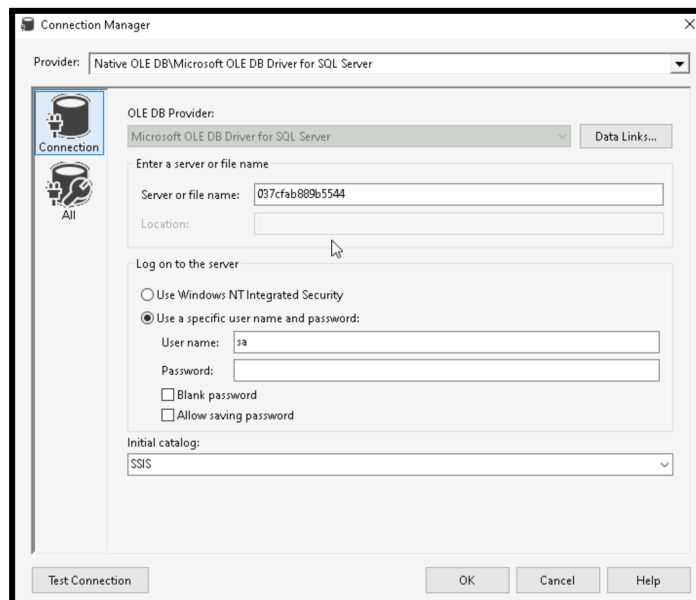


SSIS: Assignment

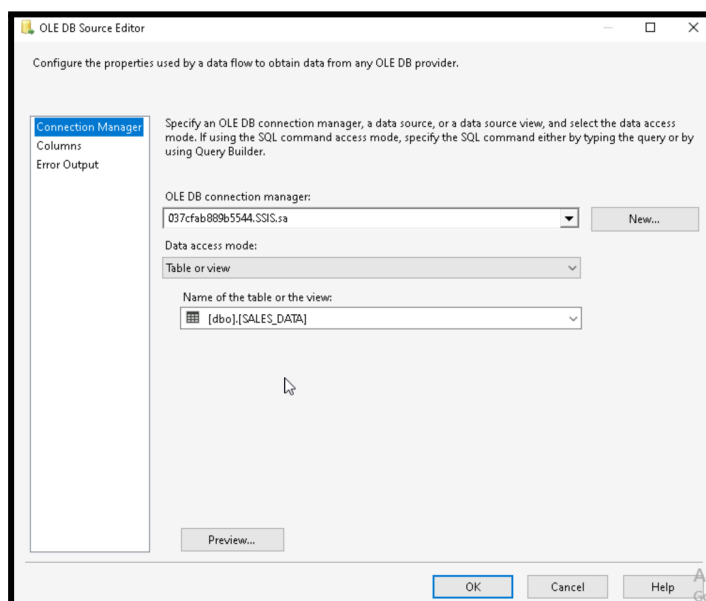
Task 1: Integration with ETL Data Warehouse (DWH)

Scenario: Your company has a data warehouse designed to consolidate data from various sources for analytical purposes. You need to create an SSIS package that extracts data from a transactional database and loads it into the data warehouse.

