# SSIS Use Case

Requirement

We have a Sales data we need to do some cleaning. First, need to add a Full\_Name column with the First\_Name and Last\_Name. Next, we don’t need the Computer data. We need to find the Total\_Sales with the formula [ItemSold\*SoldPrice]. Sort TotalSales in Descending. Replace the ProductName Cell Phone with Mobiles. we need to separate the data based on the Region.

* The source Table contains Sales records with ID, First\_Name, Last\_Name, ProductName, ItemSold, SoldPrice, Country, and Region details.
* The target files should contain the following columns: ID, Full\_Name, ProductName, ItemSold, SoldPrice, TotalSales, Country and Region details.

Note: Copy the below Source data paste it into the notepad and save it in txt format. And it should be a comma “,“ delimited file. The file name should be “Sales”. The target file also should be in text format.

Sample Source Data:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ID | First\_Name | Last\_Name | ProductName | ItemSold | SoldPrice | Country | Region |
| 1 | Aamir | Shahzab | TV | 1 | 700 | USA | North America |
| 2 | M | Raza | Cell Phone | 2 | 800 | USA | North America |
| 3 | Christy | Ladson | TV | 3 | 1600 | USA | North America |
| 4 | John | Rivers | Laptop | 5 | 2400 | USA | North America |
| 5 | Najaf | Ali | Computer | 1 | 300 | Pakistan | Asia |
| 6 | Sukhjeet | Singh | TV | 2 | 900 | India | Asia |
| 7 | Chirag | Patel | Cell Phone | 5 | 1500 | India | Asia |
| 8 | Aleena | Aman | Laptop | 2 | 800 | Pakistan | Asia |
| 9 | Petra | Henry | TV | 10 | 5000 | France | Europe |
| 10 | Rita | Roger | Laptop | 7 | 2100 | France | Europe |
| 11 | Tamara | Tony | Cell Phone | 2 | 1200 | Germany | Europe |

Sample Target North\_America\_Sales:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ID | Full\_Name | ProductName | ItemSold | SoldPrice | Total\_Sales | Country | Region |
| 4 | John Rivers | Laptop | 5 | 2400 | 12000 | USA | North America |
| 3 | Christy Ladson | TV | 3 | 1600 | 4800 | USA | North America |
| 2 | M Raza | Mobiles | 2 | 800 | 1600 | USA | North America |
| 1 | Aamir Shahzad | TV | 1 | 700 | 700 | USA | North America |

Sample Target Asia\_Sales:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ID | Full\_Name | ProductName | ItemSold | SoldPrice | Total\_Sales | Country | Region |
| 7 | Chirag Patel | Mobiles | 5 | 1500 | 7500 | India | Asia |
| 6 | Sukhjeet Singh | TV | 2 | 900 | 1800 | India | Asia |
| 8 | Aleena Aman | Laptop | 2 | 800 | 1600 | Pakistan | Asia |

Sample Target Europe\_Sales:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ID | Full\_Name | ProductName | ItemSold | SoldPrice | Total\_Sales | Country | Region |
| 9 | Petra Henry | TV | 10 | 5000 | 50000 | France | Europe |
| 10 | Rita Roger | Laptop | 7 | 2100 | 14700 | France | Europe |
| 11 | Tamara Tony | Mobiles | 2 | 1200 | 2400 | Germany | Europe |

Table Format

Source File: - Source file records are separated by comma “,” Target Files: - These file records are separated by a comma “,”

Source file details:

The Associates need to do the following tasks as per the naming standards. Source Table Name: TotalSales

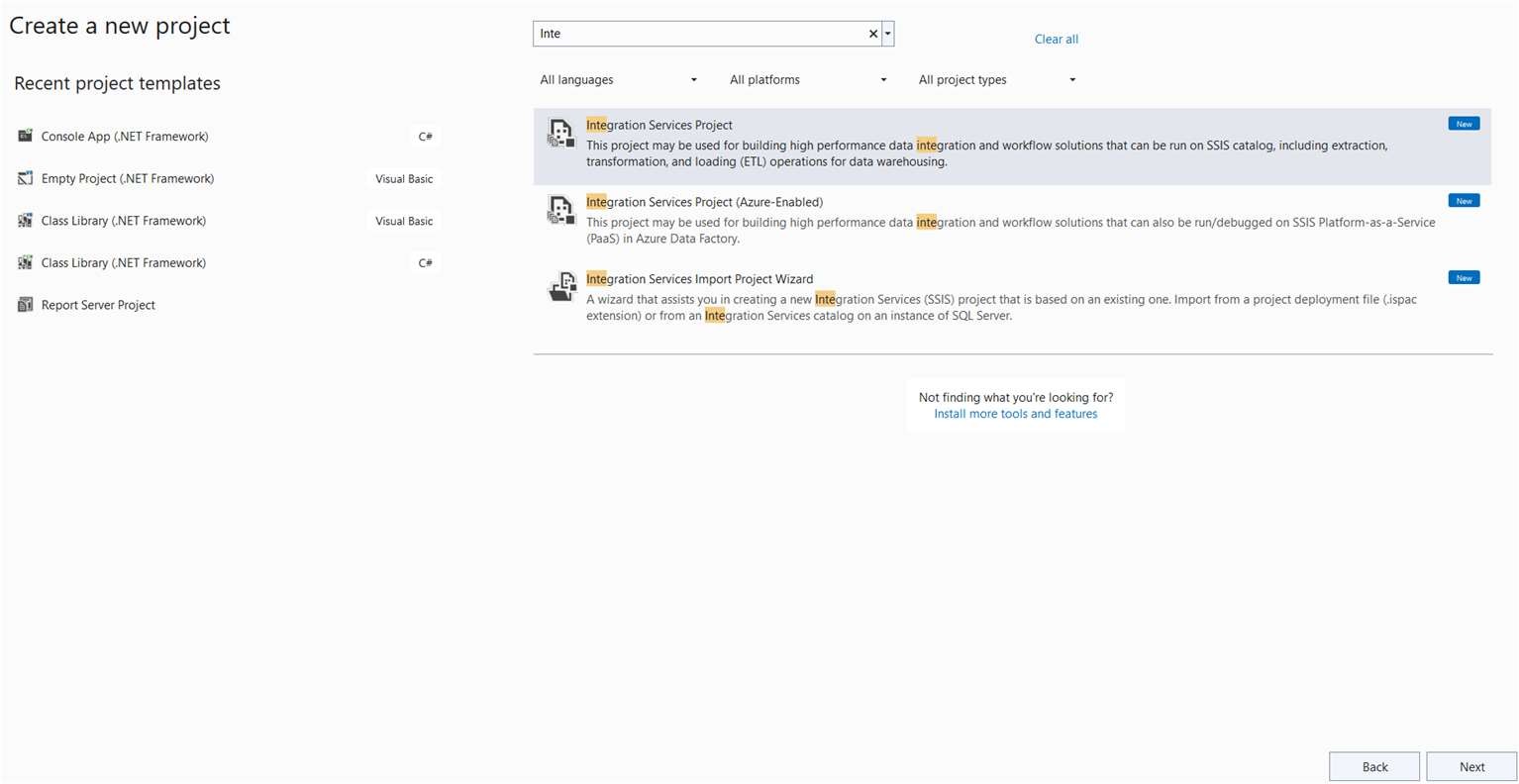
Target Table Names: North\_America\_Sales, Asia\_Sales and Europe\_Sales

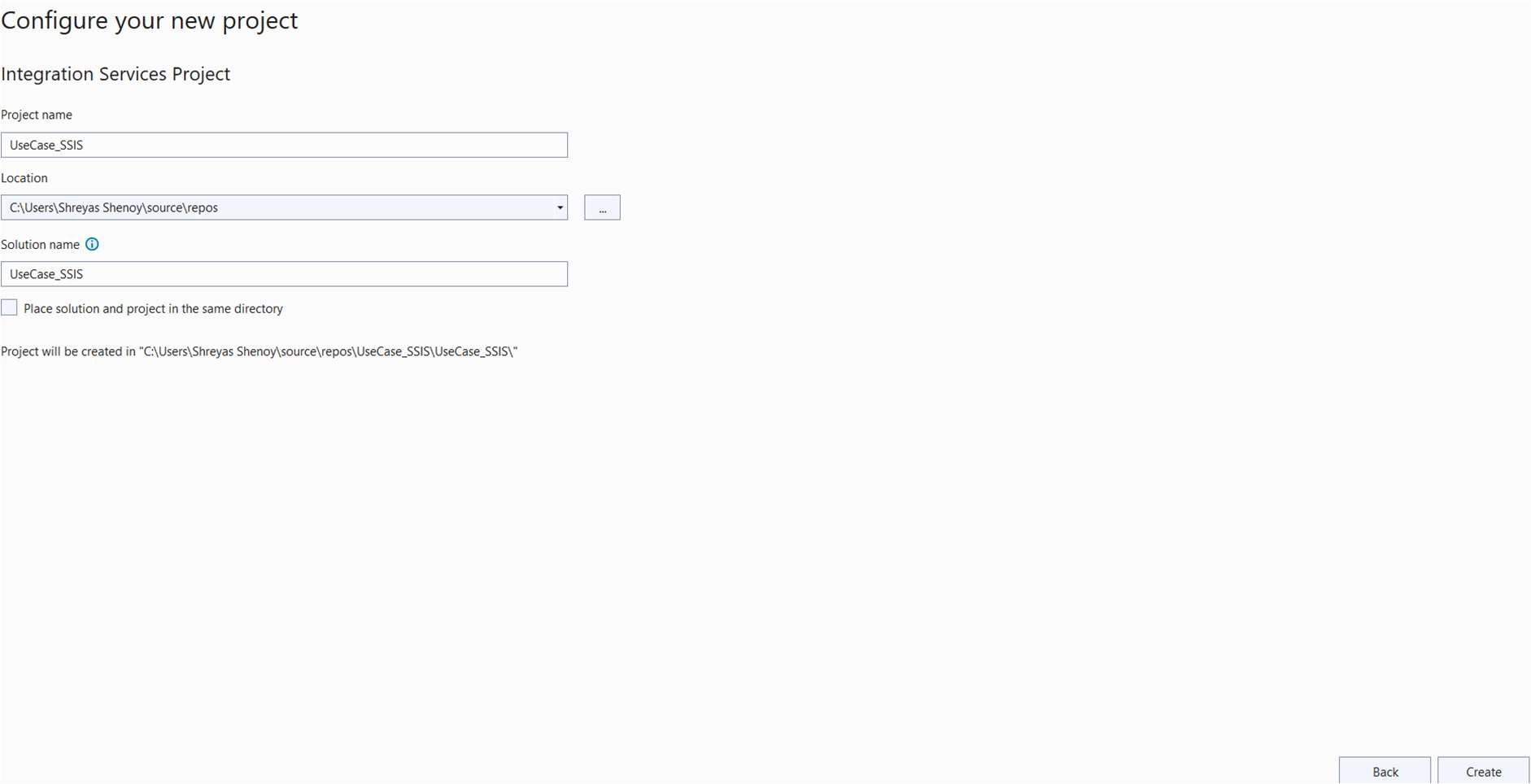
# Implementation

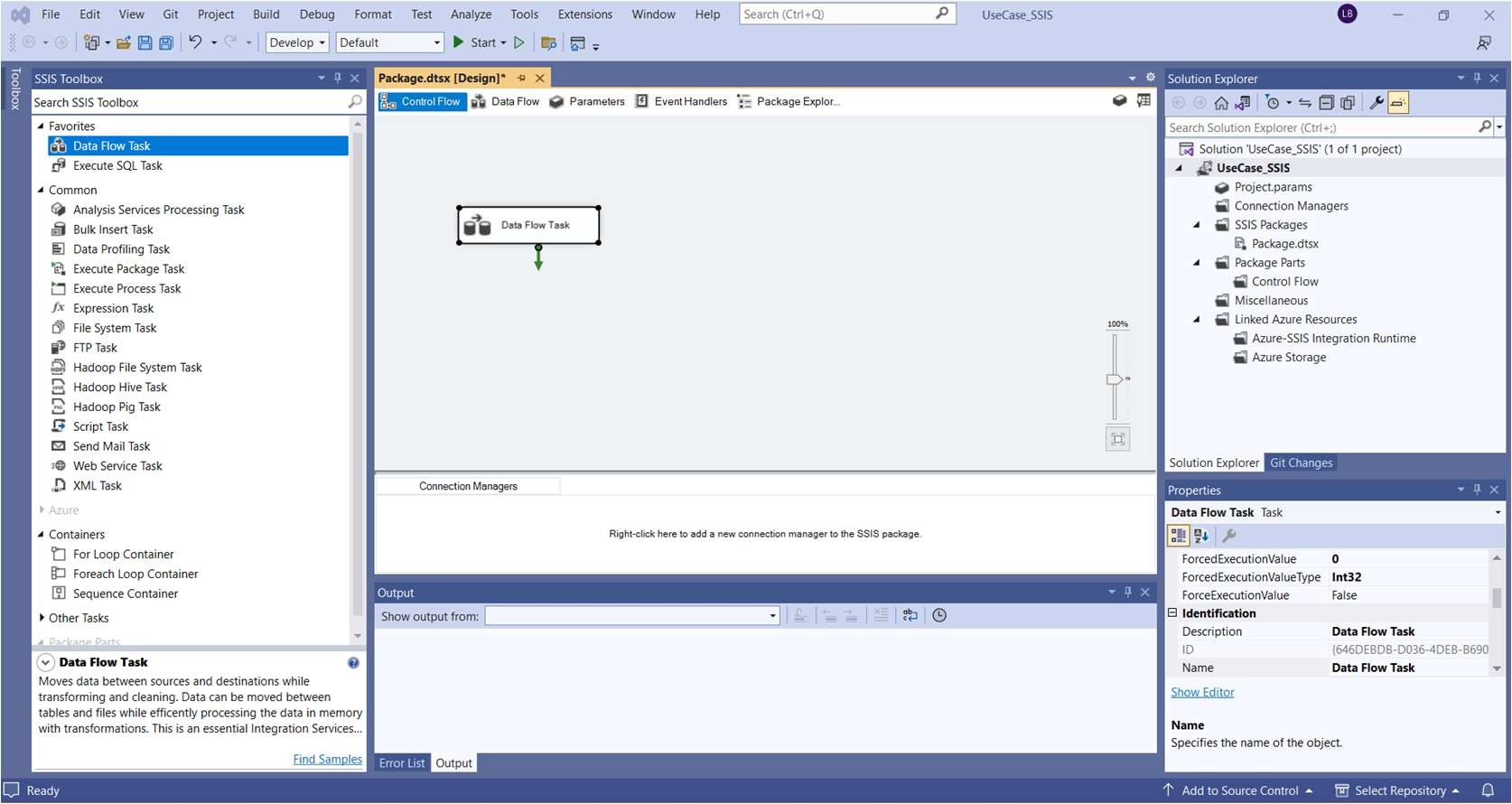
* Open the Visual Studio. Click on Create a New Project.



