Experiment No 2

Aim:- To perform Extraction ,Transformation and Loading Operation using SSIS(SQL Server Integration Services) tool of Microsoft.

Introduction:

SQL Server Integration Service (**SSIS**) is a component of the Microsoft SQL Server database software that can be **used** to execute a wide range of data migration tasks. **SSIS** is a fast & flexible data warehousing tool **used** for data extraction, loading and transformation like cleaning, aggregating, merging data, etc

What is an ETL process ?

ETL stands for Extraction, Transformation and Loading. It is a process in data warehousing to extract data, transform data and load data to final source. ETL covers a process of how the data are loaded from the source system to the data warehouse. Let us briefly describe each step of the ETL process.

Extraction

Extraction is the first step of ETL process where data from different sources like txt file, XML file, Excel file or various sources collected.

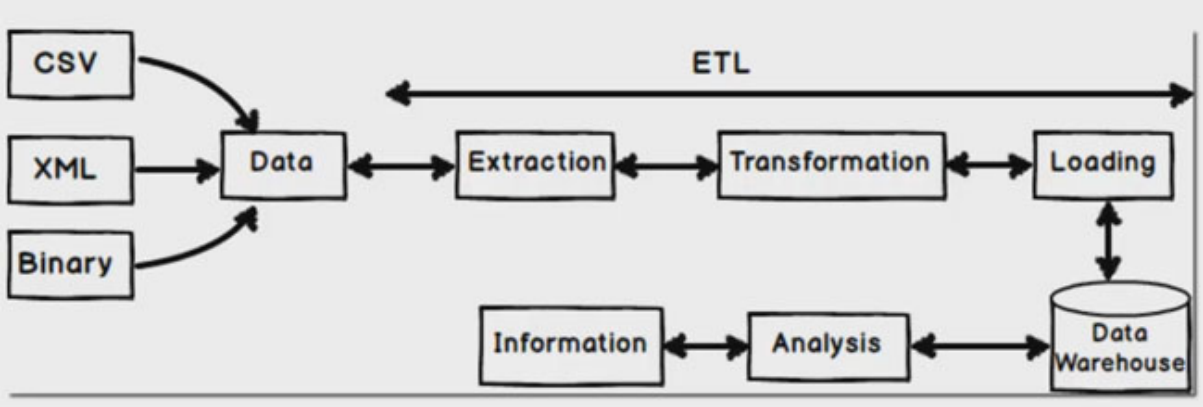
Transformation

Transformation is the second step of ETL process where all collected data is been transformed into same format i.e. format can be anything as per our requirement before loading it to data-warehouse i.e. it may be data-type format, data merge format, splitting format, alphabet joining format, currency format etc.

Loading

Final step of ETL process, The big chunks of data which is collected from various sources and transformed then finally load to our data warehouse.

Steps To Perform ETL using SSIS are as given below:

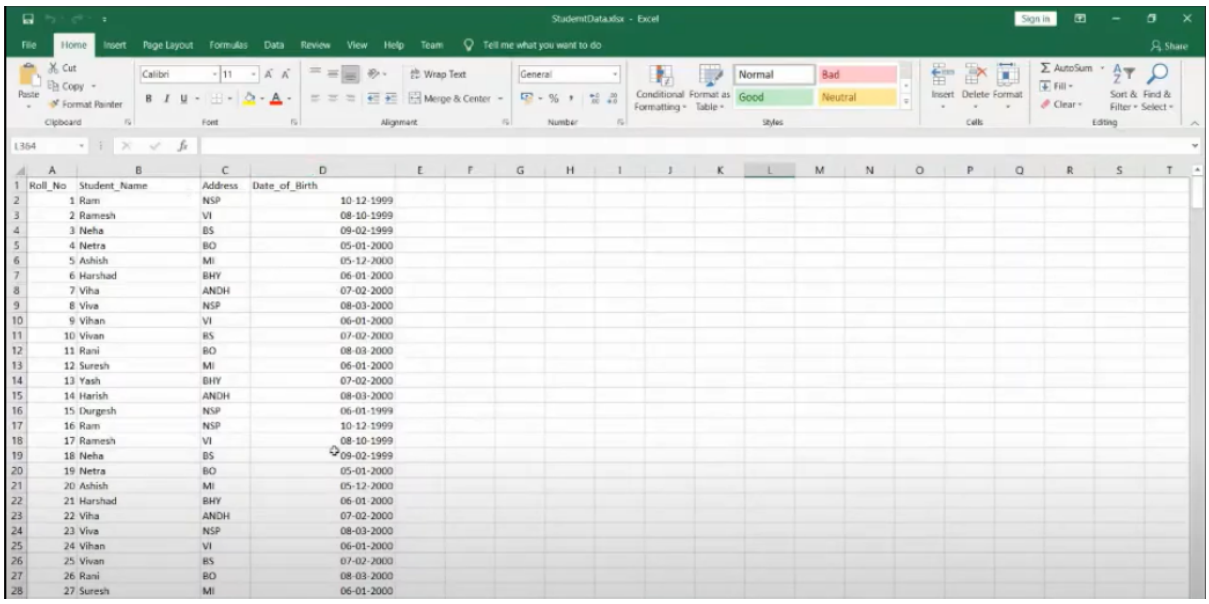


Step 1:- Extract Data From Excel File

Before you read this steps kindly make sure you have installed Microsoft business intelligence along with SQL Server.

A) I thin step we will create a simple excel file with a columns names as Roll_No,Student_Name,Address,DateofBirth.

