removes the need of hard core program → Tight integration with other products of Micro • SQL Server Integration Service (SSIS) is a component of the Microsoft SQL Server Database Software that can be used to perform a broad range of Data Integration and Data Transformation tasks.







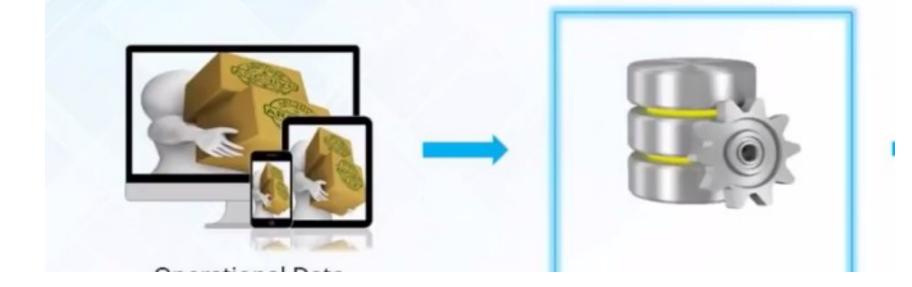
• The task of Data Transformation and Workflow creation will be carried out using SSIS Package.



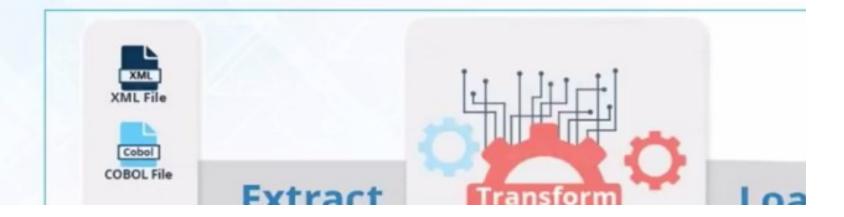


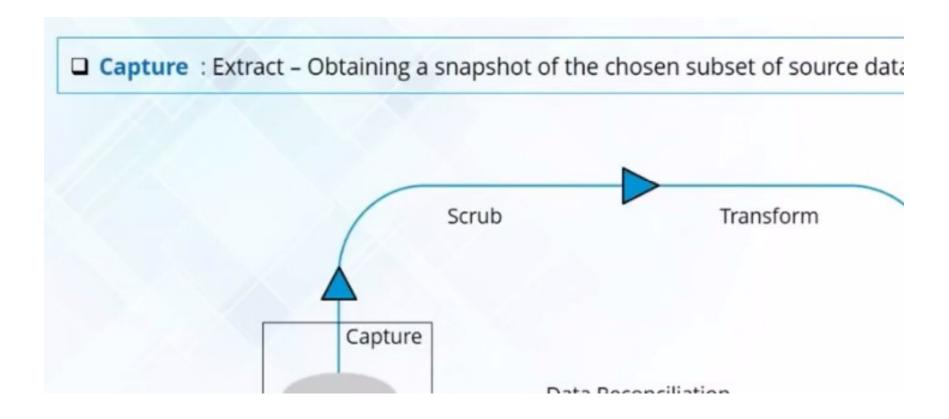
- An operational data store (ODS) is a to integrate data from multiple source operations on the data.
- · This is the place where most of the da

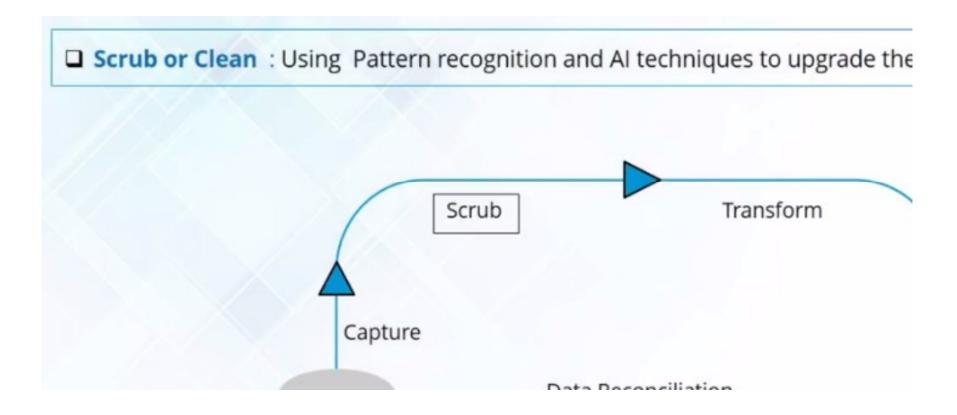
ETL is a process to Extract, Transform and Load t