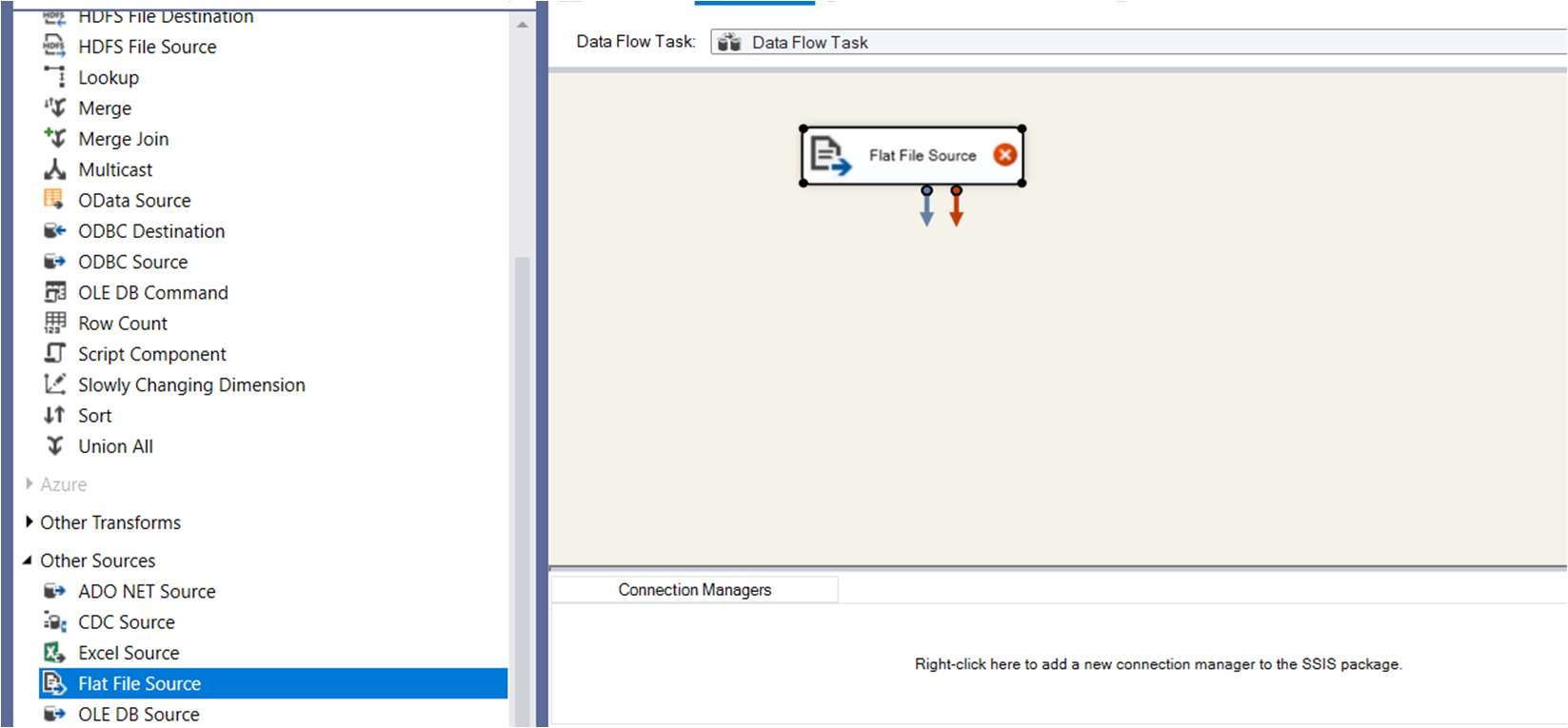
* Select Integration Service and click on Next.



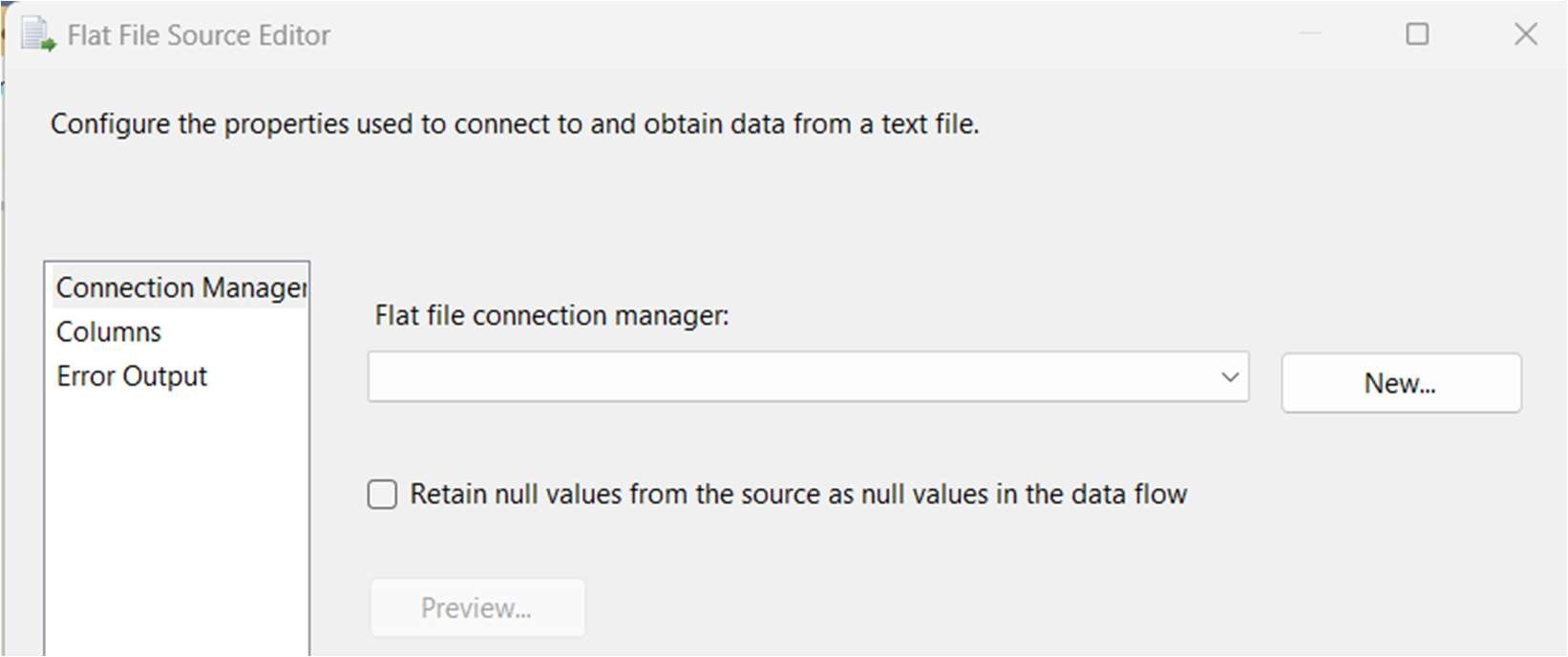
* Give a Name and click on create.
* This is our Integration Service package, drag and drop the Data Flow task into the package.



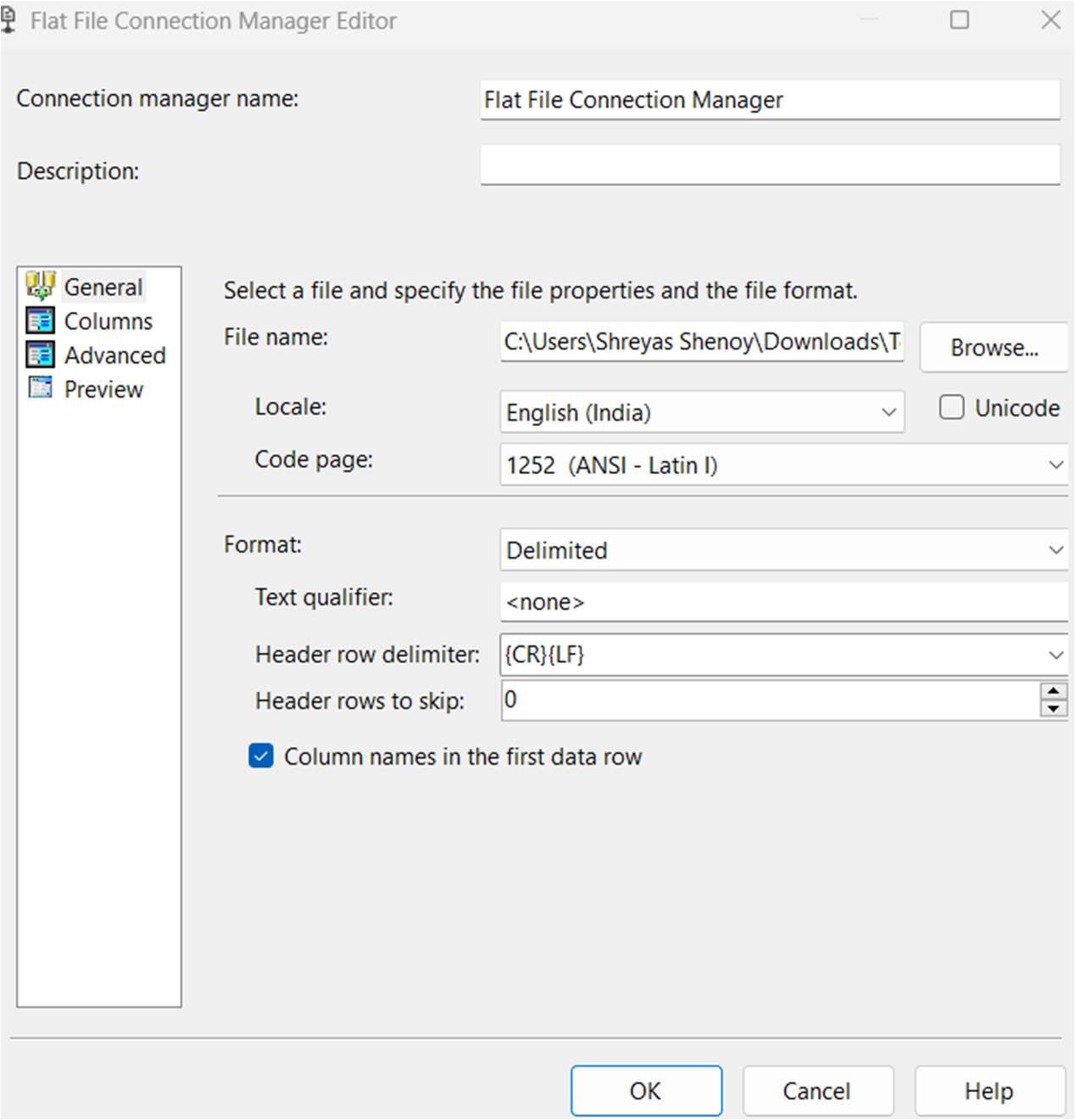
* Double-click on it and drag and drop the Flat File source into the flow.