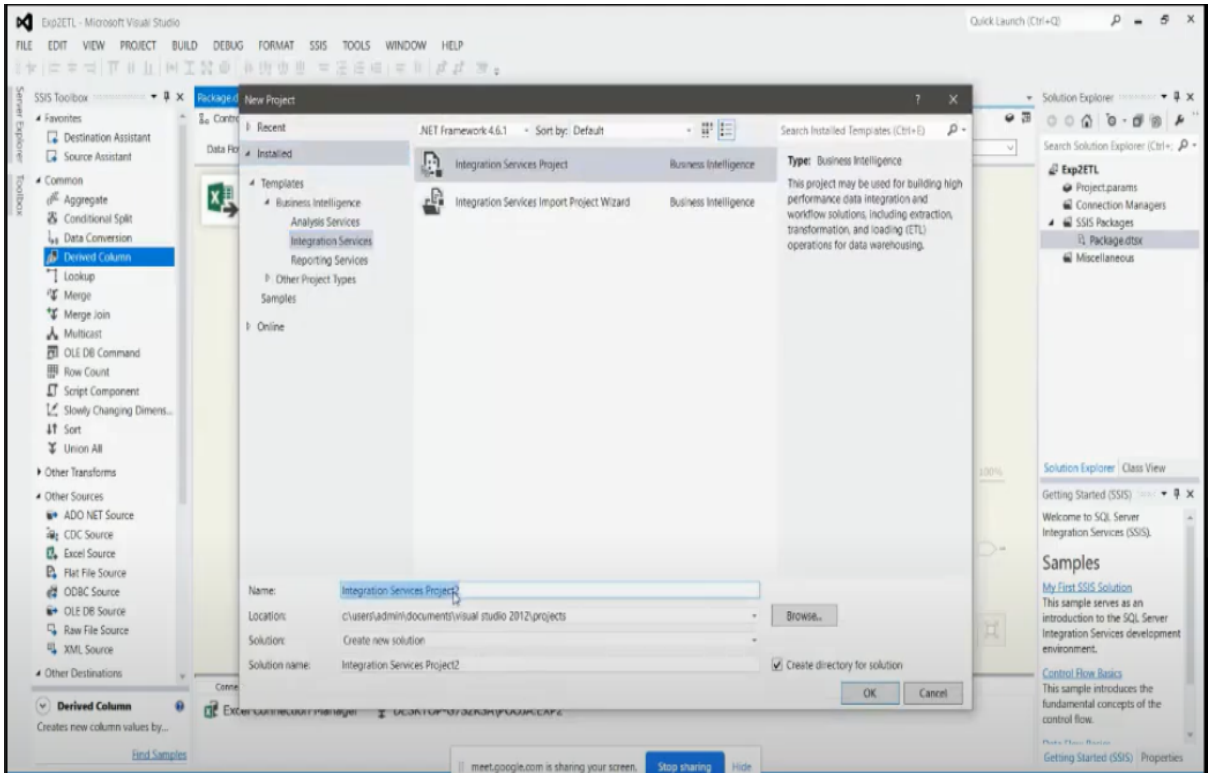
Add some data as shown in below image.

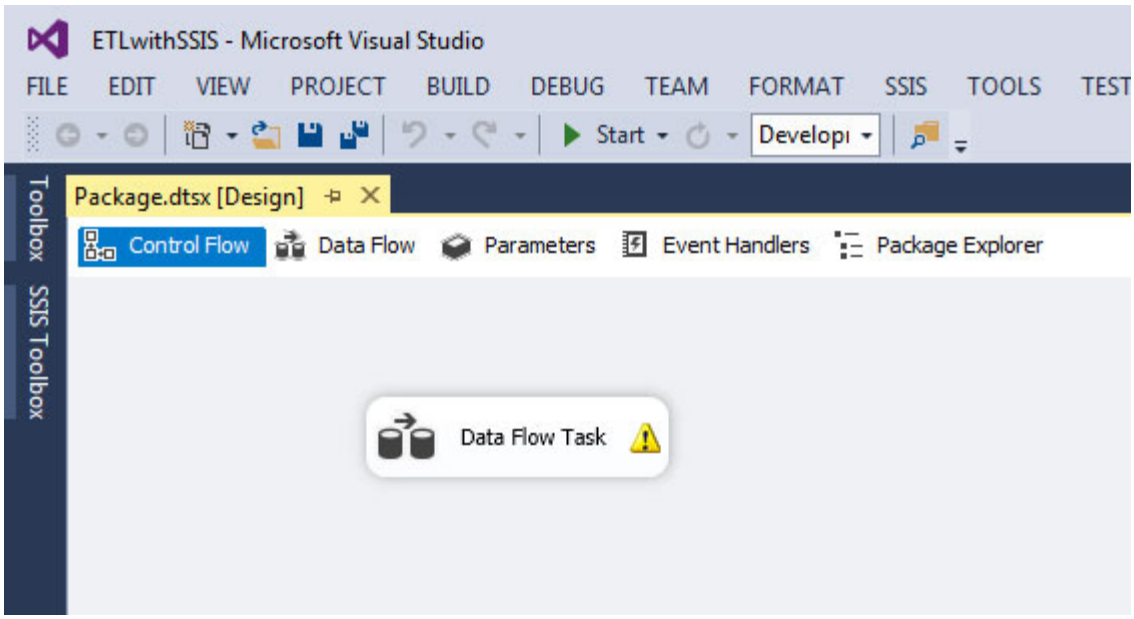


Roll_No	Student_Name	Address	Date_of_Birth
1	Ram	NSP	10-12-1999
2	Ramesh	VI	08-10-1999
3	Neha	BS	09-02-1999
4	Netra	BO	05-01-2000
5	Ashish	MI	05-12-2000
6	Harshad	BHY	06-01-2000
7	Viha	ANDH	07-02-2000
8	Viva	NSP	08-03-2000
9	Vihan	VI	06-01-2000
10	Vivan	BS	07-02-2000
11	Rani	BO	08-03-2000
12	Suresh	MI	06-01-2000
13	Yash	BHY	07-02-2000
14	Harish	ANDH	08-03-2000
15	Durgesh	NSP	06-01-1999
16	Ram	NSP	10-12-1999
17	Ramesh	VI	08-10-1999
18	Neha	BS	09-02-1999
19	Netra	BO	05-01-2000
20	Ashish	MI	05-12-2000
21	Harshad	BHY	06-01-2000
22	Viha	ANDH	07-02-2000
23	Viva	NSP	08-03-2000
24	Vihan	VI	06-01-2000
25	Vivan	BS	07-02-2000
26	Rani	BO	08-03-2000
27	Suresh	MI	06-01-2000