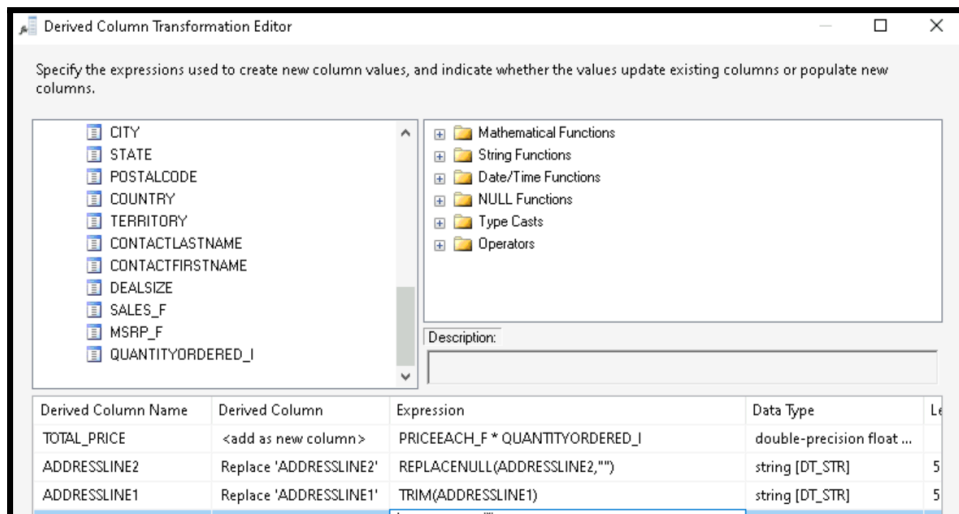
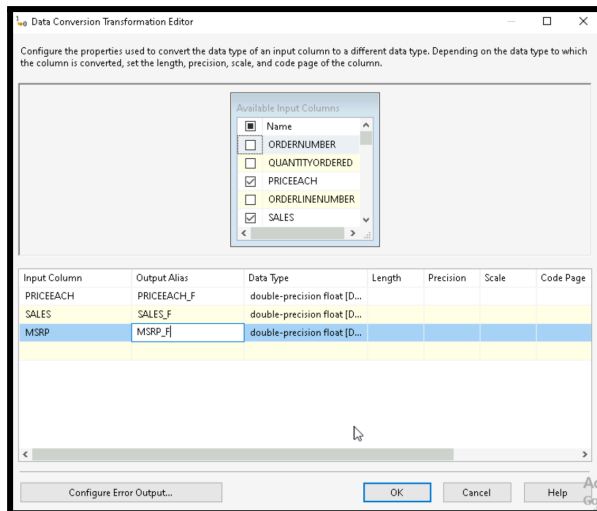
1. Create a Connection Manager to connect to the transactional database and the data warehouse.



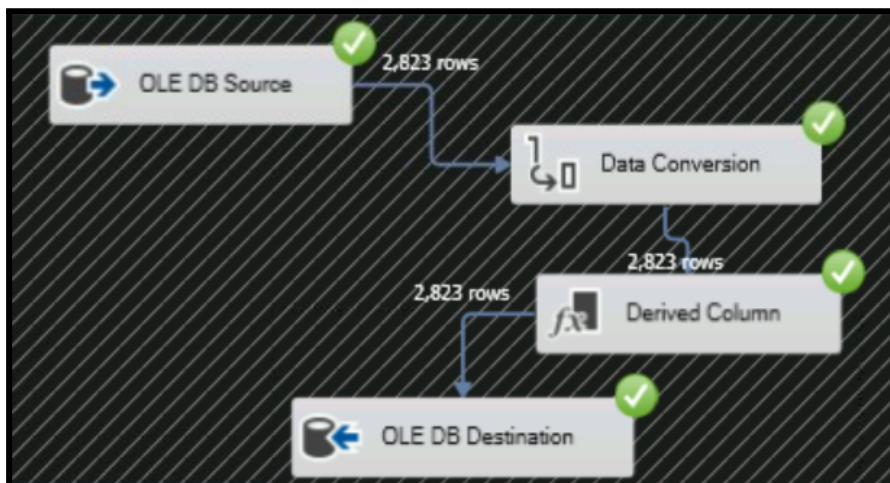
2. Extract Data from a transactional table (e.g., SalesData) using an OLE DB Source.



3. Transform Data: Apply necessary transformations such as data type conversions, data cleansing, and calculations.



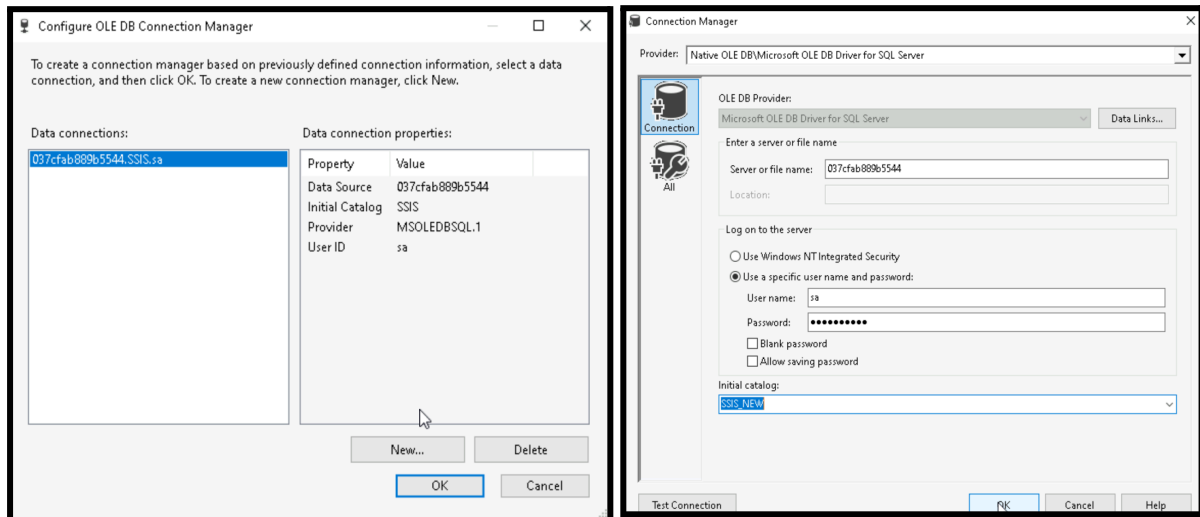
4. Load Data into the data warehouse (e.g., FactSales table).