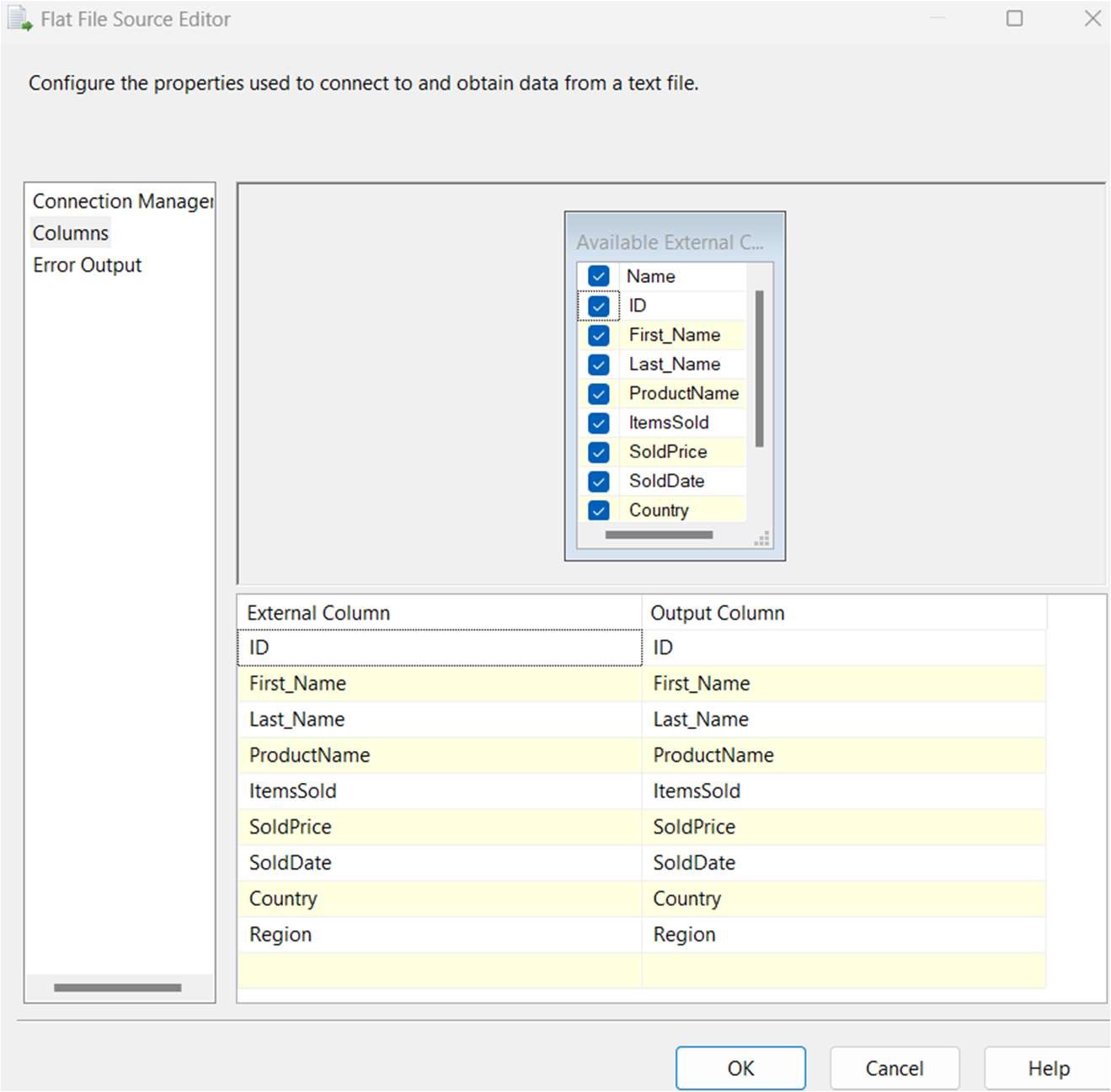
* Double-click on it, and Click on New.



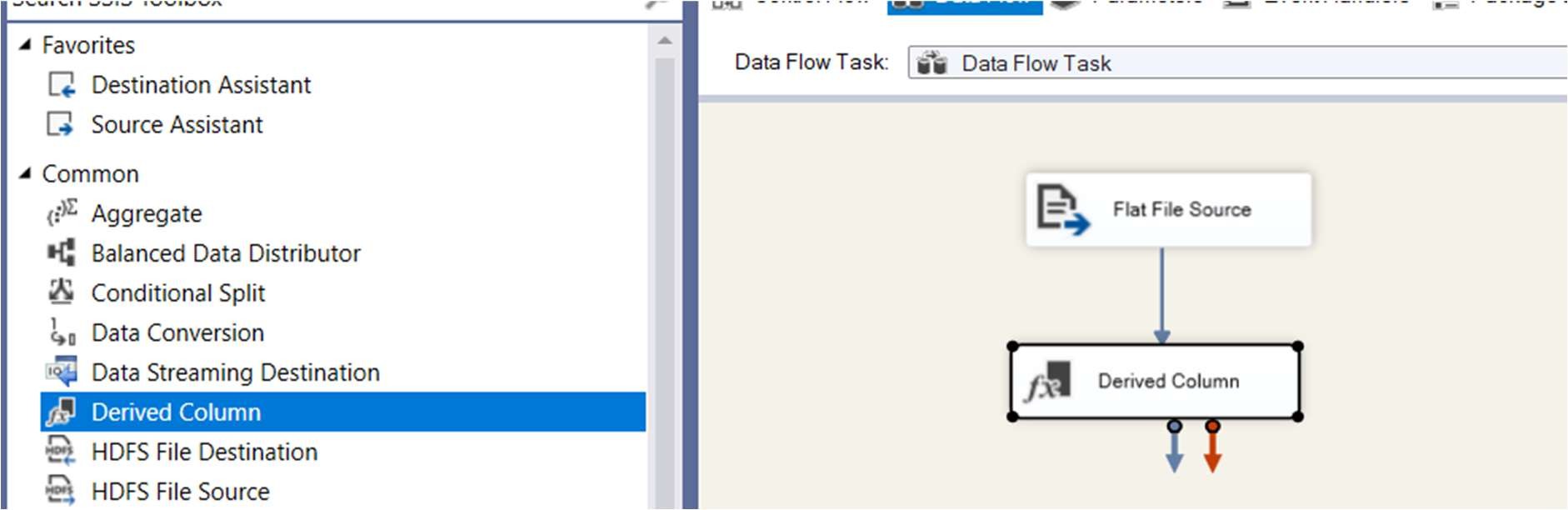
* Browse the file from your local system and click on ok.



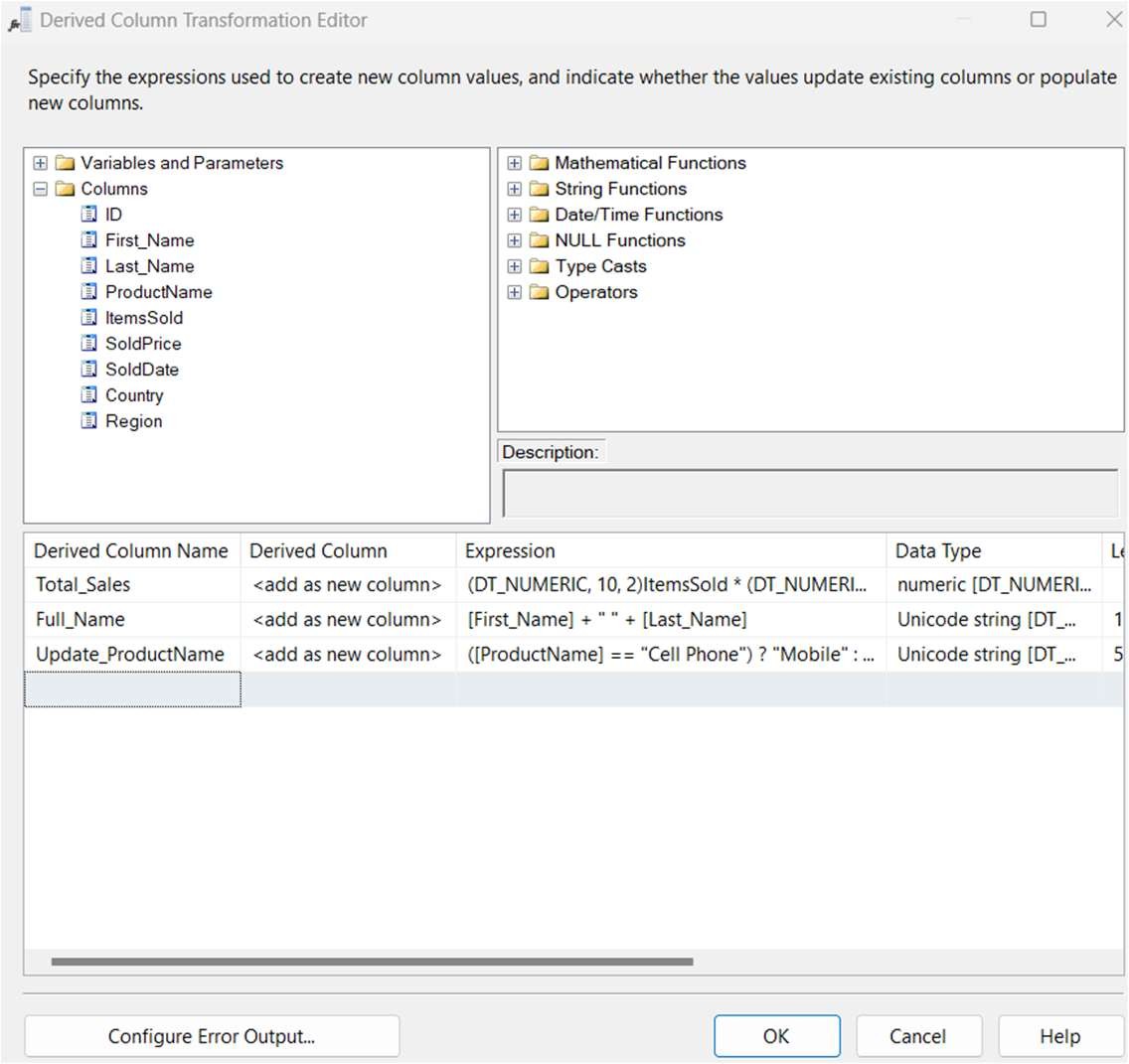
* Next check the columns and click on OK.



* Drag and drop the Derived column and make the connection as shown below.



* Double-click on it. Add three columns as new and give the below expressions. Total\_Sales: (DT\_NUMERIC, 10, 2)ItemsSold \* (DT\_NUMERIC, 10, 2)SoldPrice Full\_Name: [First\_Name] + " " + [Last\_Name]
  + Update\_ProductName: ([ProductName] == "Cell Phone") ? "Mobile" : [ProductName]

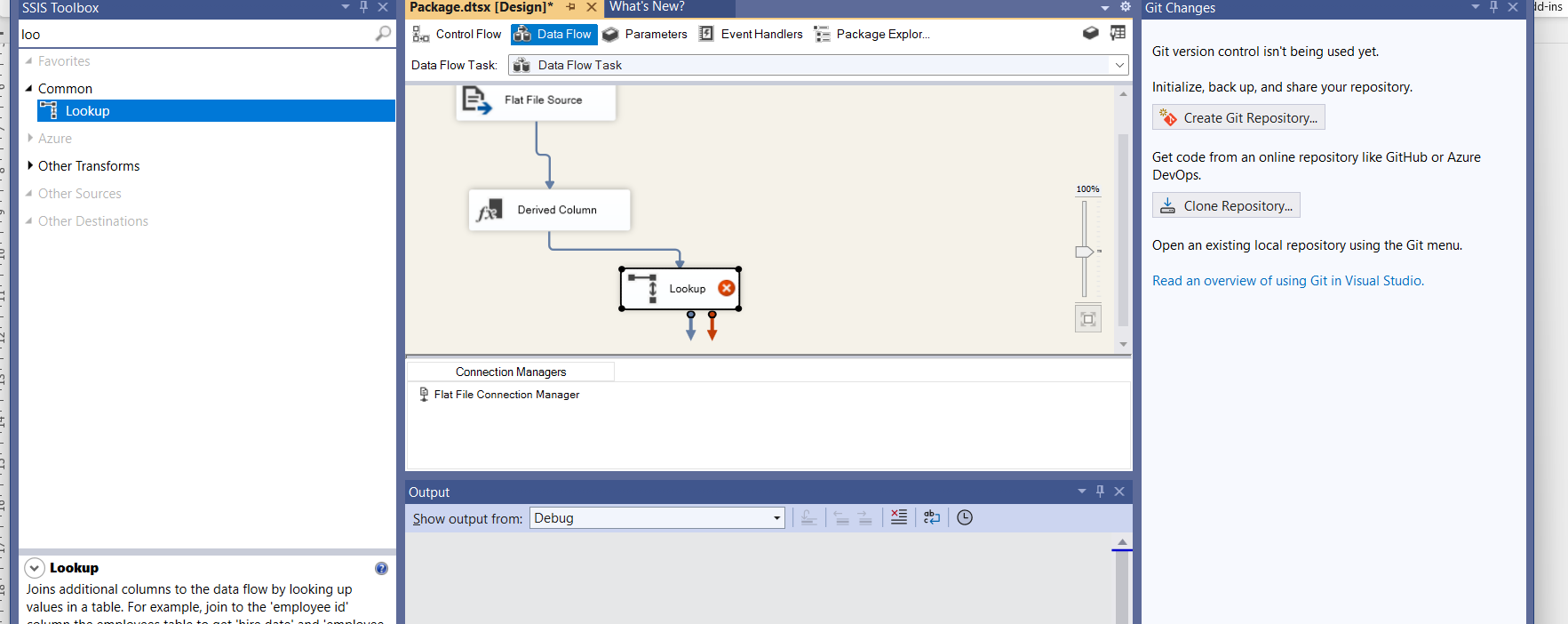


* + Add Lookup Transformation (For Region)
  + Add Lookup Transformation:
  + Drag a Lookup Transformation into the designer and connect it to the output of the Derived Column.
  + In the Lookup transformation editor, choose to perform a lookup based on the Country (e.g., to lookup missing Region information).