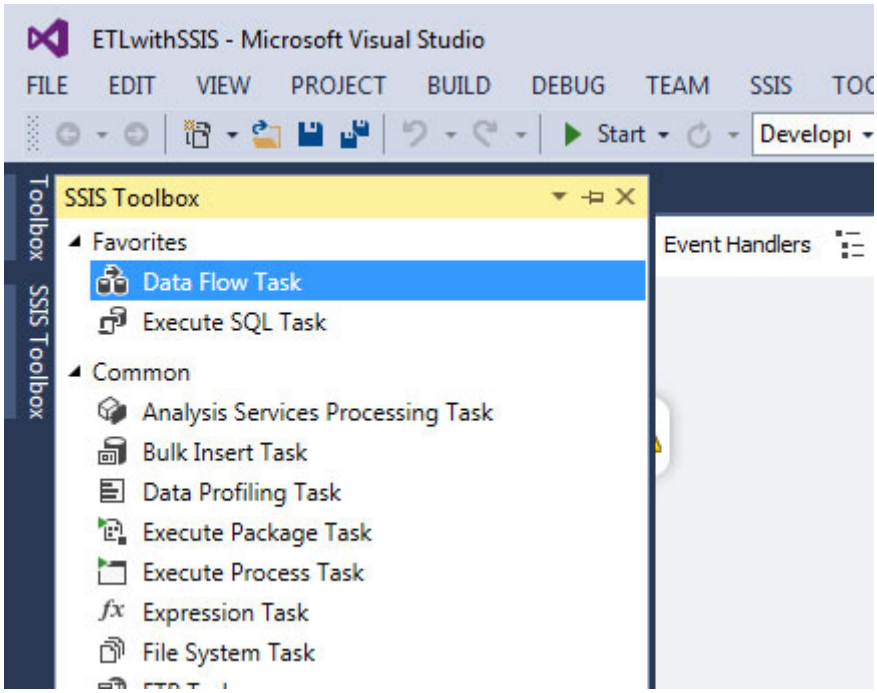
B) Create New Integration Service Project

Now open your SQL Server data tools just open it and go to FILE -> NEW -> PROJECT and Since we are doing ETL process and that process comes under SSIS so we need to create Integration Services so choose that -> Integration Service Project.





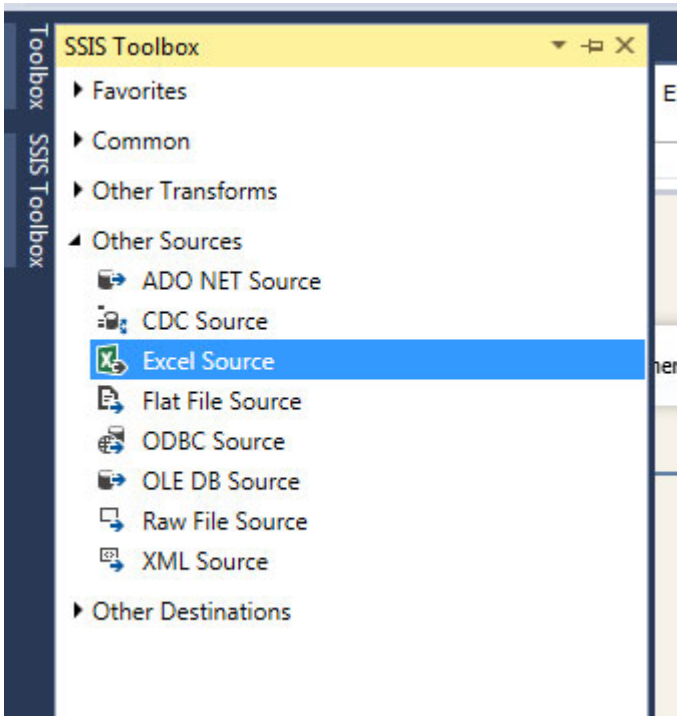
Now go to SSIS Toolbox and drag and drop "Data Flow Task" to control panel as show in below image.



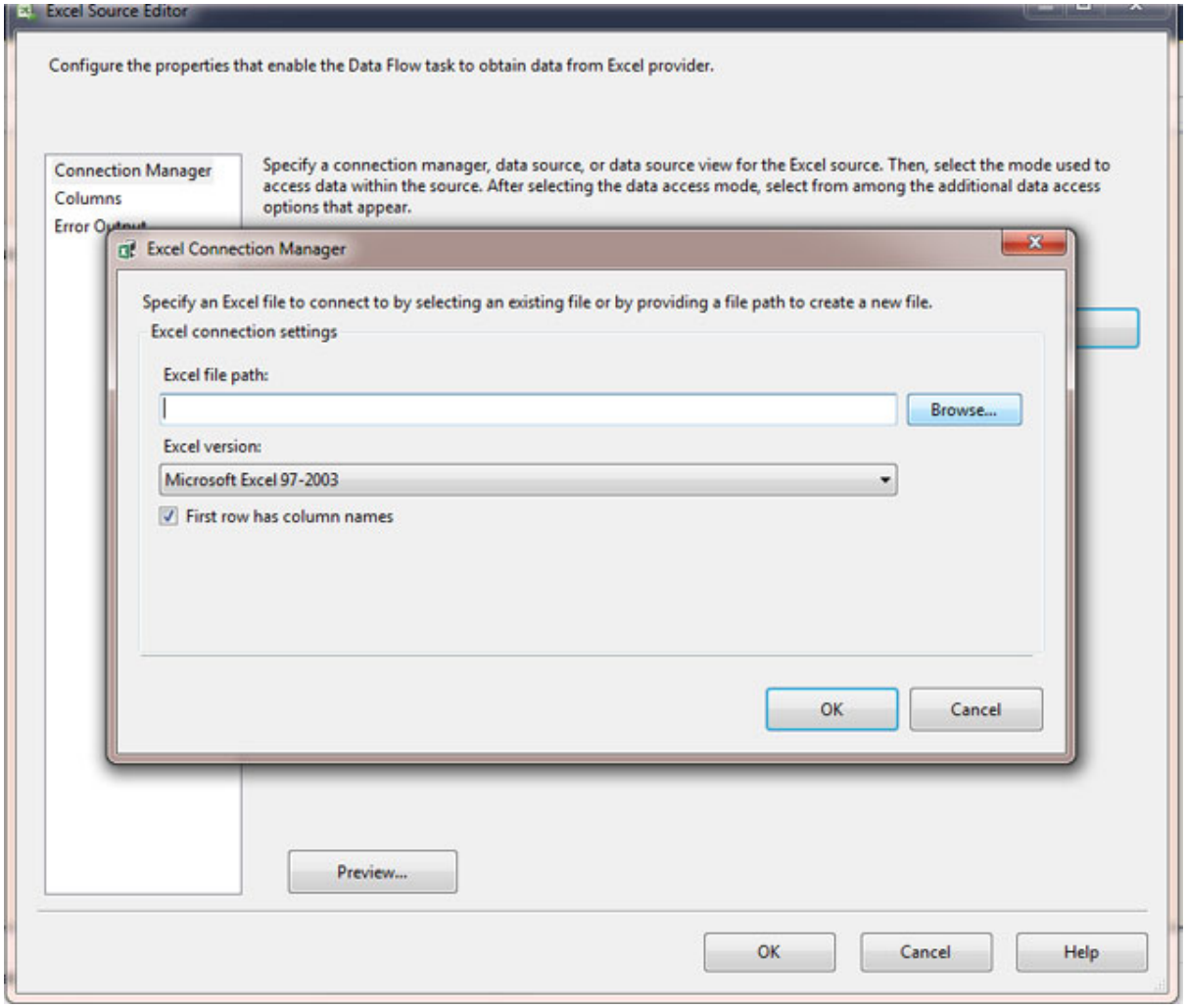
Just double click on "Data Flow Task" to take you to "Data Flow".