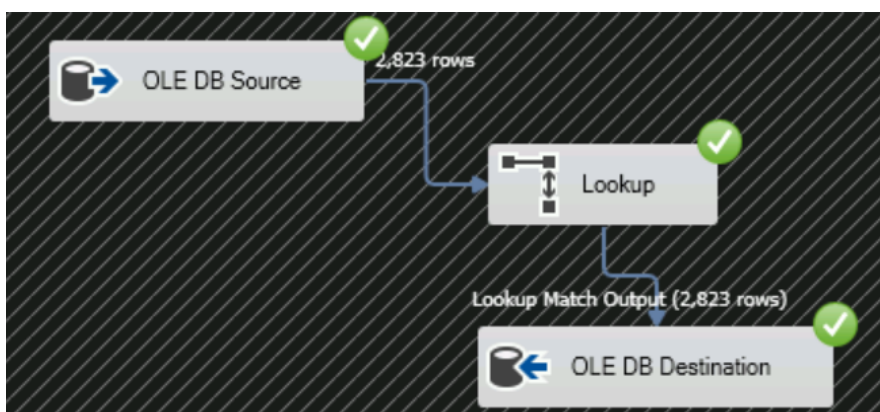
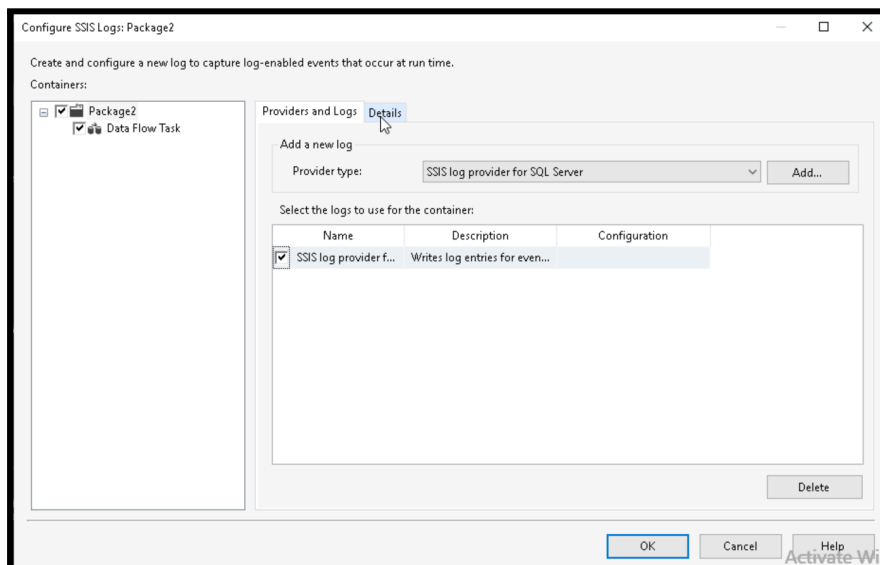
Task 2: Data Warehouse Migrations

Scenario: Your organization is migrating its data warehouse from one server to another. You need to create an SSIS package that facilitates this migration.

1. Create Connection Managers for both the source and destination data warehouses.



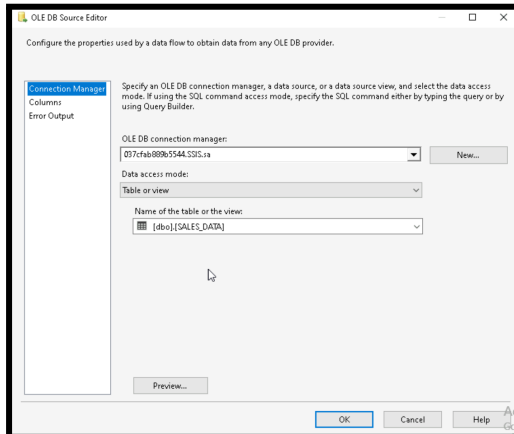
2. Transfer Data from the source data warehouse to the destination using the Data Flow Task. (Create Connection Managers for both the source and destination data warehouses, Transfer Data from the source data warehouse to the destination using the Data Flow Task.)