Configure Lookup:

* + Choose an OLE DB Connection or another file source where your lookup table is stored.
  + Map the Country column from your source to your lookup reference.
  + Map the resulting Region column back into the data flow.

Output: This will ensure that all records have the correct Region



In the SSMS you need to create database name region

Use database Region

CREATE TABLE Region (

Country VARCHAR(100) PRIMARY KEY,

Region VARCHAR(100)

);

INSERT INTO Region (Country, Region)

VALUES

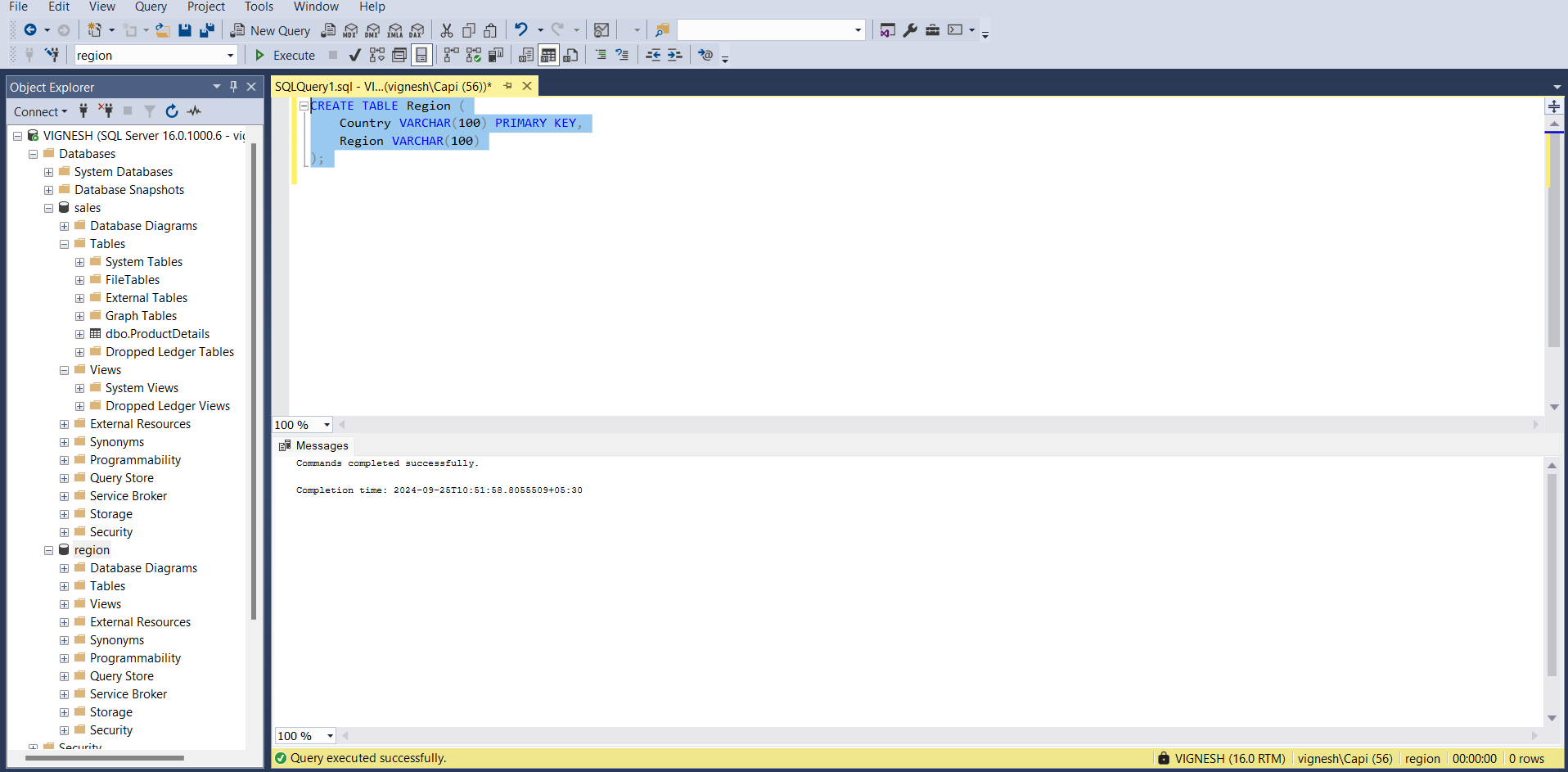
('USA', 'North America'),

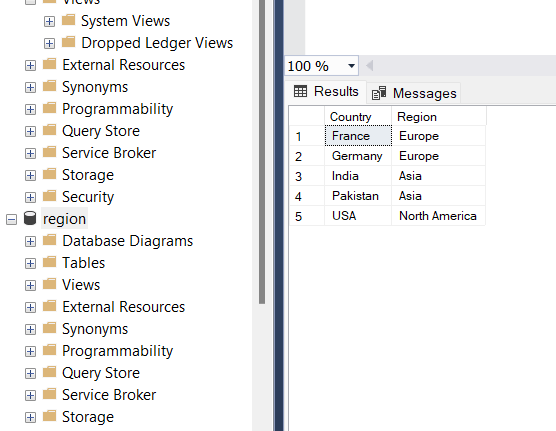
('Pakistan', 'Asia'),

('India', 'Asia'),

('France', 'Europe'),

('Germany', 'Europe');





* + Select the appropriate table or reference file where your Country to Region mapping is stored (e.g., CountryRegionReference).

Configure Lookup Columns:

* + Go to the Columns tab in the Lookup editor.
  + Under Available Input Columns, check the Country column from your input (main data flow).
  + Under Available Lookup Columns, check the Region column from your reference table (CountryRegionReference).

Set the Lookup Matching Condition:

* + In the same Columns tab, set the Lookup Condition:
    - Map the Country from your input columns to Country in the lookup reference table.
    - This will ensure the Lookup matches based on the Country.

Handle Lookup Failures:

* + In the General tab, configure what should happen if a lookup fails:
    - Redirect rows to a failure output: If no match is found, you can send these rows to a different path for further handling.
    - Use a default value: For instance, you can use a default region if no matching record is found in the lookup table.

11 )Map the Resulting Region into Data Flow

Output the Region Column:

* + After configuring the lookup, the Region column from the reference table will now be available as part of the data flow.
  + In the Columns tab, make sure the Region column is checked for inclusion in the output.

Replace or Add the Region Column:

* + You can decide whether to replace the existing Region column (if one exists in the input) or add the Region column from the lookup as a new column.
    - If there is an existing Region column, rename the new Region column from the lookup in the Derived Column transformation or within the Mapping in the destination.
  + If the existing Region column in your source is empty or incomplete, you can use the Lookup Region column to fill in the missing data.
  + In the Derived Column transformation, you can add logic like:

ISNULL([Region]) ? [Lookup\_Region] : [Region]

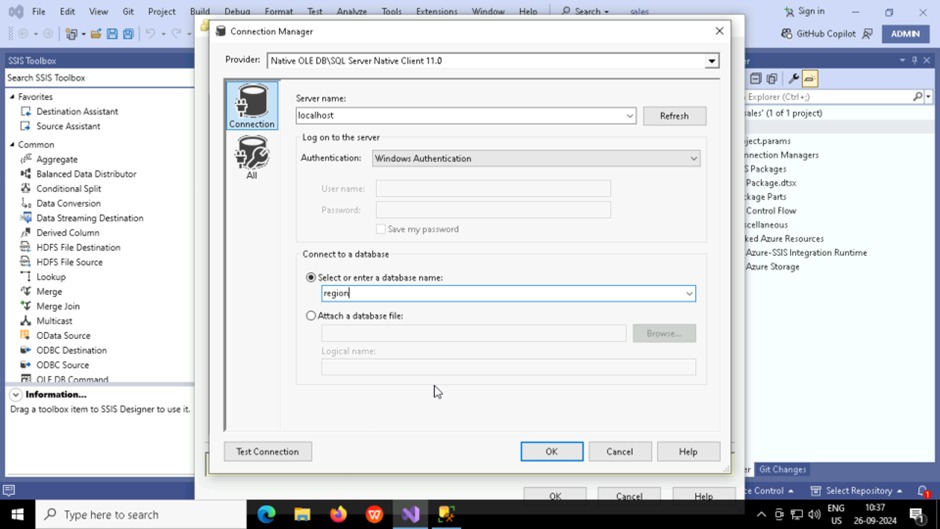
SSIS Lookup Transformation for Country to Region

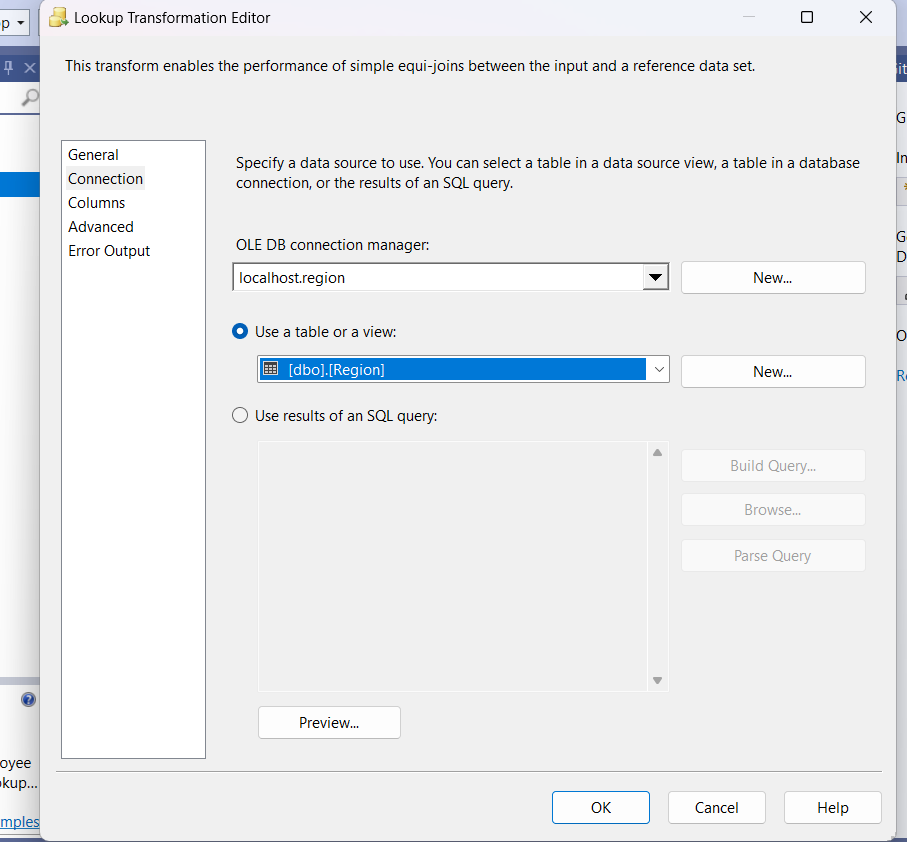
Example Input Data Before Lookup:

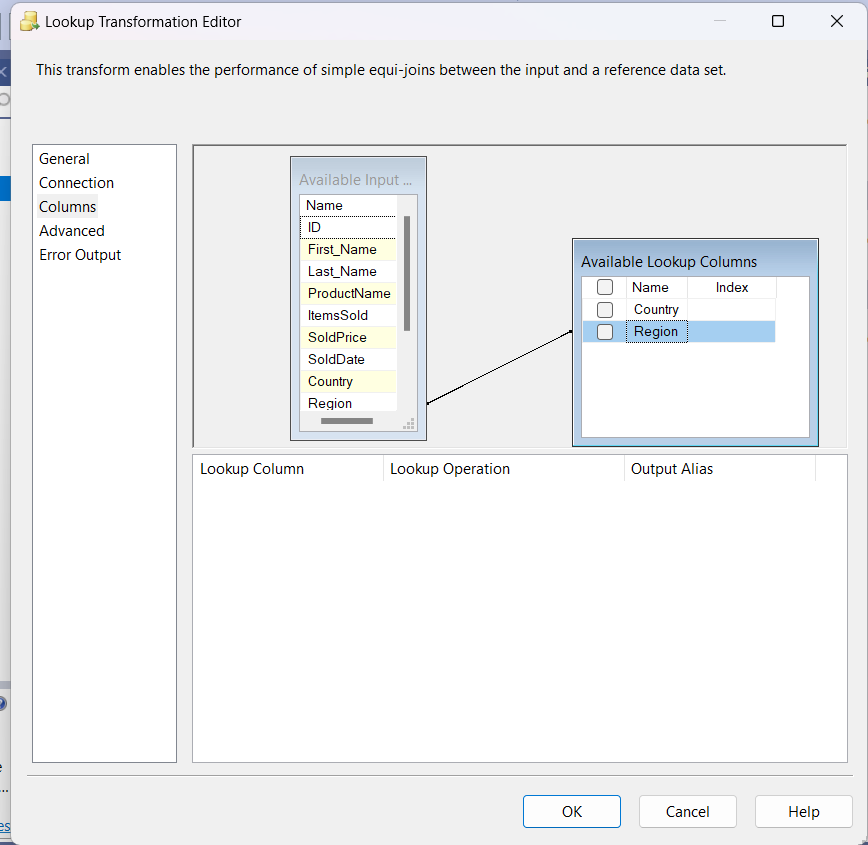
| ID | First\_Name | Last\_Name | ProductName | ItemsSold | SoldPrice | Country | Region (incomplete) |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | Aamir | Shahzad | TV | 1 | 700 | USA | NULL |
| 2 | M | Raza | Cell Phone | 2 | 800 | USA | NULL |
| 3 | Christy | Ladson | TV | 3 | 1600 | USA | NULL |
| 4 | John | Rivers | Laptop | 5 | 2400 | USA | NULL |
| 5 | Najaf | Ali | Computer | 1 | 300 | Pakistan | NULL |