Now go to SSIS Toolbox and from Other Souces tab just drap and drop Excel Souces. Why excel source because our inital data which we want to extract it is in excel format.

So just drag excel file as shown below image and right click and rename it so that if any developer reads it can easily able to understand.

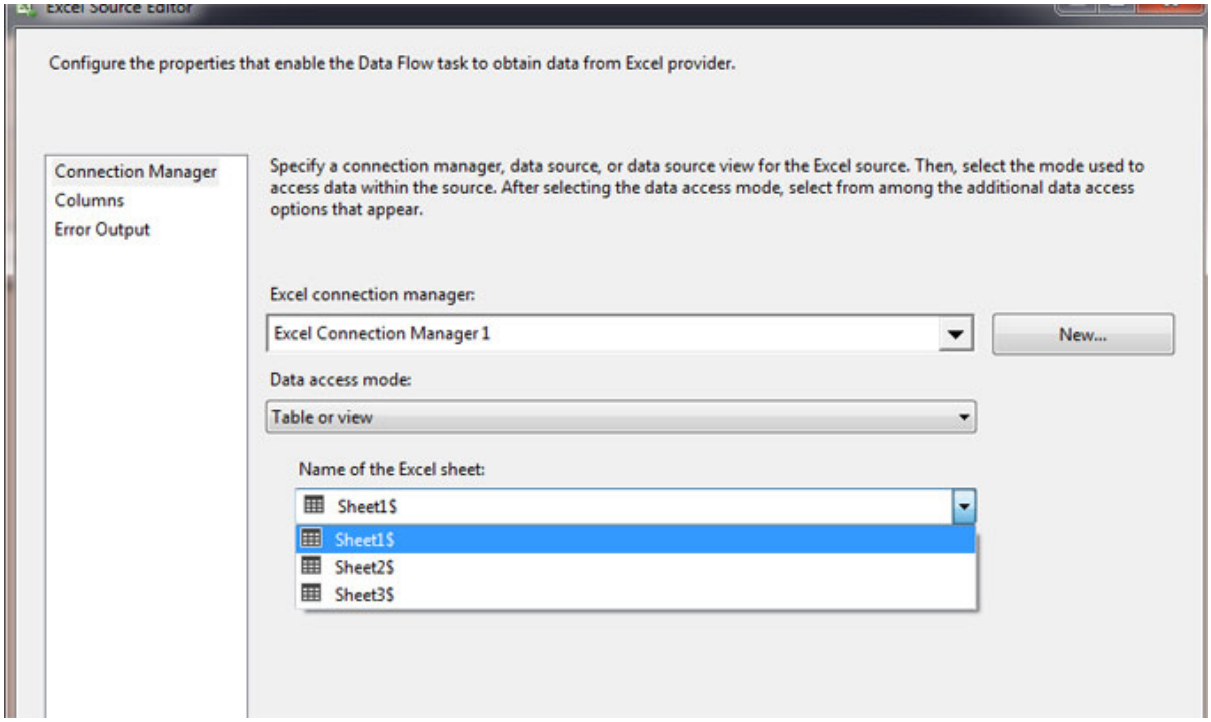


Now right click on that excel source -> Edit -> Click on New button - Browse Excel file from your computer as shown in below image.



Since our first column of excel file is having column names so we need to check this below check box as you see in above image.

Now select Data access mode as "Table on view" then select Excel sheet name from drop down.



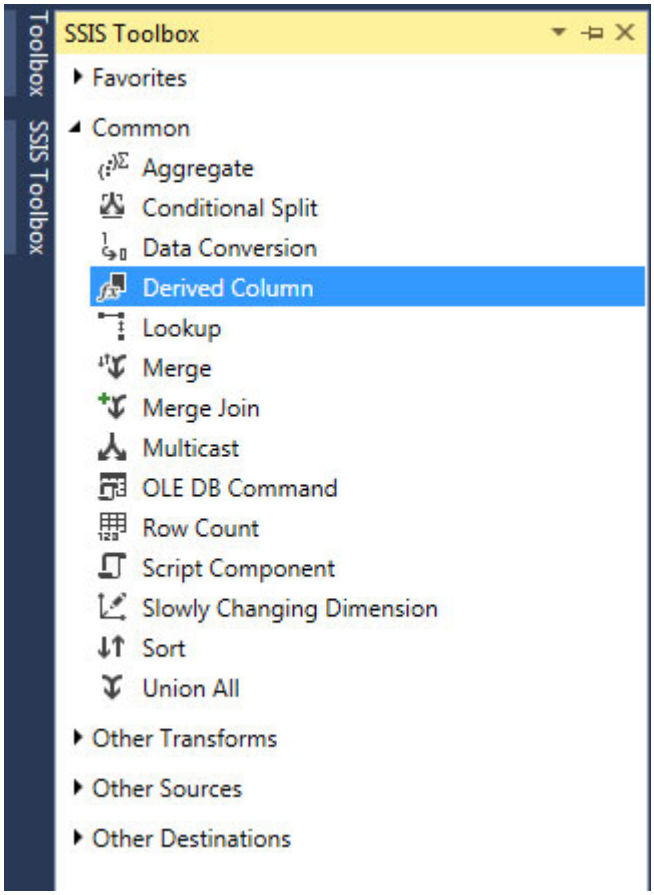
If you want to preview you can also do that by clicking preview button.

Finally click on OK button. So now your excel source is ready.It means we have successfully extracted our excel data file to SSIS excel data source.