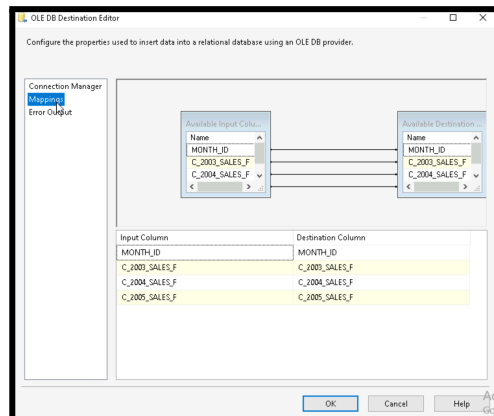
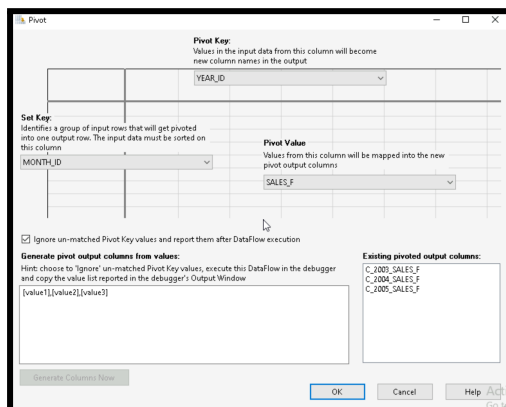
Task 3: Implementing a Pivot Transformation

Scenario: You have data in a normalized format and need to pivot it for reporting purposes.

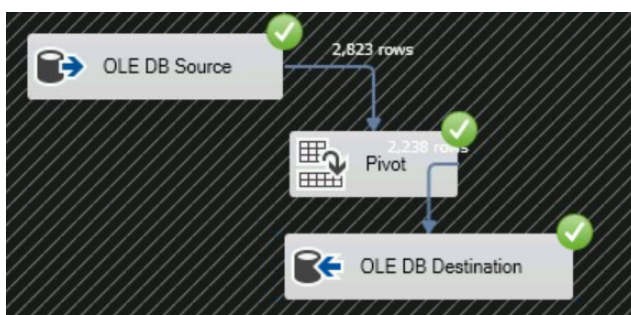
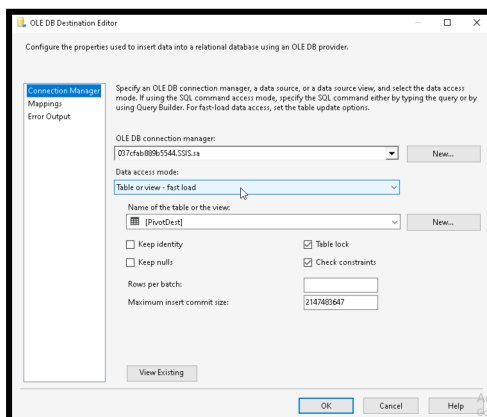
1. Extract Data from the source table using an OLE DB Source.