Example Output After Lookup Transformation:

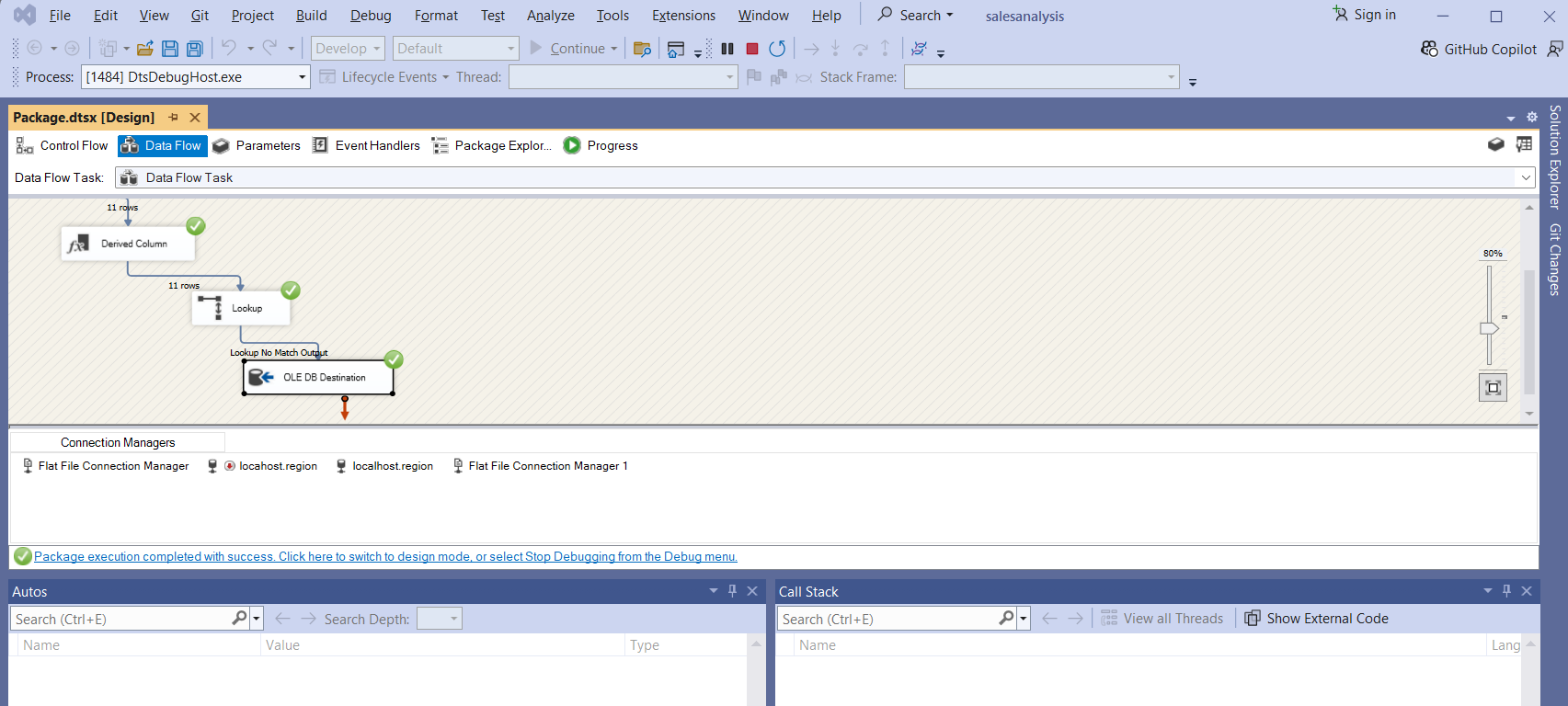
| ID | First\_Name | Last\_Name | ProductName | ItemsSold | SoldPrice | Country | Region |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | Aamir | Shahzad | TV | 1 | 700 | USA | North America |
| 2 | M | Raza | Mobiles | 2 | 800 | USA | North America |
| 3 | Christy | Ladson | TV | 3 | 1600 | USA | North America |
| 4 | John | Rivers | Laptop | 5 | 2400 | USA | North America |
| 5 | Najaf | Ali | Computer | 1 | 300 | Pakistan | Asia |

 To connect to the database you use the server name localhost and select windows authentication give the database as “**region** “





Verify the lookup data in the table after the lookup is performed

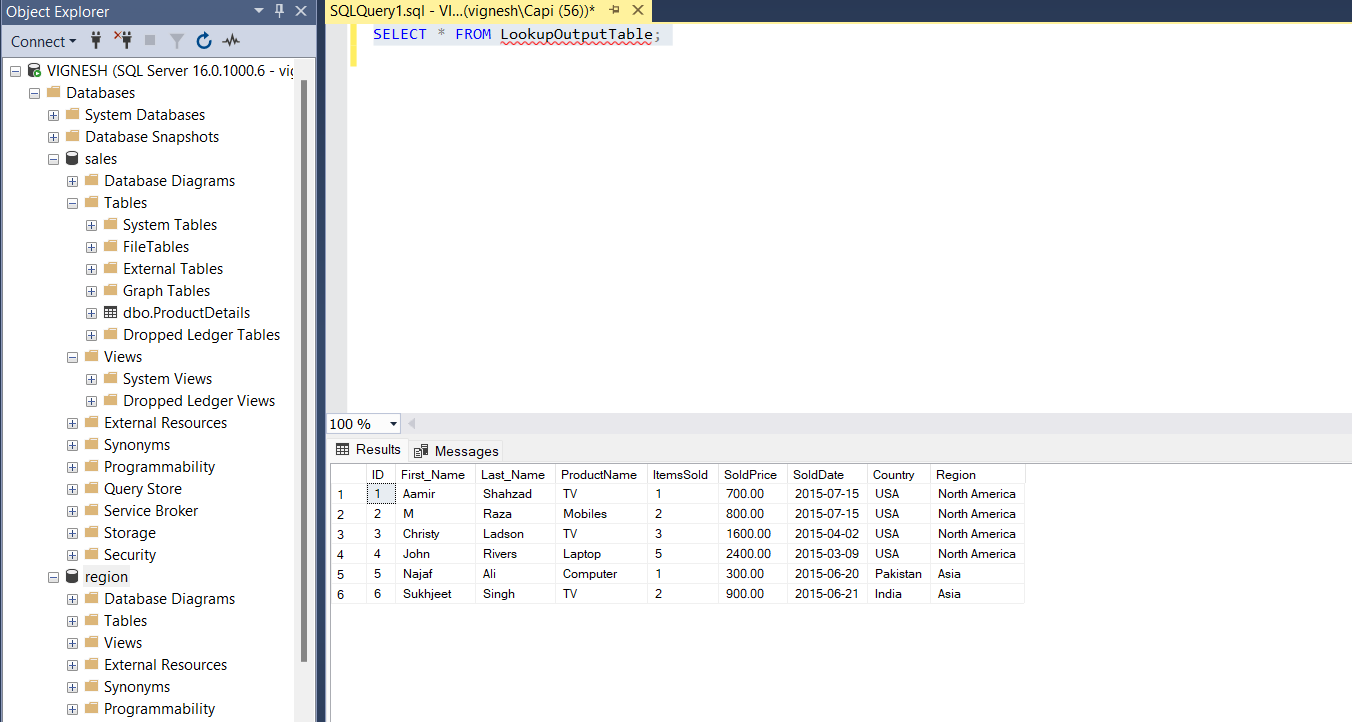


Add an OLE DB Destination:

* After your Lookup Transformation, drag an OLE DB Destination into the data flow.
* Connect the output of the Lookup to the OLE DB Destination.

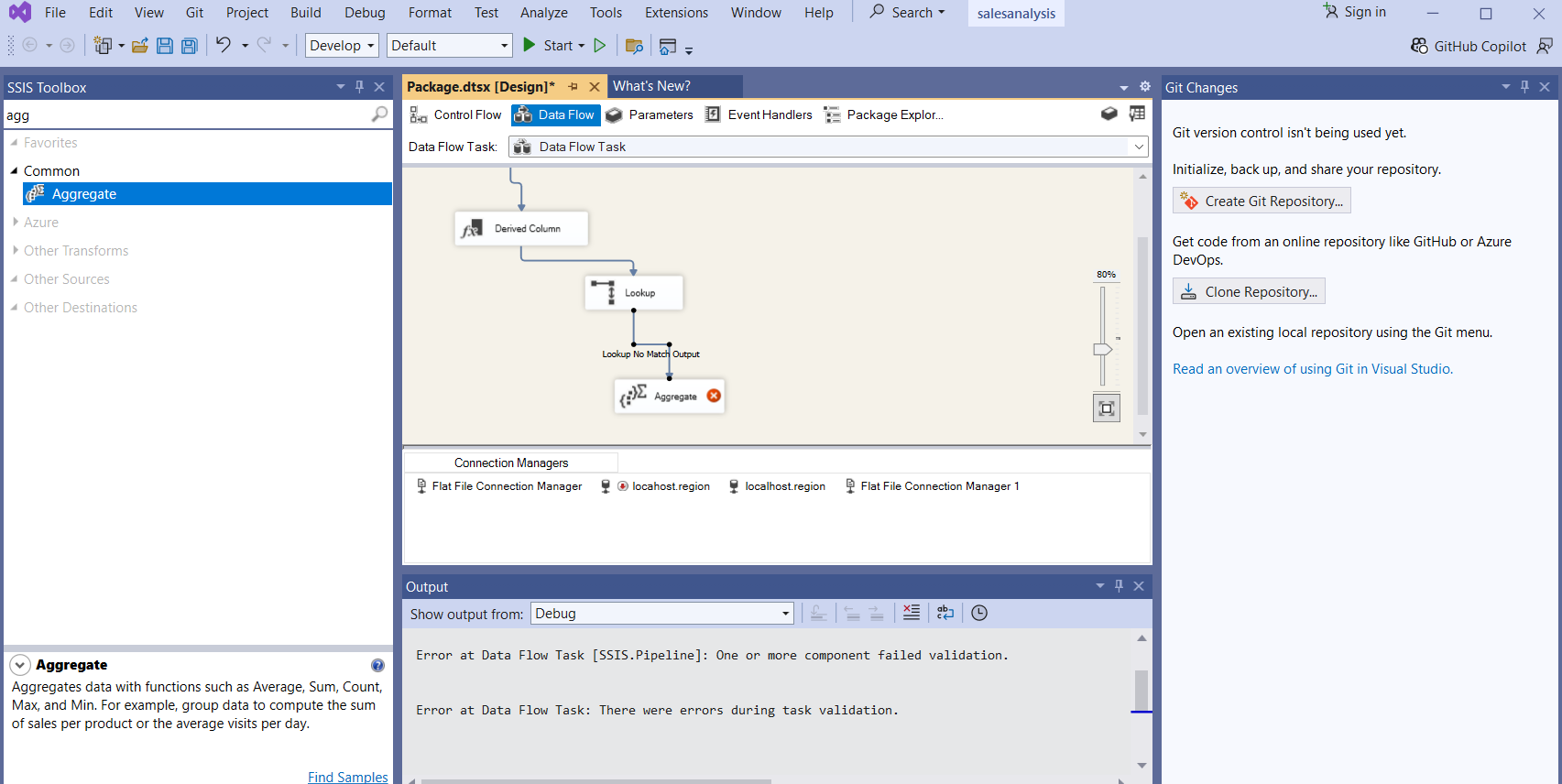
Configure OLE DB Destination:

* Select the LookupOutputTable as the destination table.
* Map the columns from the Lookup Transformation output to the respective fields in the LookupOutputTable.

