Module 2: SSIS Control and Data Flow

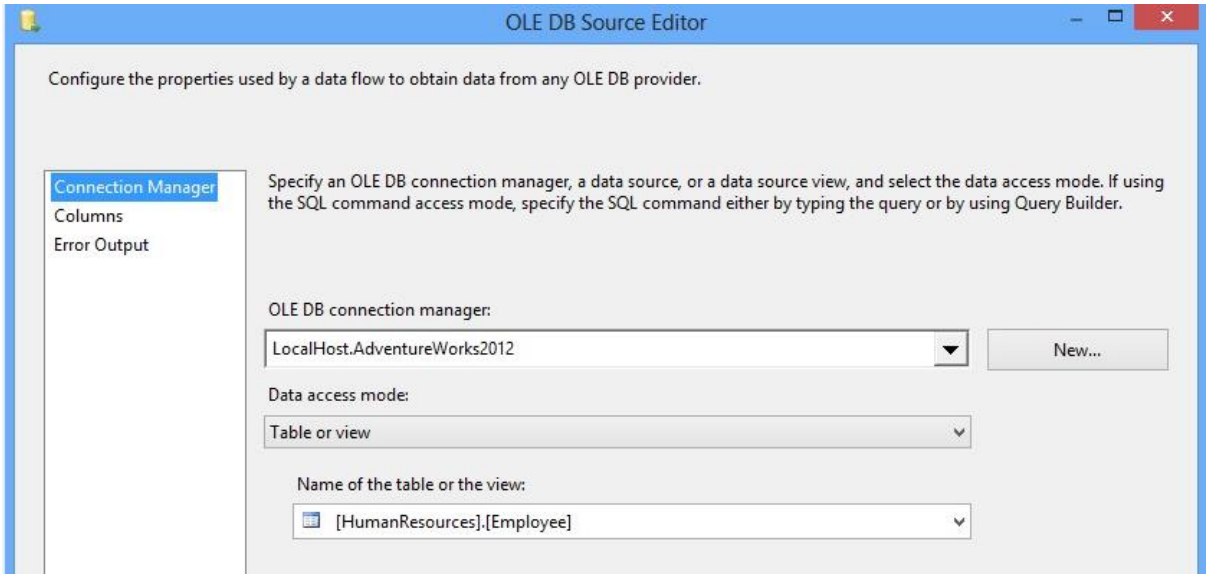
Assignment Solution

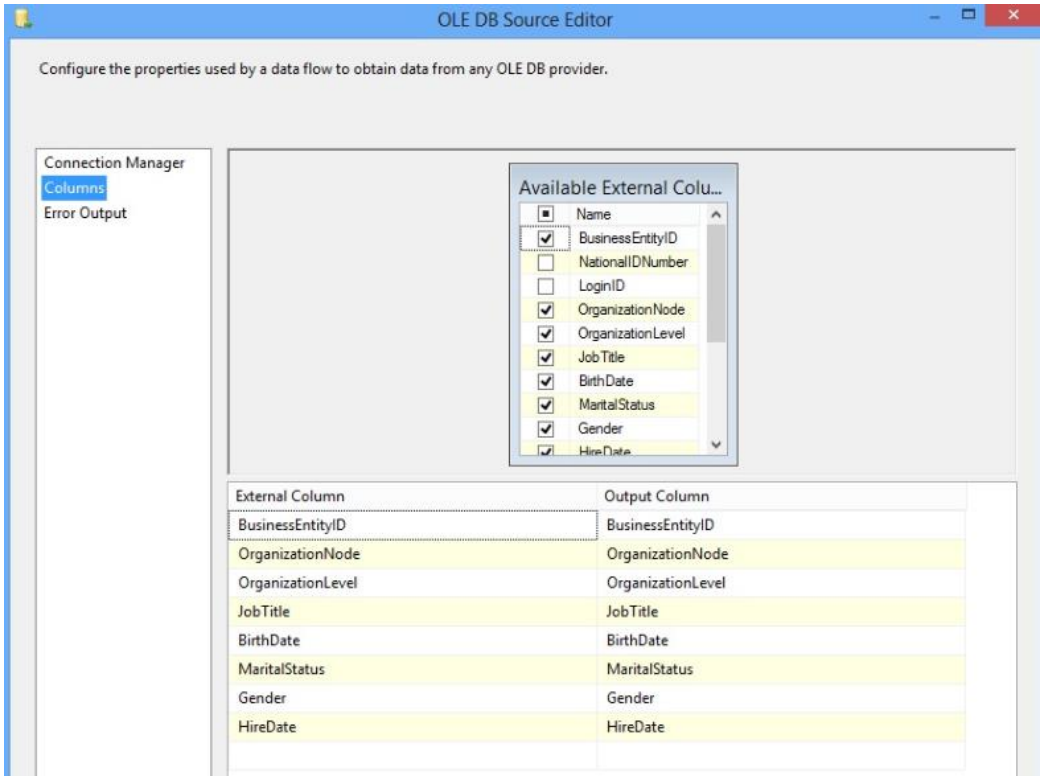