2. Apply a Pivot Transformation to transform the normalized data into a pivoted format.

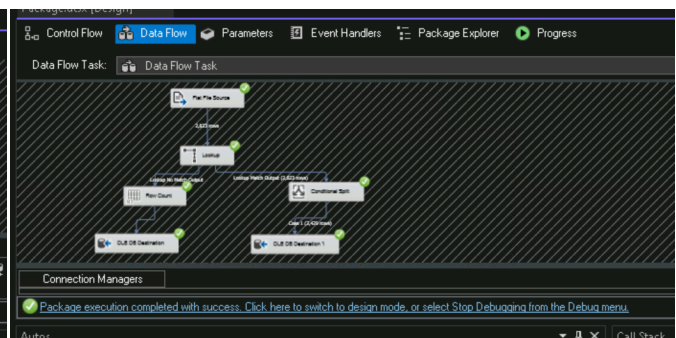
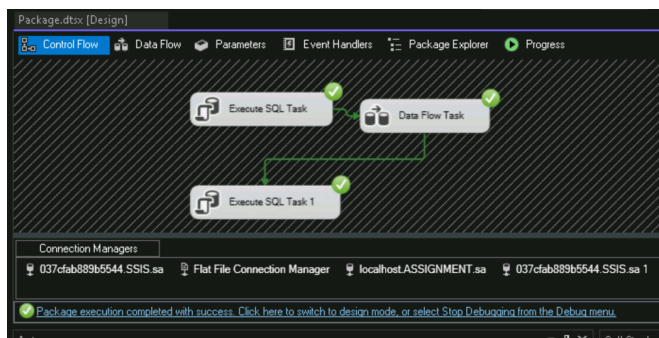
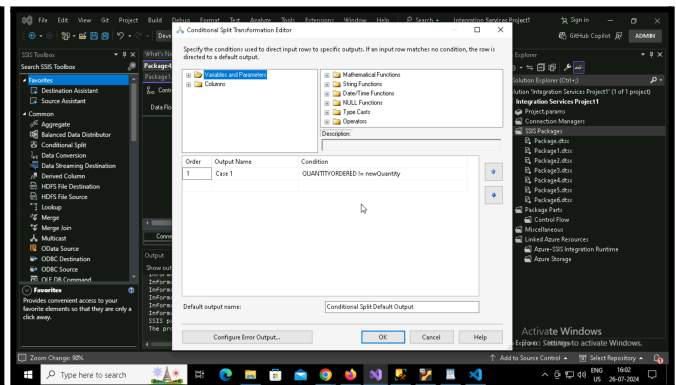
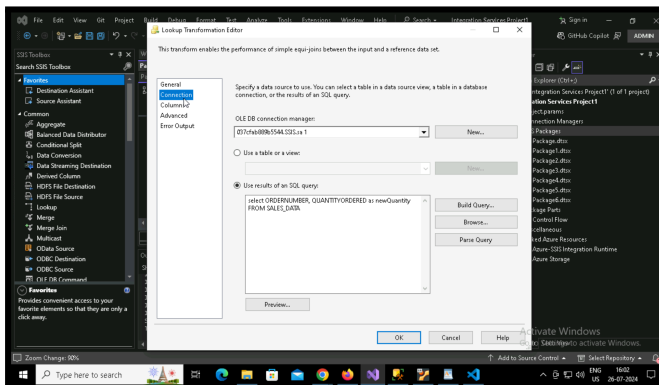
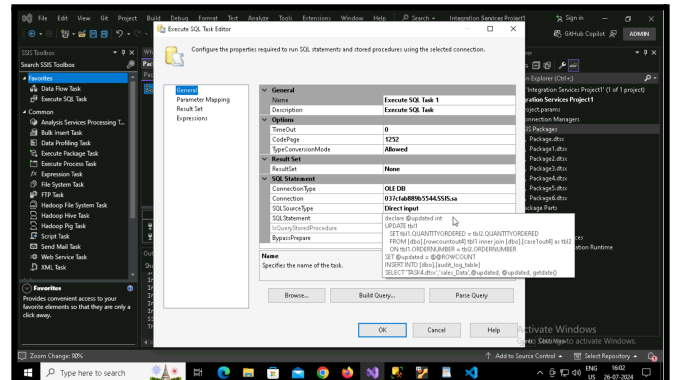
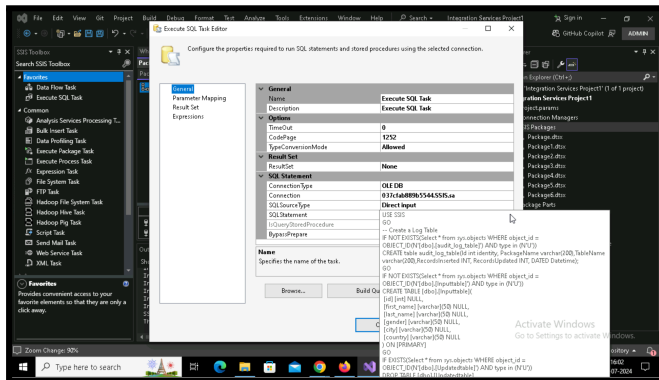