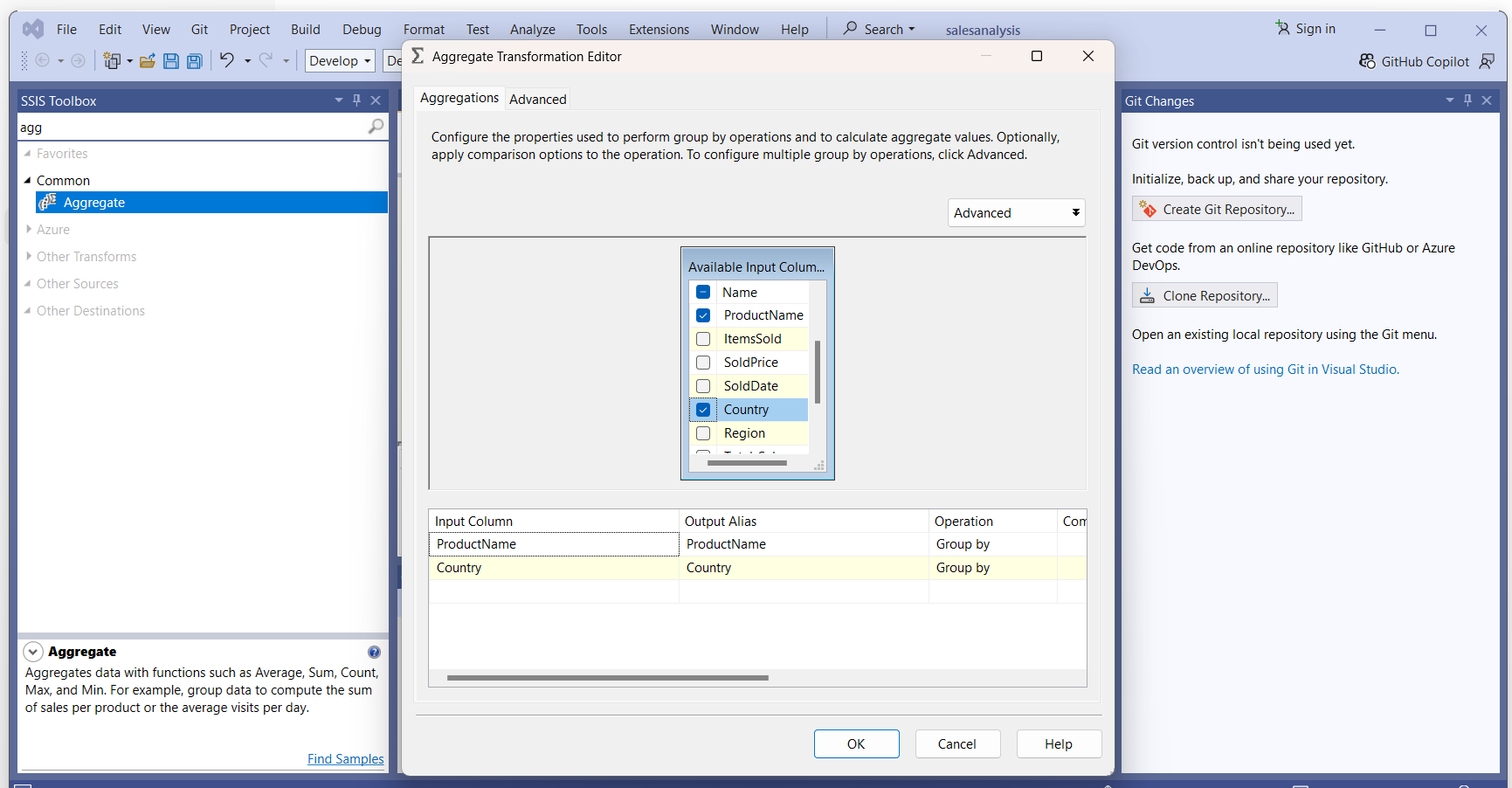


Add Aggregate Transformation (For Totaling Sales)

Add Aggregate Transformation:



* + Drag an Aggregate Transformation onto the designer and connect it to the Lookup Transformation.
  + In the Aggregate Transformation, group by:
    - ProductName
    - Country



* + Add the following operations:
    - Total\_Items\_Sold: Sum the ItemsSold.
    - Total\_Sales: Sum the calculated Total\_Sales.

Output: This step calculates the total sales and items sold for each product by country

12 Add Conditional Split (For Region-Based Separation)

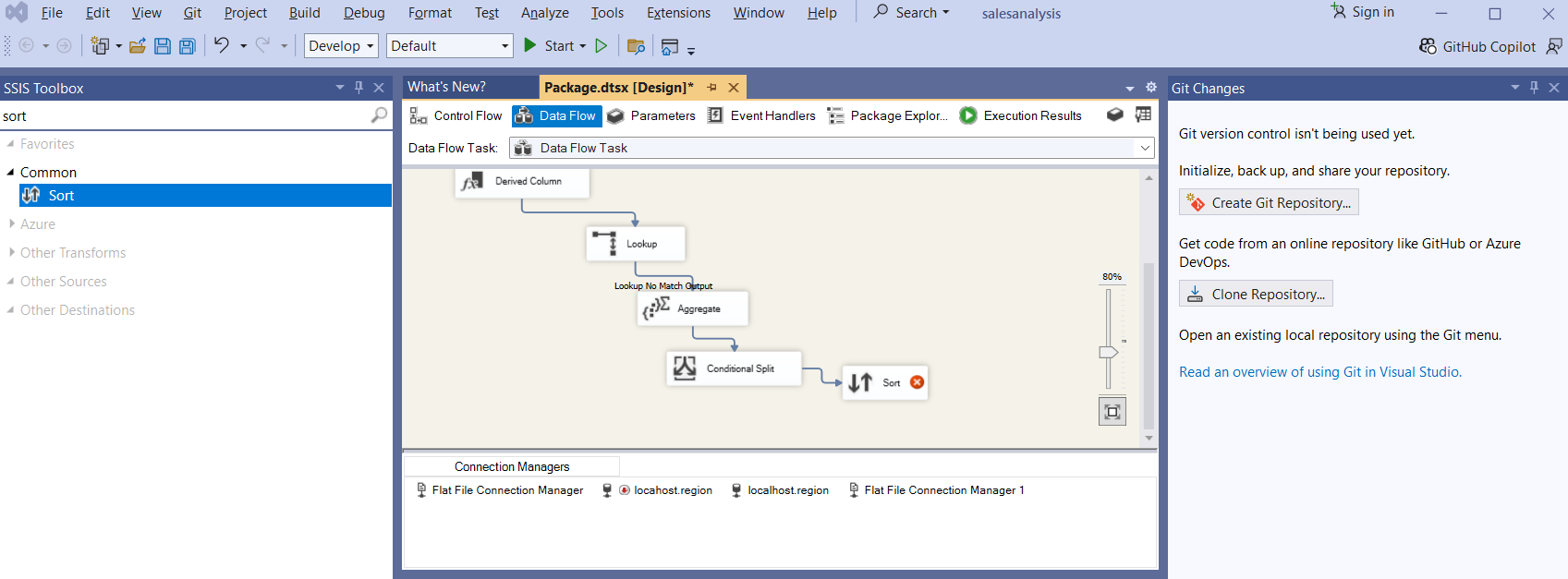
Add Conditional Split Transformation:

* + Drag a Conditional Split transformation onto the designer and connect it to the output of the Aggregate Transformation.

Configure Conditional Split:

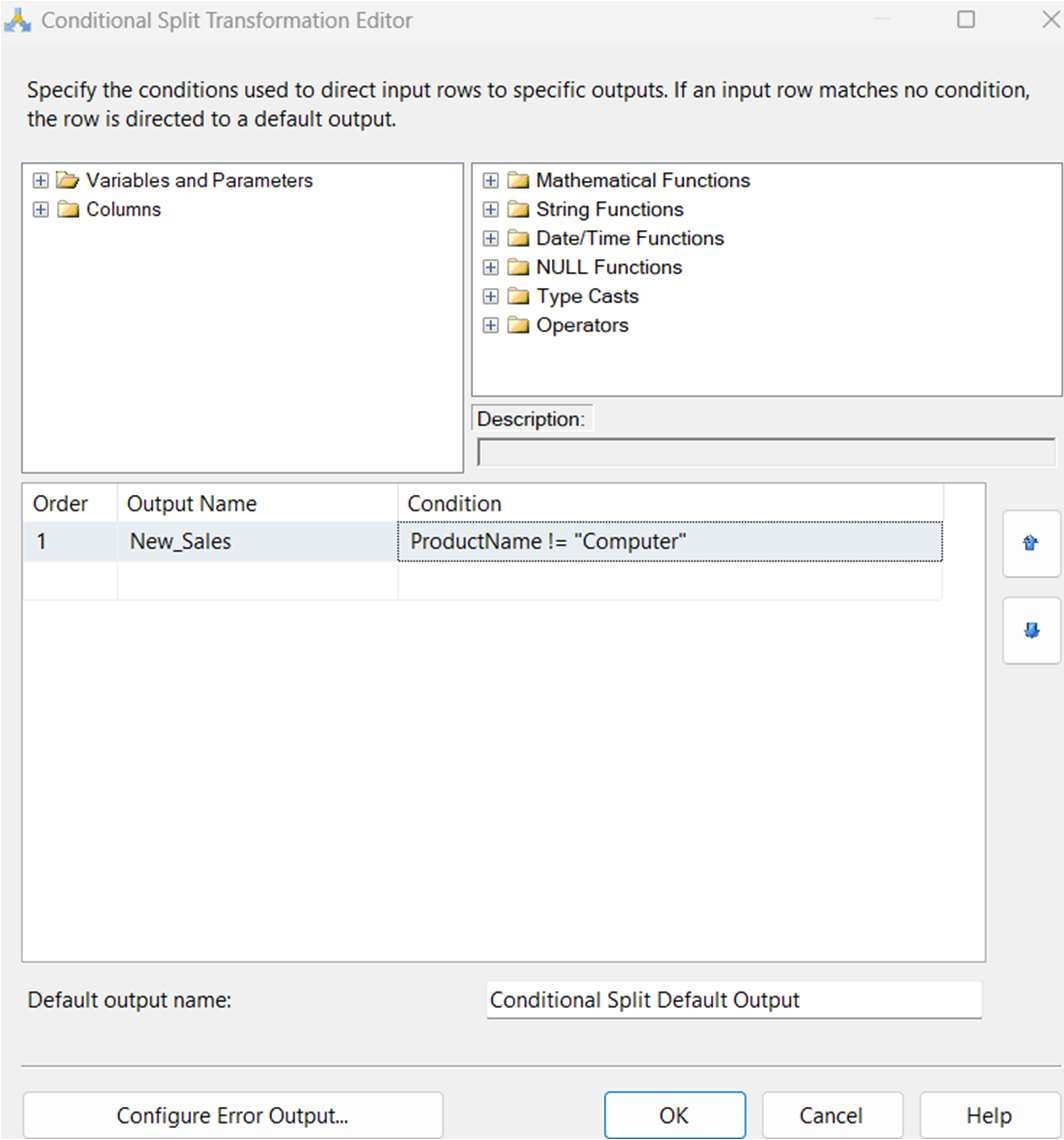
* + Add conditions to split the data based on the Region:
    - Region == "North America" -> North America Output
    - Region == "Asia" -> Asia Output
    - Region == "Europe" -> Europe Output

Drag and drop the Conditional Split and make the connection.