Version 1.0

Solution 1

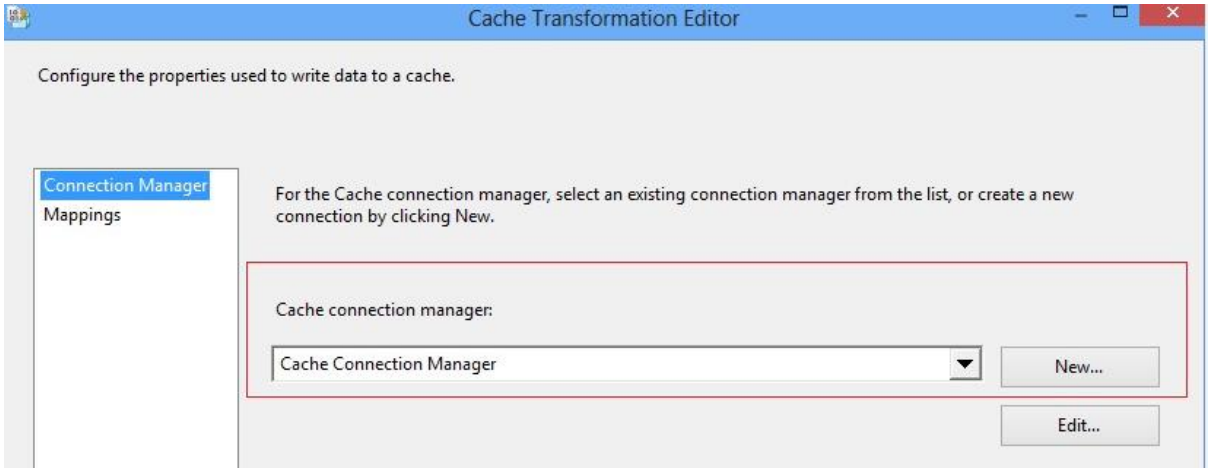
Following are the steps for cache connection:

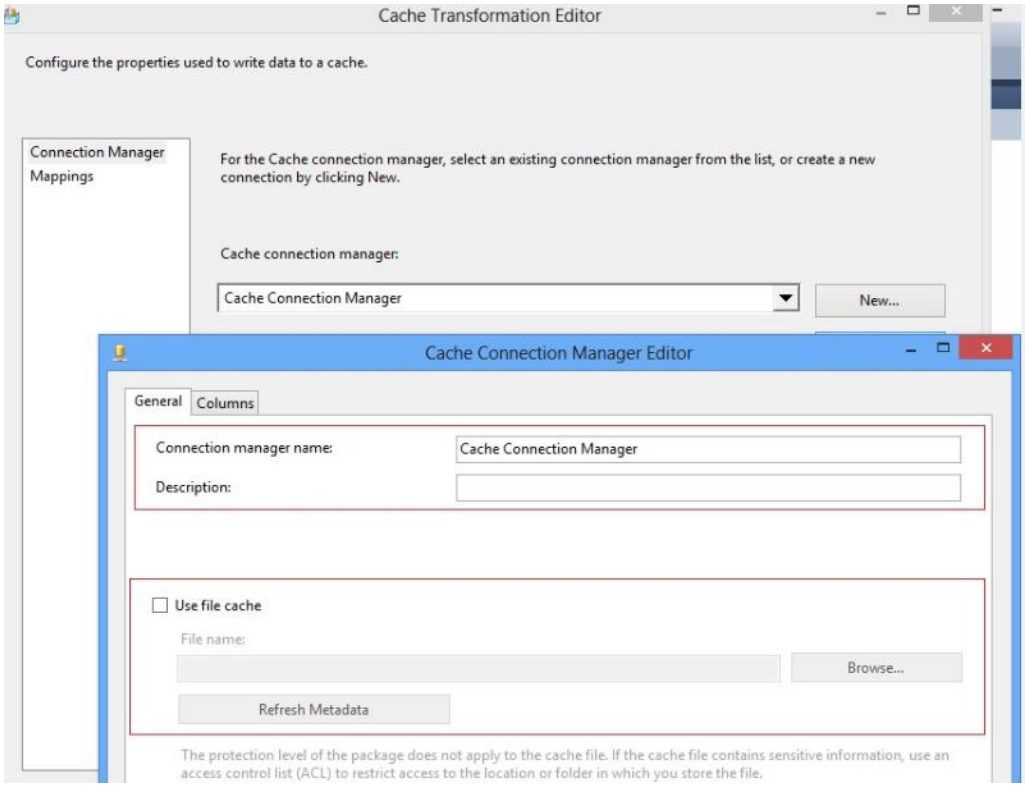
1. Drag and Drop Data Flow Task onto Control Flow and double click on it.
2. In Data Flow Tab, drag and drop OLEDB Source and map it to [HumanResources].[Employee] table of AdventureWorks2012 database.
3. Go to the Columns section, uncheck the columns which are not required and click OK.
4. Drag and drop Cache Transformation onto Data flow and join it to OLEDB Source.
5. Double click on Cache Transformation to configure it.