3. Load the Pivoted Data into a destination table.



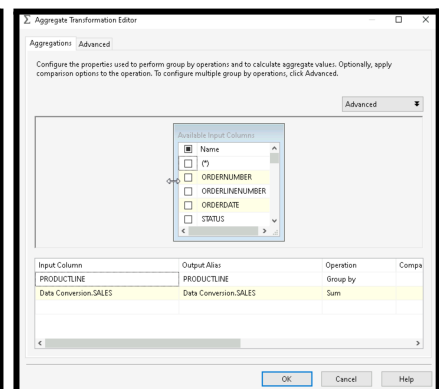
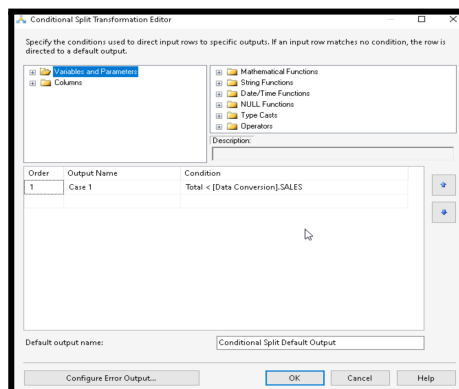
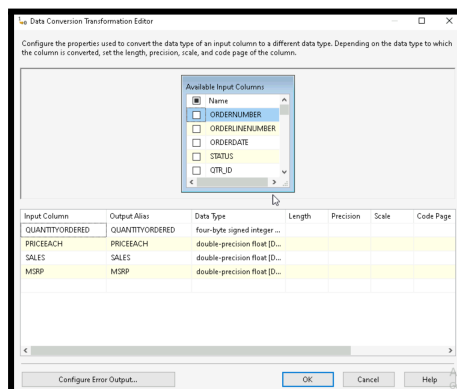
Task 4: Incremental Load

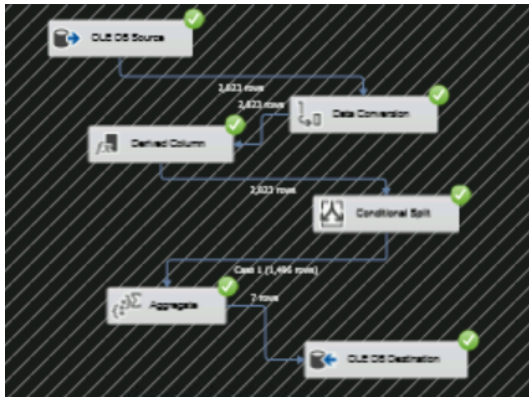
Scenario: To optimize ETL processes, you need to implement an incremental load to update only the changed data in the data warehouse.



Task 5: Transformations

Scenario: Your company needs to transform raw data into a format suitable for reporting. You need to perform multiple transformations within an SSIS package.





Task 6: MERGE & FUZZY LOOKUP

Scenario: You need to merge two datasets and use fuzzy matching to handle potential duplicates.

Fuzzy Lookup Transformation Editor

Configure the properties used to perform a lookup operation between an input dataset and a reference dataset using a best-match algorithm.

Reference Table Columns Advanced

Specify the join columns and the use of reference columns.

Available Input Columns		Available Lookup...	
Name	Pass Thru...		
id	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> Name	
first_name	<input checked="" type="checkbox"/>	<input type="checkbox"/> ID	
last_name	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> City	
email	<input checked="" type="checkbox"/>		
gender	<input checked="" type="checkbox"/>		

Lookup Column	Output Alias
City	City

OK Cancel Help

Merge Join Transformation Editor

Configure the properties used to join two sources of sorted data. Select the join type and then specify the columns to be used as the join key. Join keys must be used in the order specified by the sort-key position of the column.

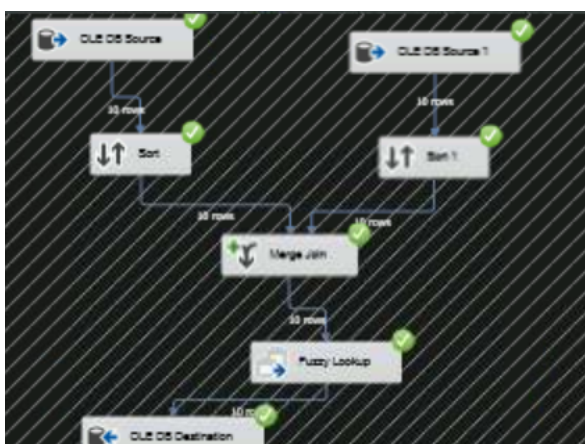
Join type: Left outer join Swap Inputs

Sort	Name	Order	Join K...
<input type="checkbox"/>	id	1	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	first_name	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	last_name	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	email	0	<input type="checkbox"/>