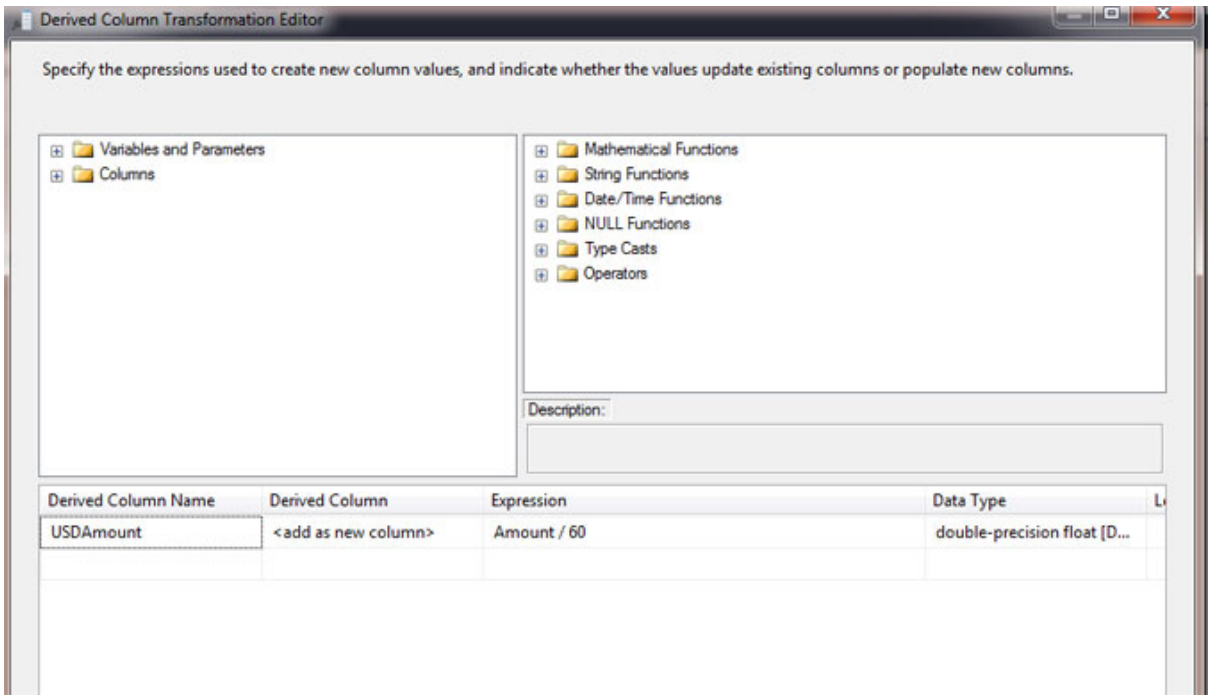
Step 2 : Transform Data (Convert to US currency and Upper case)

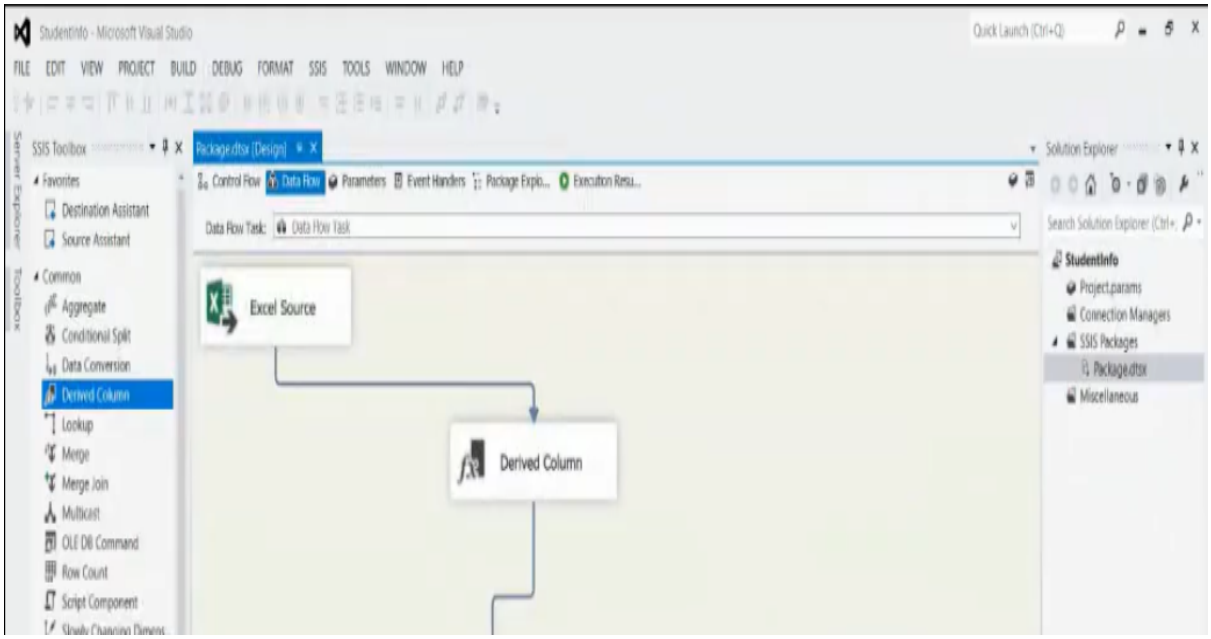
As you now in our excel file we have column name called "Amount" and that amount is in Indian currency. So we need to convert that Indian amount to USD amount so to do that we will drag and drop "Derived Column" from SSIS toolbox.