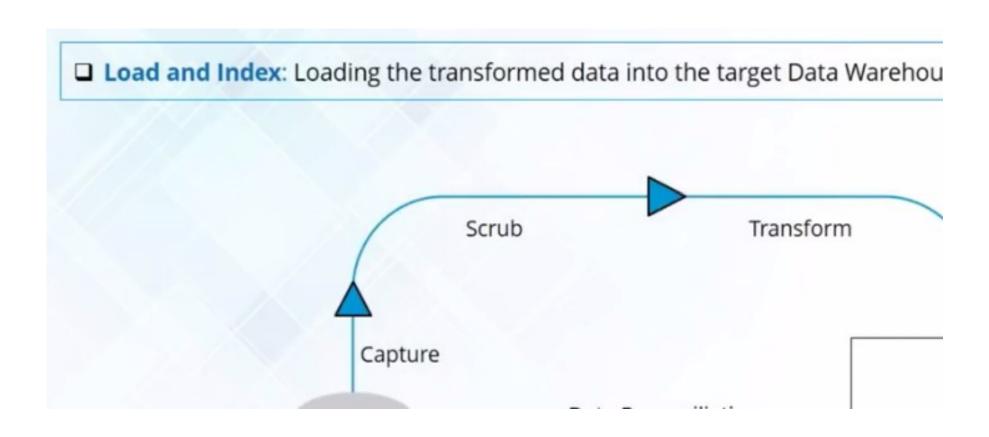
Extract, Transform and Load (ETL) is the process of extracting the data from vato meet your requirement and then loading into a target data warehouse.







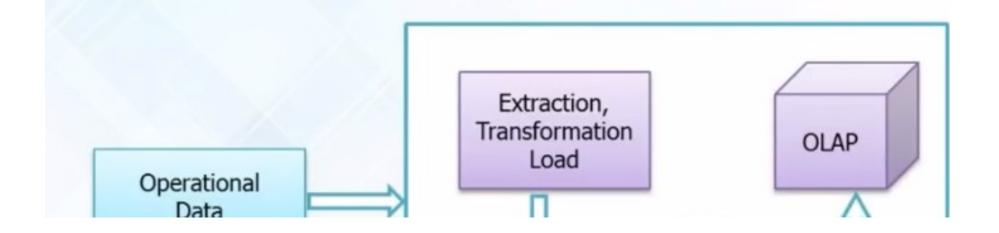
☐ Transformation: Converting the data format from that of the Source System Scrub Transform Capture Data Reconciliation



Datawarehouse captures the data from diverse sources for use



Datawarehousing is a large set of data accumulated which is used for assembling sources for the purpose of answering business questions. Hence, helps in making



- Download SQL server 2014 or any other latest edition from Microsoft site.
- Download SQL Server Data Tools (SSDT) which is also available on Microsoft Site.

It is an object that implements Integration Services functionality to extract, tr is composed of:

- Connections
- · Control flow elements
- · Data flow elements

Table 2.1 Control-Flow

Machine Model	(control-driven)
Basic Definition	Conventional computation; token of control indicates when a statement should be executed
Advantages	Full control
	Complex data and control structures are easily implemented



- Create a comma separated (CSV) file using notepad++ (You can use any other source also) Save this file as normal txt file.
- Open SQL Server Data Tools from your start menu
- Start -> new -> project
- Select Integrated Service Project from the screen and give name to your project

- Now Start SQL Server database
- Start -> SQL server -> SQL server
 Management Studio
- Click on Connect
- Databases -> New DataBase
- Give name to your DataBase
- Select your DataBase and Create a new Table with required format

- Save the Table
- Give StudentId as primarykey
- Go back to SSIS Environment
- In control flow window drag and drop the Data Flow Task from SSIS Toolbox
- Give name to DataFlowTask
- Double click on the DataFlowTask component so it will open DataFlow

- Now Click on Other sources and select Flat File
- Right click on component and click edit
- Configure the component by selecting your CSV file (already created)
- See that the component has two arrows (One Red and One Blue)

- Now Click on Other Destinations and select ADO NET Destination
- Give name to it
- Join Blue Arrow from Flat file to Destination
- Right Click on ADO NET Destination component and click edit
- Click on new -> new -> and copy the server of SQL from SQL Server manager -> File -> Connect Object Explorer

- Select Database
- Click on Test Connection
- Select the table from given screen
- Go to mappings and Map columns
- Click ok
- Now Click Start.
- So, it will transfer data from CSV file to DataBase

- Through this we have completed Extraction and Loading (EL) Part
- Now we will do some Transformation so that ETL completes
- For that first delete the connection
- From common select Derived Column and drop into DataFlow

- Make connection from Flat File to Derived Column to Destination
- Right click on Derived column and create a new column for capital names
- Right click on Destination and do mapping again
- Go to SQL Studio and check all names are now in Capital

- There are two ways to use Scripts in SSIS
 - Through Script Task
 - Through Script Component through Data Flow

Script Task

- First Take For Loop Container From Tool Box
- Second Take Executable task
- Double click on executable task
- Select connection
- SQL Statement insert into table values ("abc");
- OK
- Right Click -> Variables (count (10) and incr)-> close
- Double click on Forloopcontainer
- initExpression > @inc=o , EvalExpression: @inc<@count , AssignExpression-> @inc=@inc+1