Sort 1	Name	Order	Join K...
<input checked="" type="checkbox"/>	id	1	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	city	0	<input type="checkbox"/>

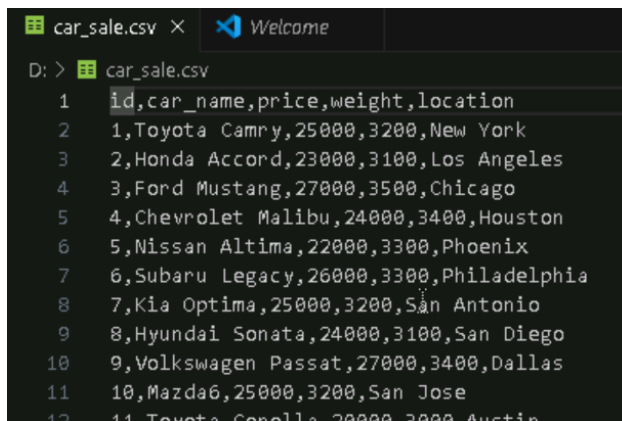
Input	Input Column	Output Alias
Sort 1	id	id
Sort	first_name	first_name
Sort	last_name	last_name
Sort	email	email
Sort	gender	gender
Sort 1	city	city

OK Cancel Help

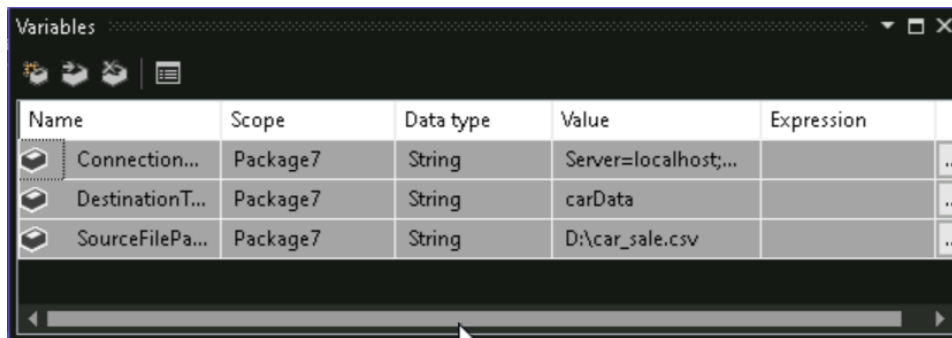
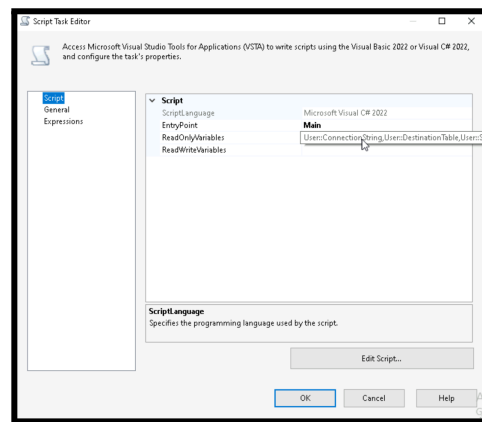