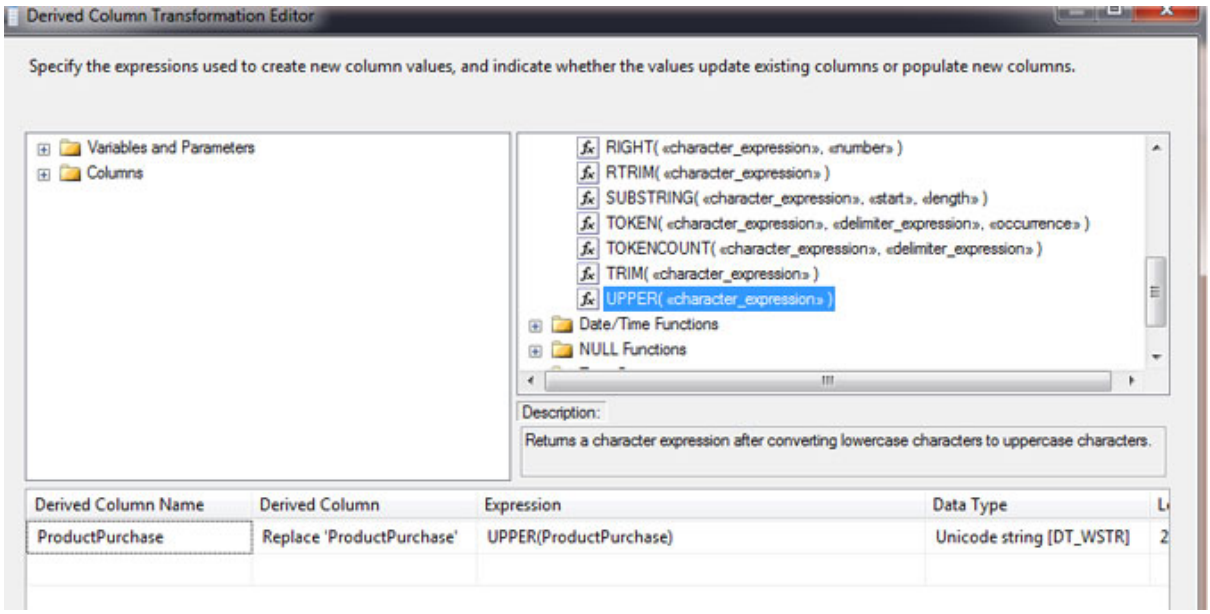
Now if you see on Excel Source file box there are two arrows "Red" and Blue". Just drag that "Blue" arrow and join it to "Derived column" as shown in below image and rename that "Derived column".



Now right click on "Derived column" i.e ConvertingAmounttoUSD and click on Edit as show below image.