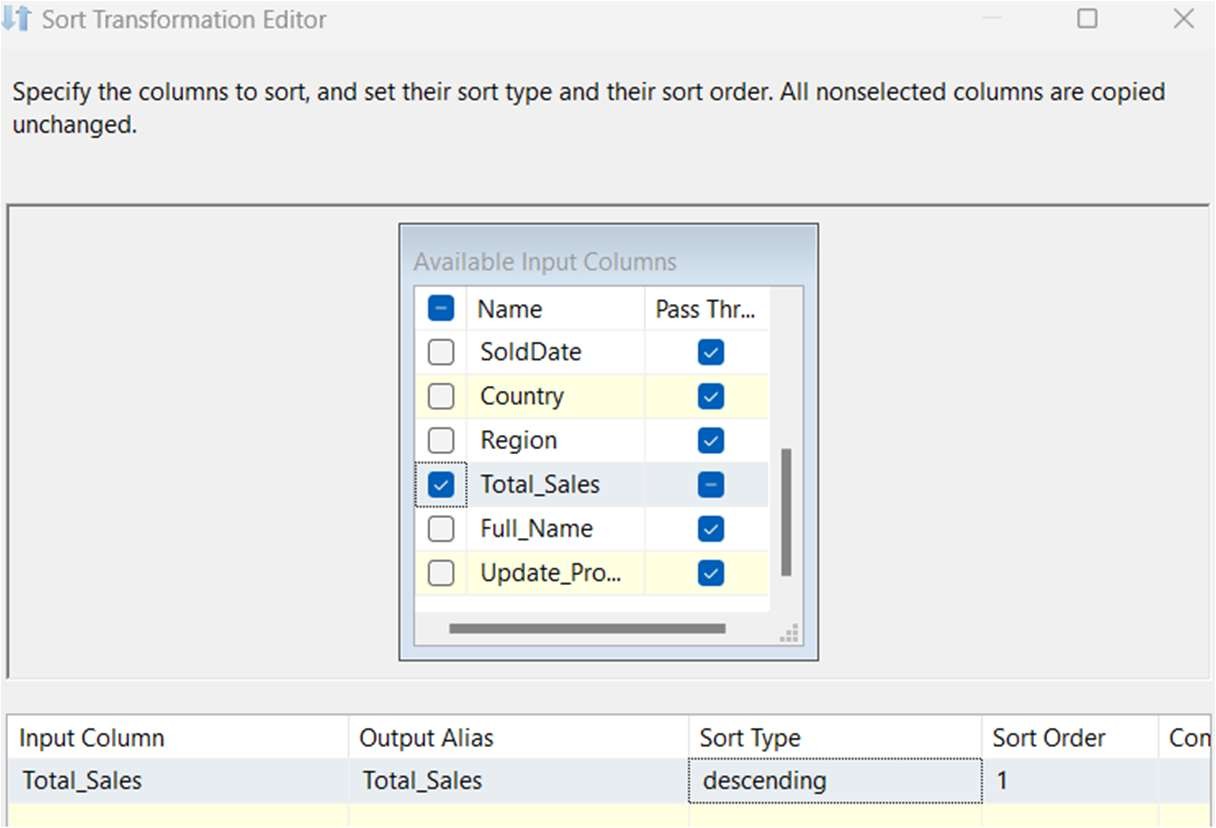


Double-click on it and give an output and condition as shown below

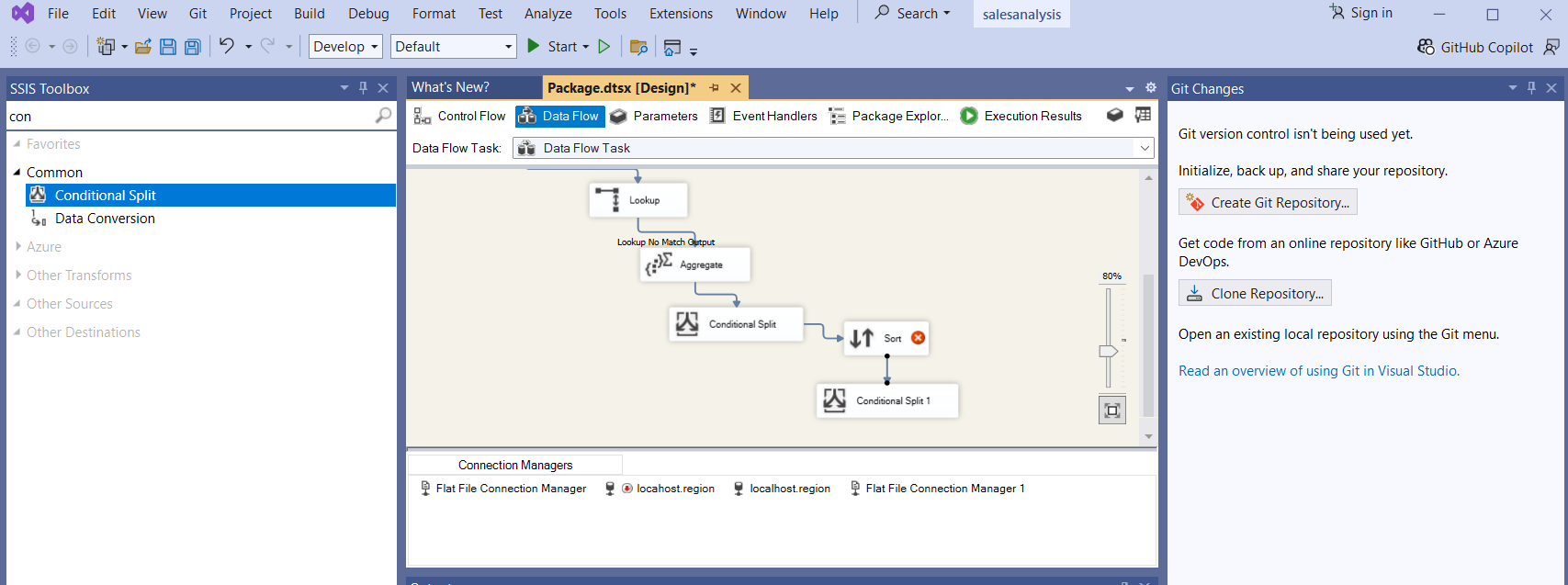
. Exp: ProductName != "Computer"



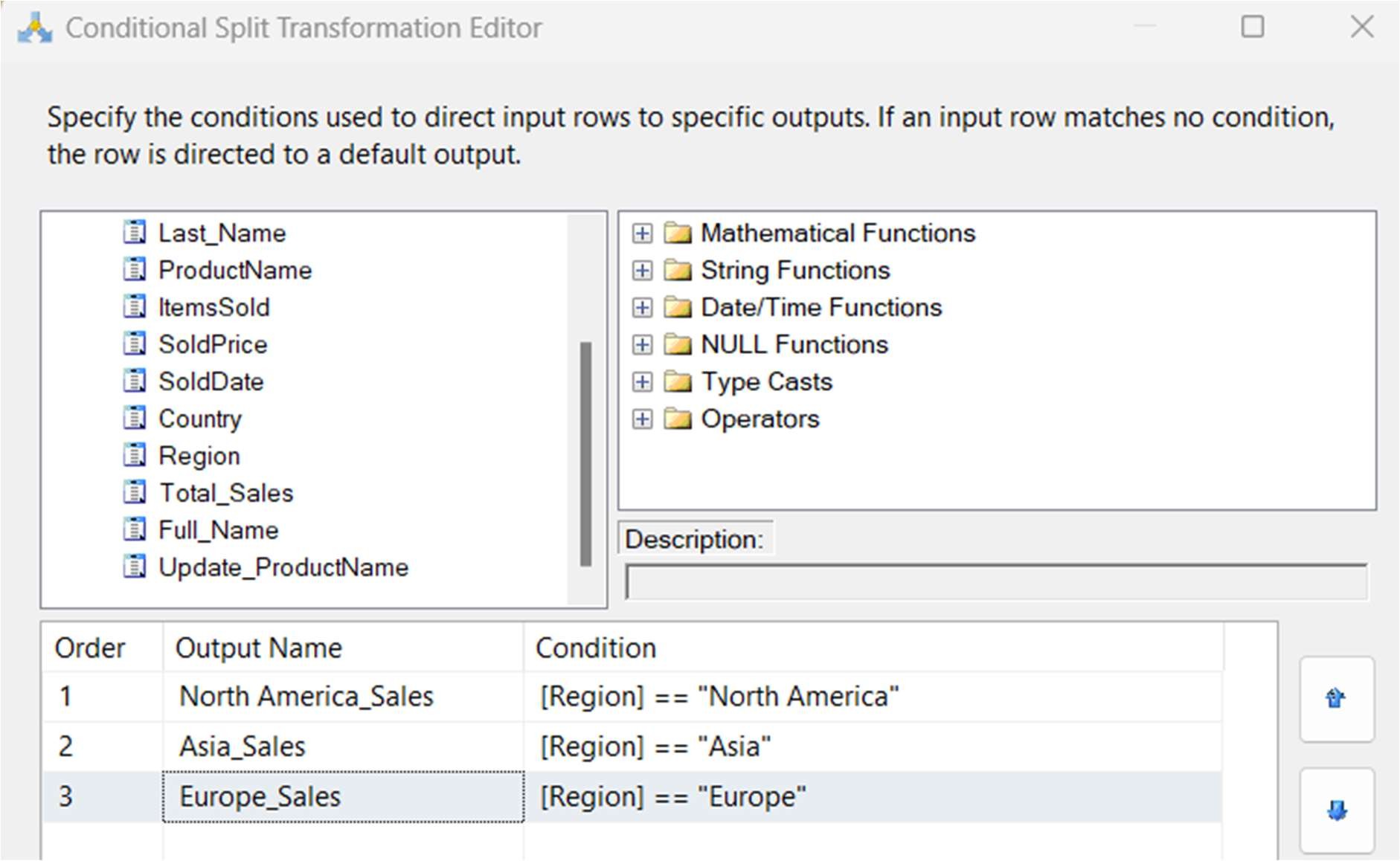
* + Drag and drop the Sort and make the connection as shown below.
    - Double-click on it and set the condition.

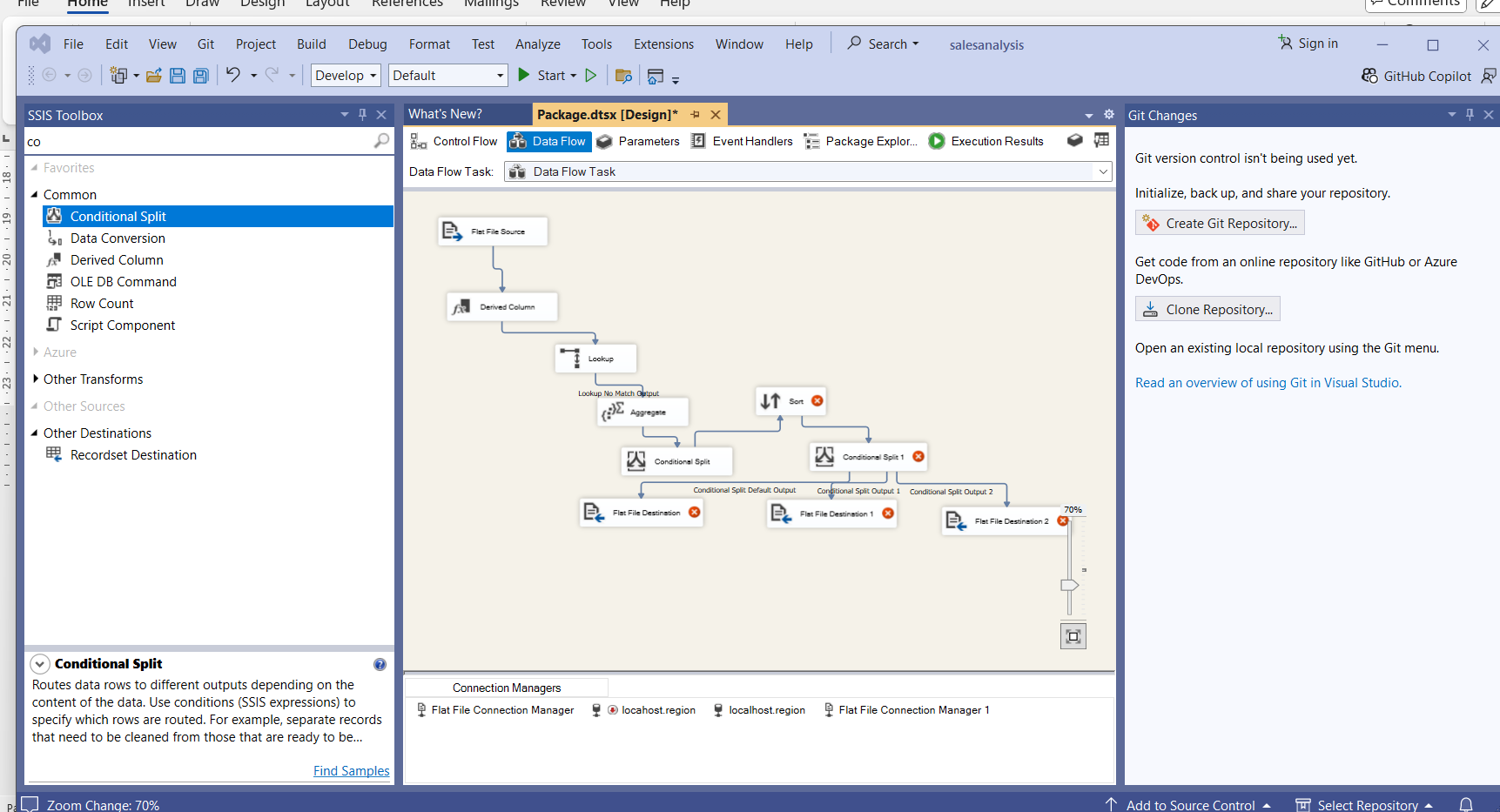


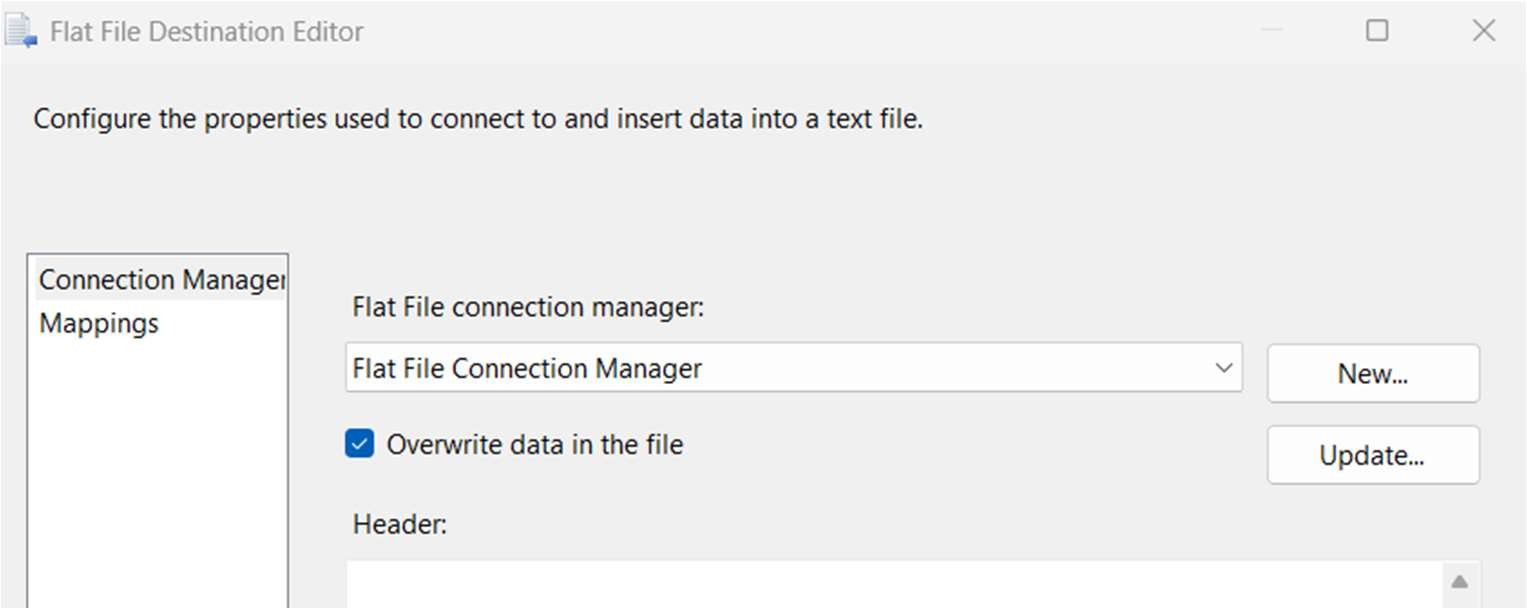
Drag and drop the conditional split again and make the connection as shown below.