Task 7: Using Script Task

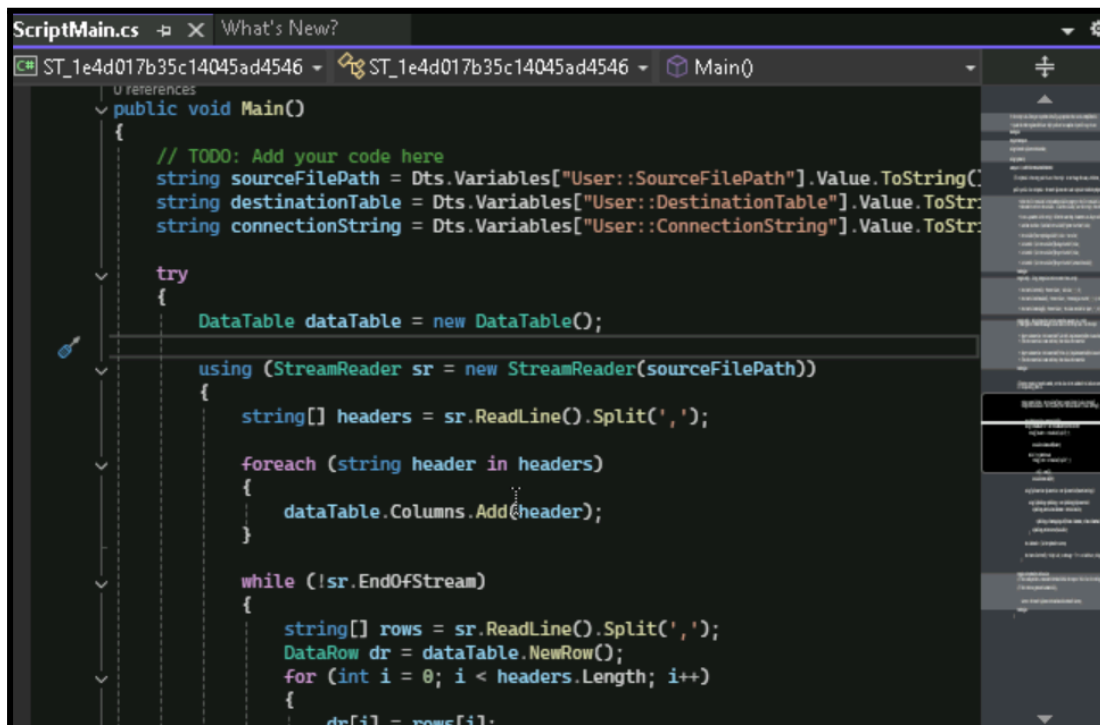
Scenario: You need to perform a complex data transformation that is not supported by the standard SSIS components. A Script Task can be used to achieve this.



```
D:\> car_sale.csv
1 id,car_name,price,weight,location
2 1,Toyota Camry,25000,3200,New York
3 2,Honda Accord,23000,3100,Los Angeles
4 3,Ford Mustang,27000,3500,Chicago
5 4,Chevrolet Malibu,24000,3400,Houston
6 5,Nissan Altima,22000,3300,Phoenix
7 6,Subaru Legacy,26000,3300,Philadelphia
8 7,Kia Optima,25000,3200,San Antonio
9 8,Hyundai Sonata,24000,3100,San Diego
10 9,Volkswagen Passat,27000,3400,Dallas
11 10,Mazda6,25000,3200,San Jose
12 11,Toyota Corolla, 20000, 3000, Austin
```



Name	Scope	Data type	Value	Expression
Connection...	Package7	String	Server=localhost,...	
DestinationT...	Package7	String	carData	
SourceFilePa...	Package7	String	D:\car_sale.csv	



```
ScriptMain.cs
What's New?
ST_1e4d017b35c14045ad4546 - ST_1e4d017b35c14045ad4546 - Main()

// References
public void Main()
{
    // TODO: Add your code here
    string sourceFilePath = Dts.Variables["User::SourceFilePath"].Value.ToString();
    string destinationTable = Dts.Variables["User::DestinationTable"].Value.ToString();
    string connectionString = Dts.Variables["User::ConnectionString"].Value.ToString();

    try
    {
        DataTable dataTable = new DataTable();

        using (StreamReader sr = new StreamReader(sourceFilePath))
        {
            string[] headers = sr.ReadLine().Split(',');

            foreach (string header in headers)
            {
                dataTable.Columns.Add(header);
            }

            while (!sr.EndOfStream)
            {
                string[] rows = sr.ReadLine().Split(',');
                DataRow dr = dataTable.NewRow();
                for (int i = 0; i < headers.Length; i++)
                {
                    dr[i] = rows[i];
                }
            }
        }
    }
}
```


Control Flow

Data Flow

Parameters

Event Handlers

Package Explorer

Progress

Script Task

Connection Managers

Right-click here to add a new connection man

Package execution completed with success. Click here to switch to design mode, or select Stop Debugging from the Debug menu.

Results		Messages			
	id	car_name	price	weight	location
1	1	Toyota Camry	25000.00	3200	New York
2	2	Honda Accord	23000.00	3100	Los Angeles
3	3	Ford Mustang	27000.00	3500	Chicago
4	4	Chevrolet Malibu	24000.00	3400	Houston
5	5	Nissan Altima	22000.00	3300	Phoenix
6	6	Subaru Legacy	26000.00	3300	Philadelphia
7	7	Kia Optima	25000.00	3200	San Antonio
8	8	Hyundai Sonata	24000.00	3100	San Diego
9	9	Volkswagen Passat	27000.00	3400	Dallas
10	10	Mazda6	25000.00	3200	San Jose