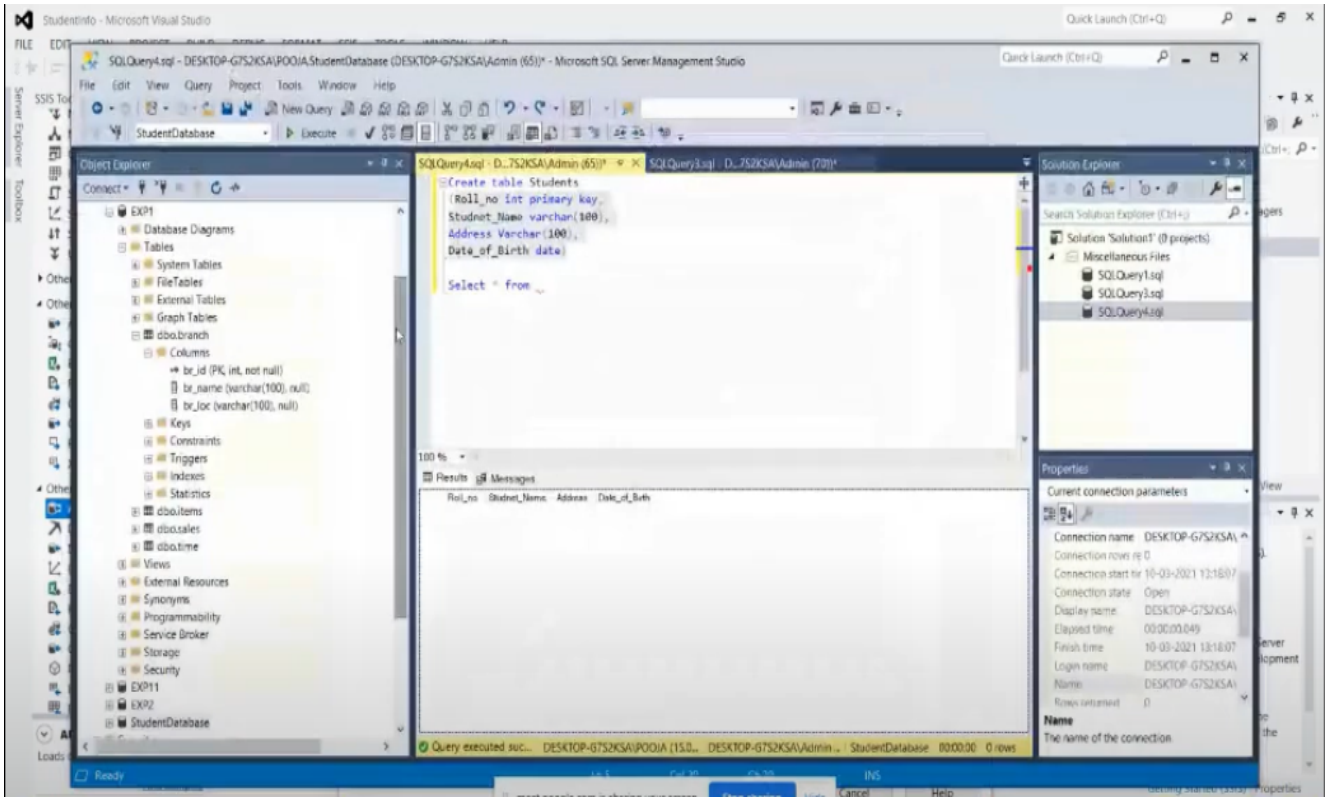




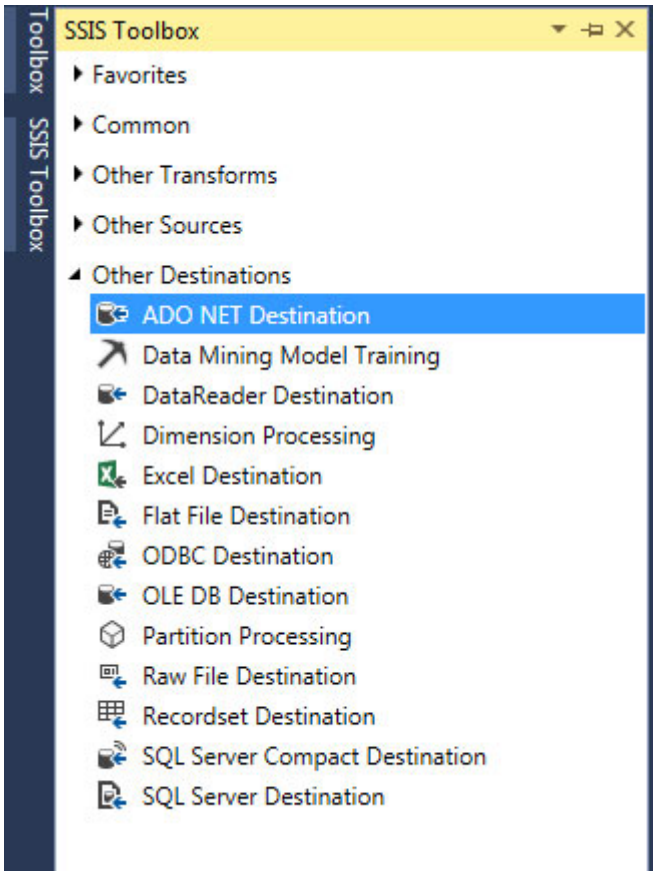


Step 3 : Loading Data To SQL Server

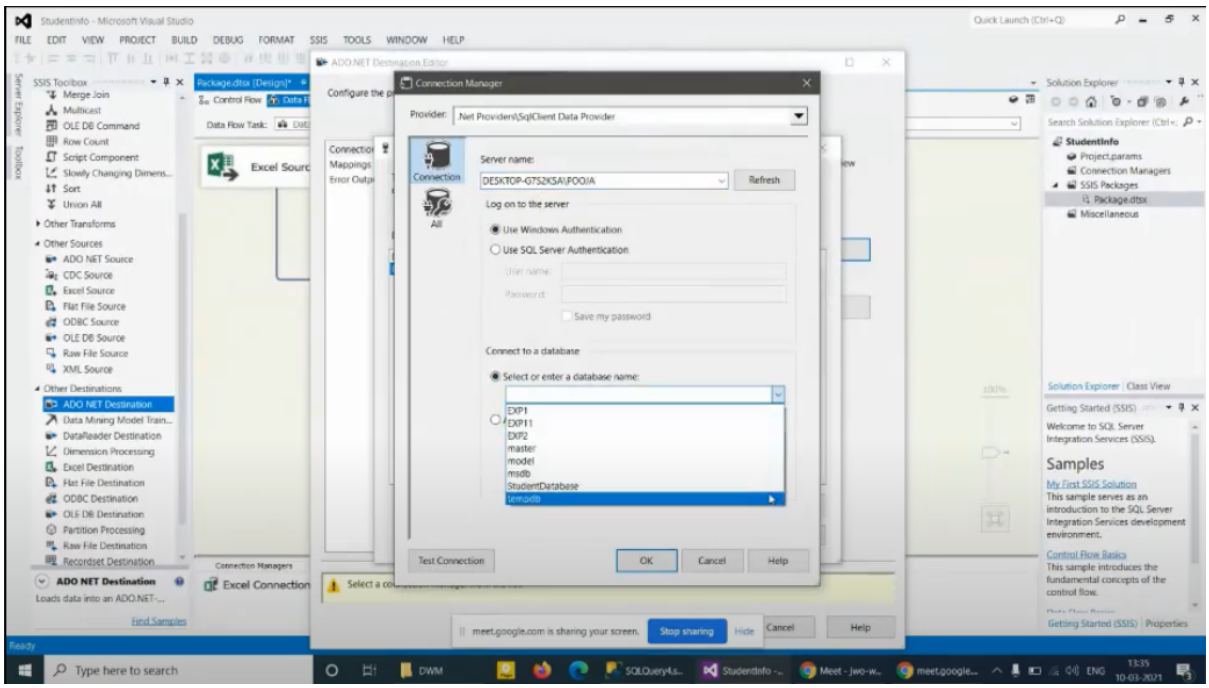
Before you start this step just open up your SQL server management studio and create a new database if need or just a new table with same excel column names as shown below image



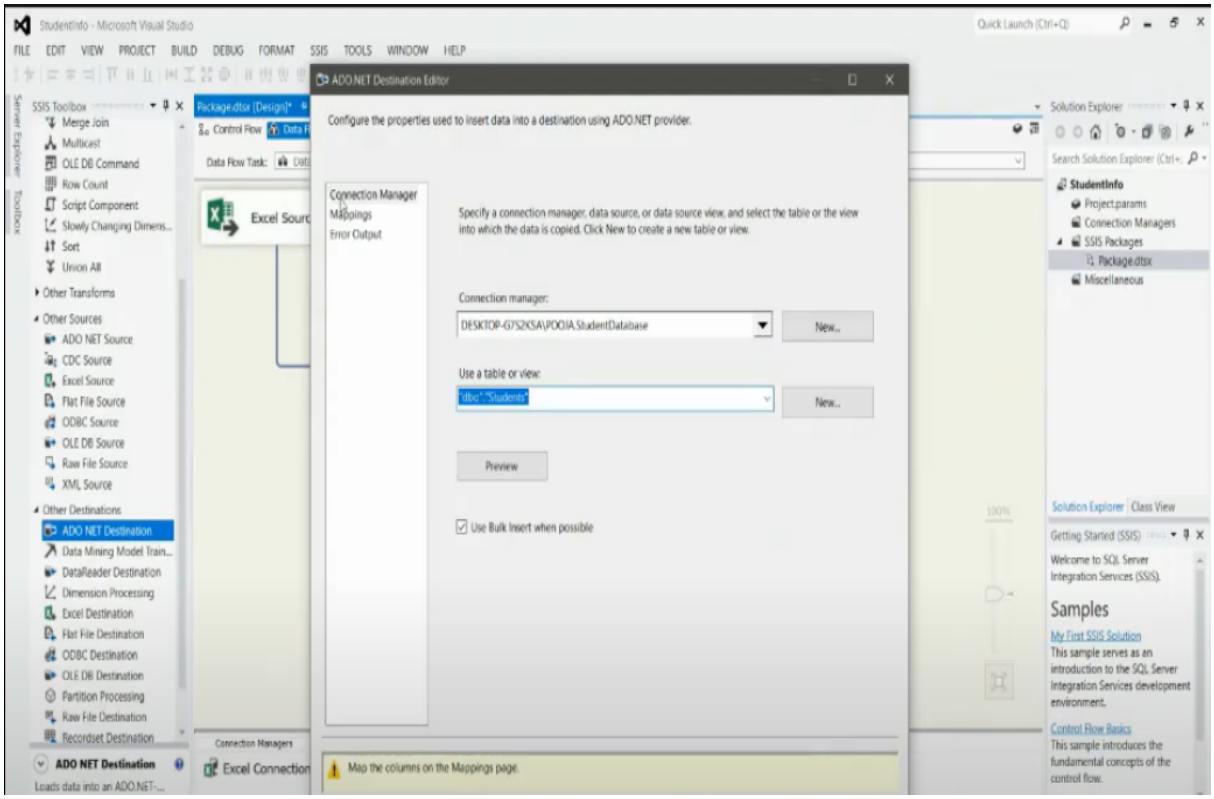
ow go to your SSDT and from SSIS toolbox under Other Destination select "ADO.NET Destination"and drag it to "Data Flow" and drag arrow from "CapitalProductName" to "ADO.NET Destination".