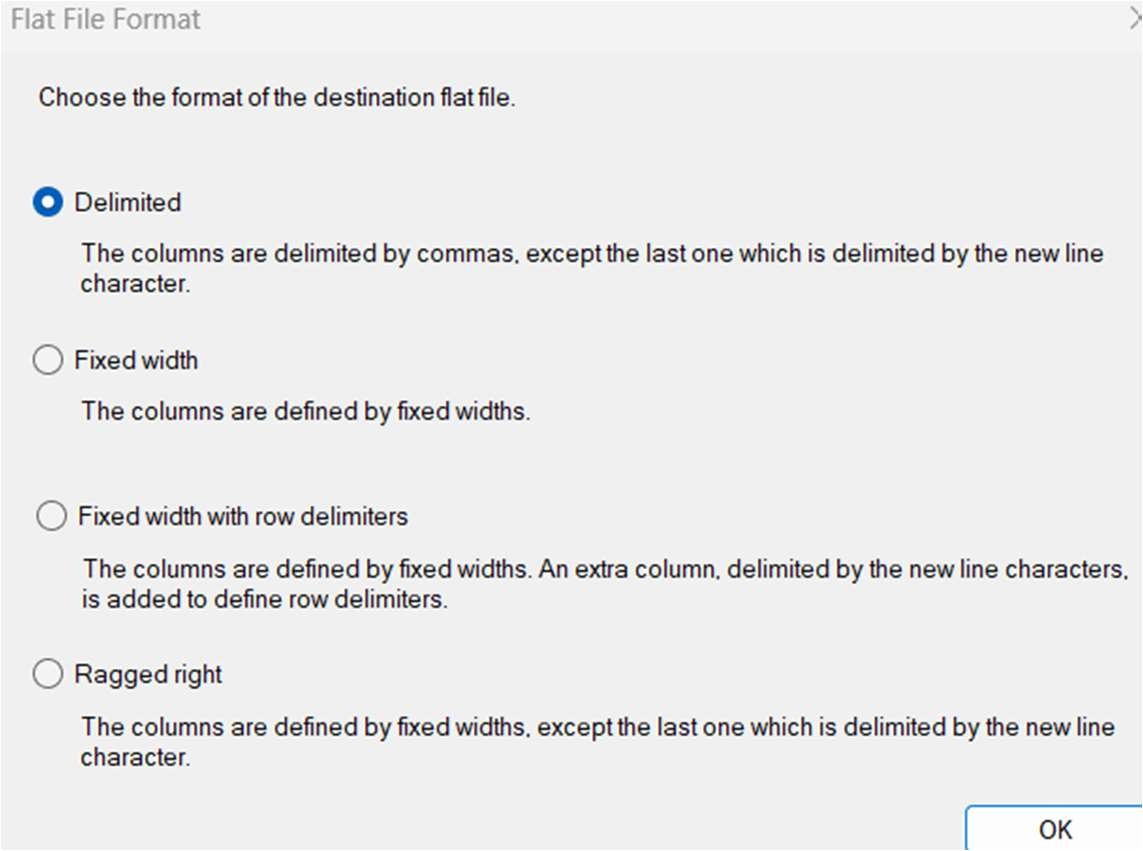
Double-click on it and add below three output conditions.



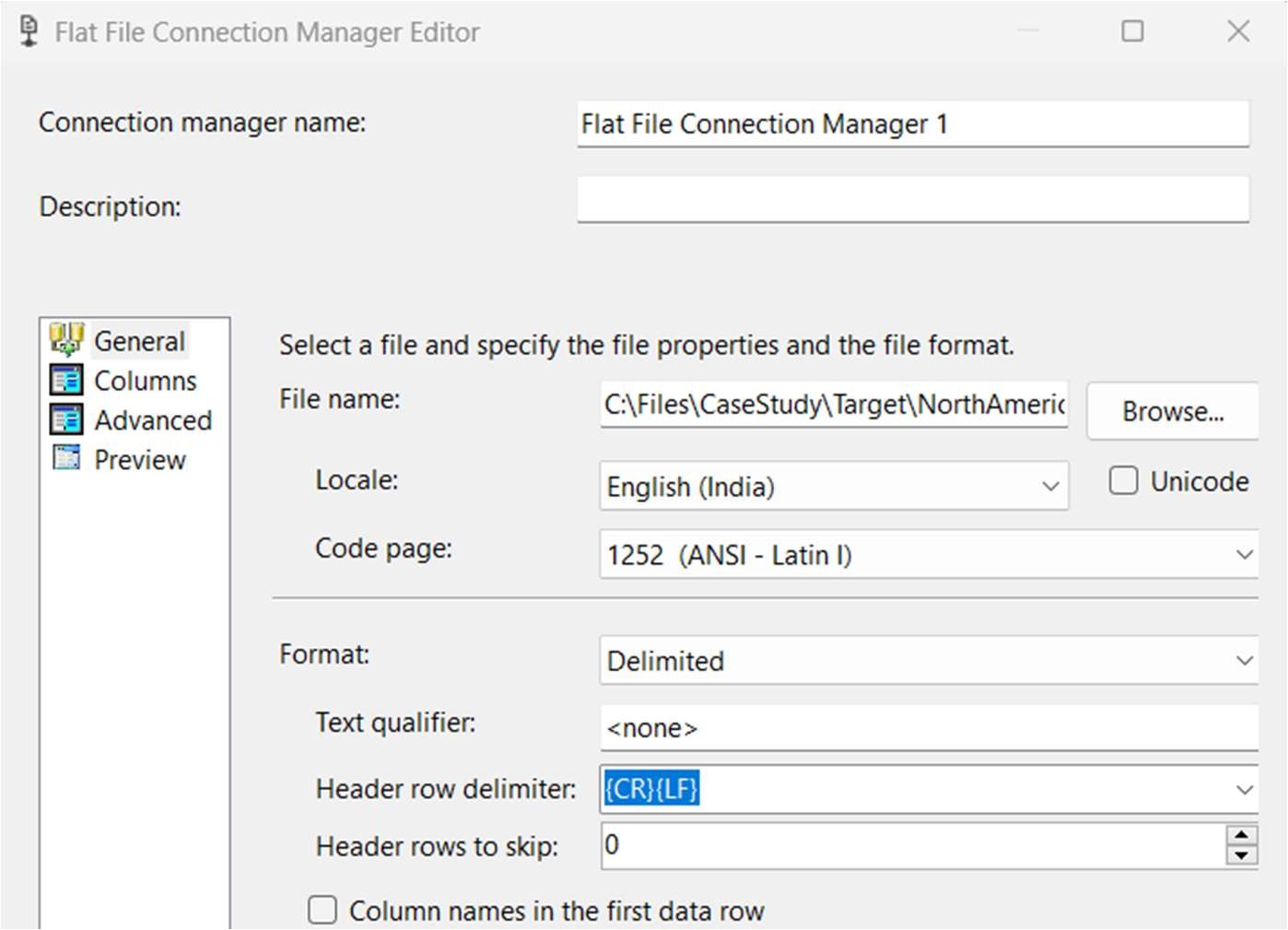
* Drag and drop three flat file destinations and make connections as shown below.
* 
* Double-click on the first destination and click on New.



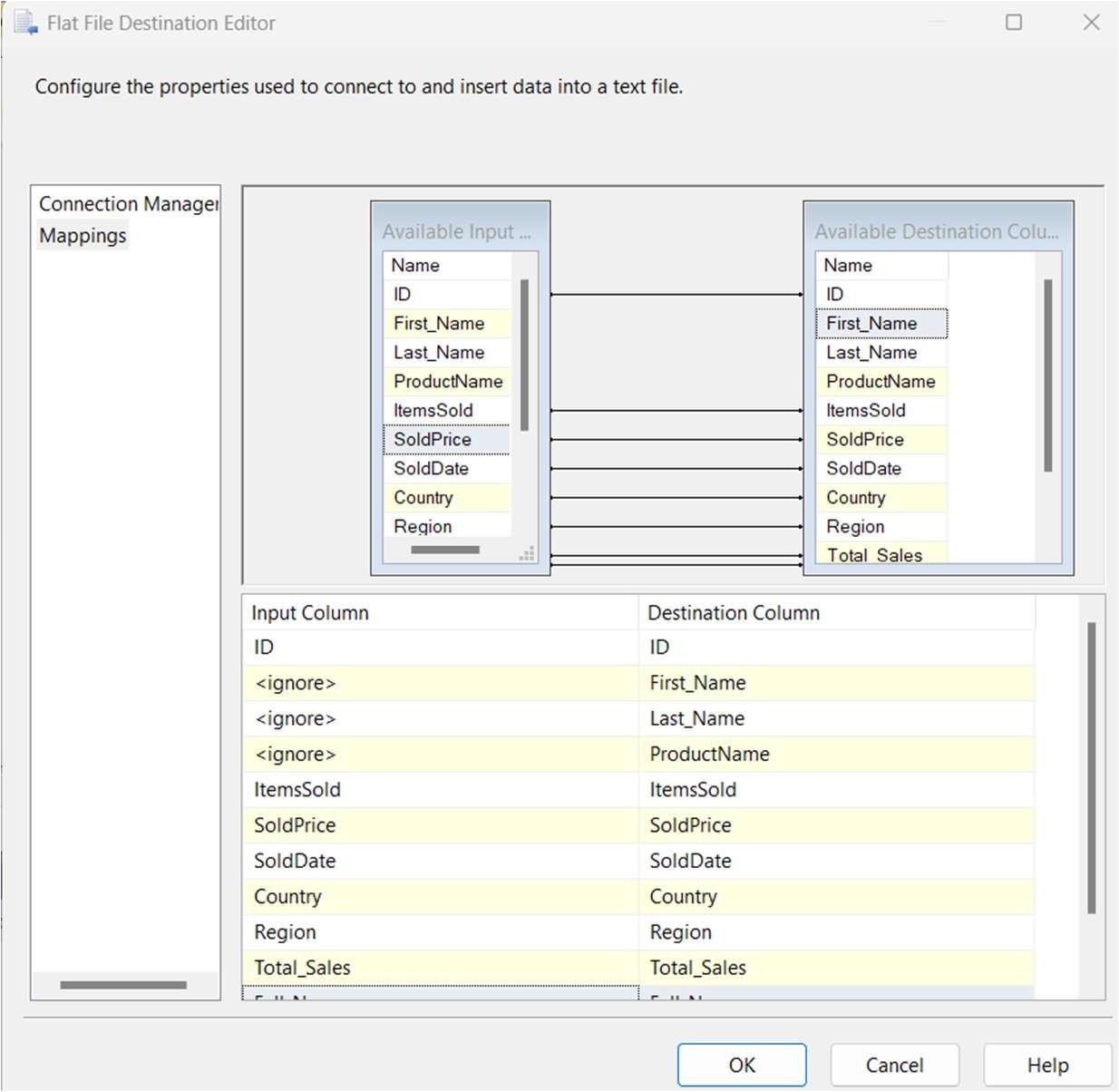
* Click on OK.



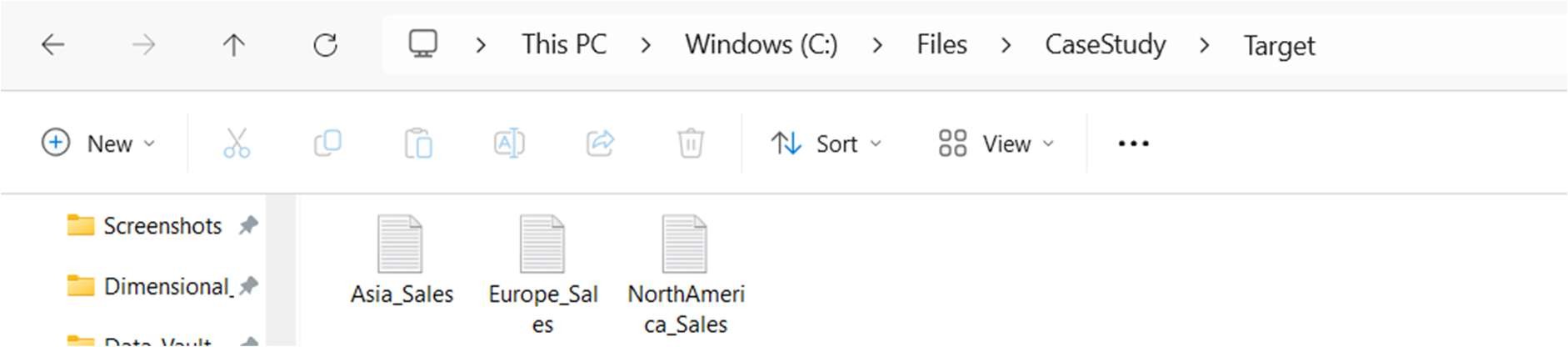
* Create a File in your local system with the North American sales.



* Under Advance delete the First Name, Last Name, and Product Name columns.
* Click OK. Under Mapping Check the columns.



* Do the Same for Other Destinations also.
* Next save the package and click on Run.
* Here our flow was executed successfully.
* Here we have three text files created.



* Check the data in those files. Export all the table you have created in the SSMS
* Submit the work to the github repository with the proper document