- Create a directory -> Files
- Create another directory > Backup
- Start Visual Studio
- Create Variables
- FilePath-> String-> E:\Files\MyFile.txt
- DirPath->String->E:\Files\MyDir
- FileNewPath->String->E:\Files\MyDir\MyFile.txt
- Take ScriptTask (Double Click)
- ReadOnlyVariable-> Select User Three Variable
- Copy Path -> EditScript

- Add System.io Name space
- In Main Method:
- String DirPath=Dts.Variables["DirPath"].Value.toString();
- String FileNewPath=Dts.Variables["FileNewPath"].Value.ToString();
- String FilePath = Dts.Variables["FilePath"].Value.toString();
- String Content = "This is my text inside the flat file";
- File.WriteAllText(FilePath,Content);
- If(!Directory.Exists(DirPath))
- Directory.createDirectory(DirPath);
- File.move(FilePath, FileNewPath);
- Close this and ok

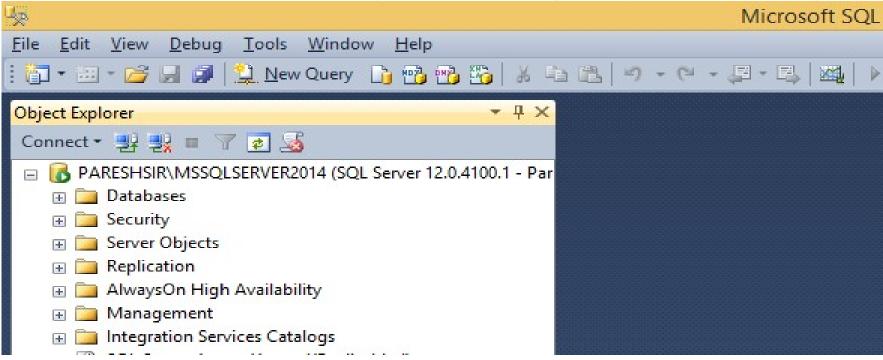
DQS (Data Quality Service)

- Cleansing Data Demo
- Data Matching Demo
- Data Activity Demo

Create SSIS Package

- Deploy SSIS Package
- First Way
- Second Way

Execute Package & Deploy Package



Run as administrator & File System -> import package & Execute Package

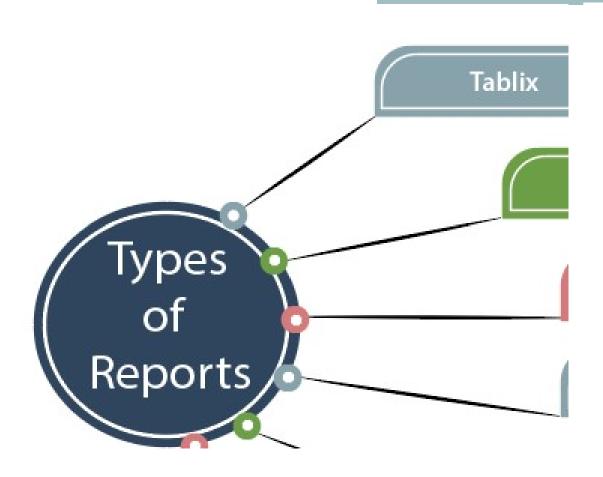
Deploying Package Video

Reporting

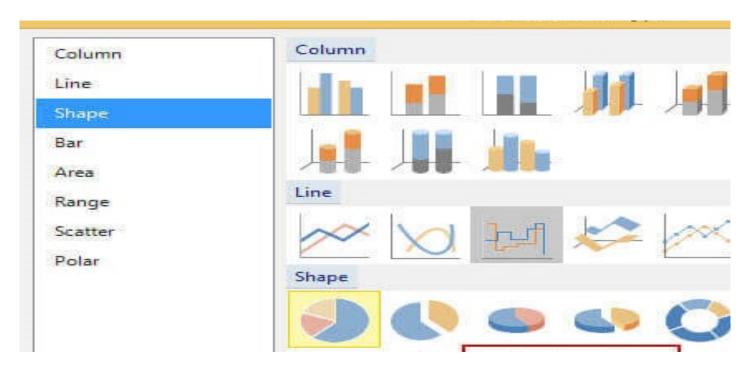
Demo

SQL Server Reporting Serv Microsoft's server-centric platform

- Reports are created locally on it
- Reports are deployed to a serv



Types of charts in SSRS



What is BI?

"a set of concepts and methodologies to improve decision making in business through use of facts and fact-based systems".

Business Intelligence (BI) is about getting

the **right information**, to the right decision makers (**right people**), at the **right time**, in the **right format**.

• BI leads to:

fact-based decision making "single version of the truth"