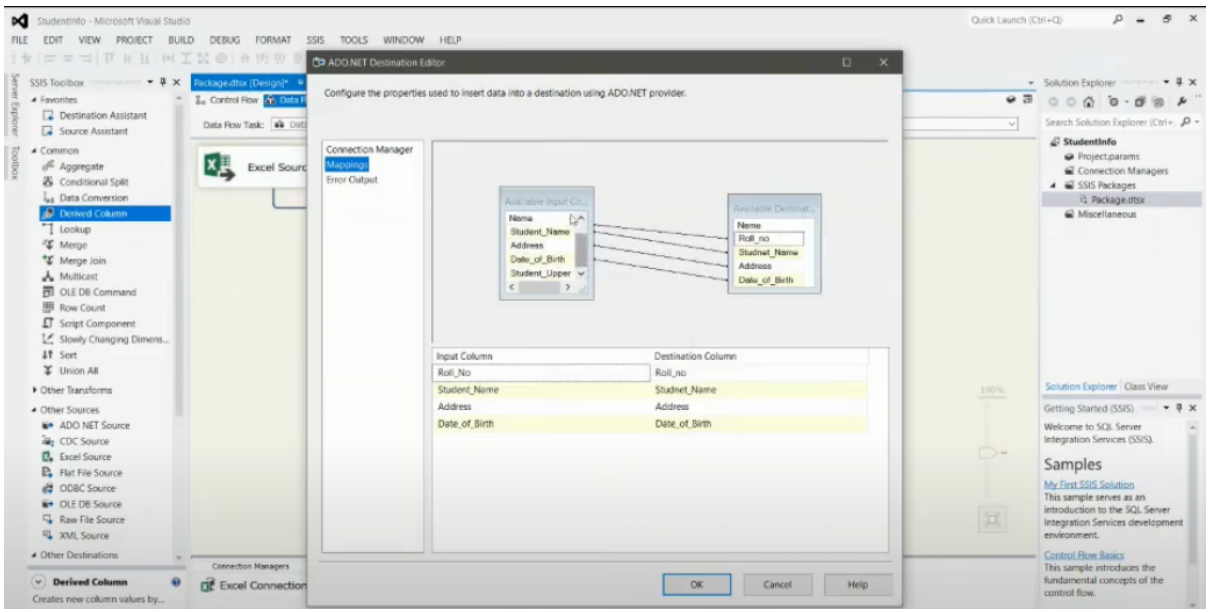


Right click and click on edit add your server name if you do not know your extact server name go to sql server management studio -> File -> click on connect and copy that server name. Now come back to SSDT and click New button -> Again New -> give here server name. Once you give server name automatically database dropdown will populate. Select database and click on OK button. Finally choose table as shown in below image and save it (OK button).



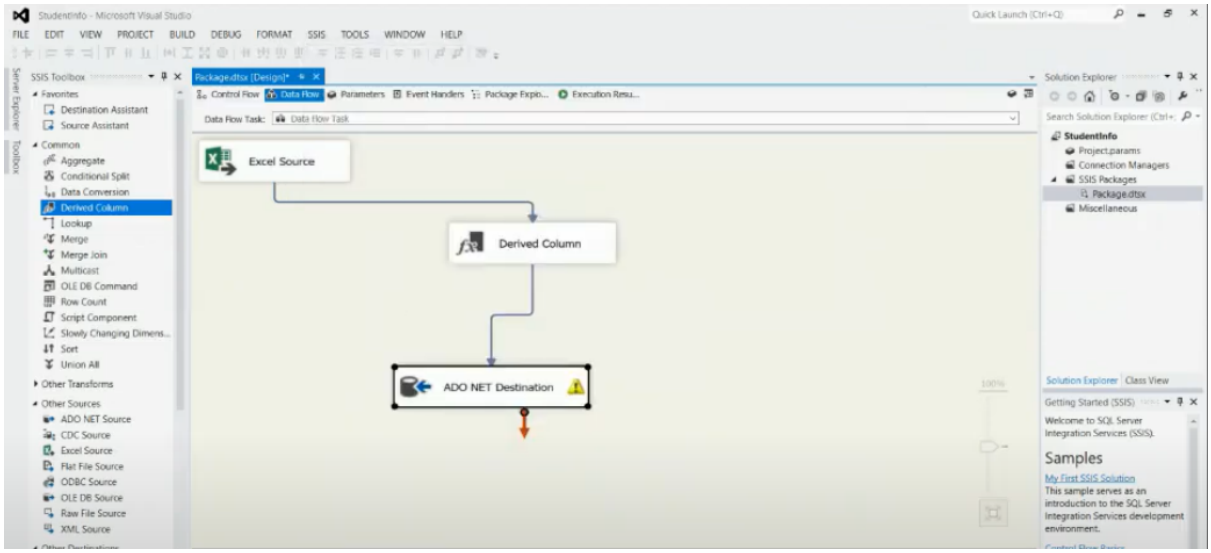
Right click and click on edit add your server name if you do not know your extact server name go to sql server management studio -> File -> click on connect and copy that server name. Now come back to SSDT and click New button -> Again New -> give here server name.Once you give server name automatically database dropdown will populate. Select database and click on OK button. Finally choose table as shown in below image and save it (OK button).





So what we did here it that we have created a path from where excel data flow from excel file -> Transform Name & UPPER CASE -> SQL Server Data Warehouse.

Final step just go to DEBUG and click on Start button from top menu of SSDT or just click on F5. Automatically data will flow from Excel Source -> SQL Server.



This is how ETL process is executed using SSIS.