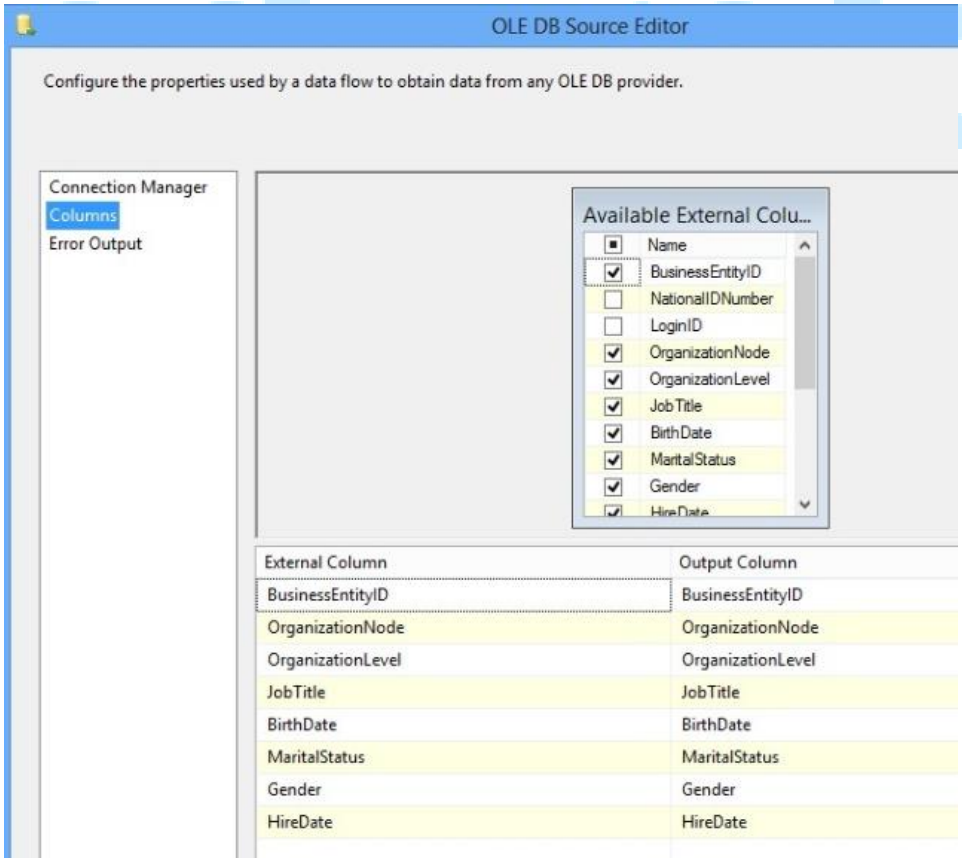




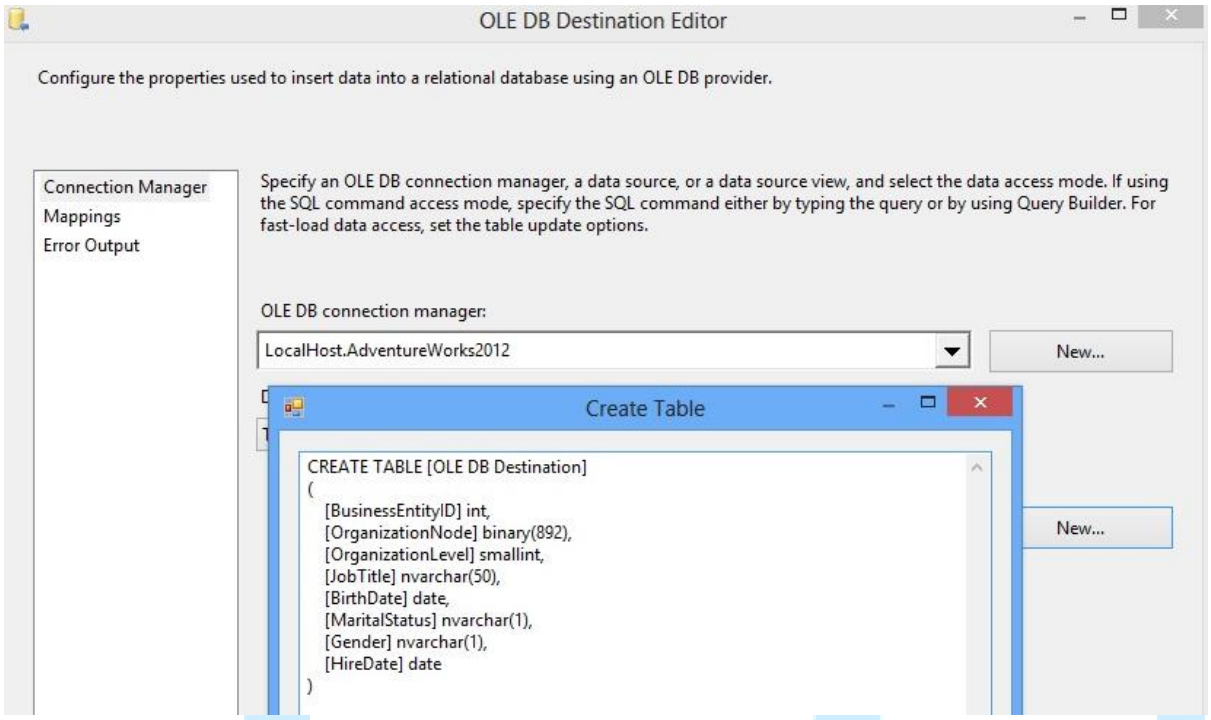
6. Click on New button.
7. Select Cache Connection Manager name or use file cache option.
Cache connection can be made by file or in memory.

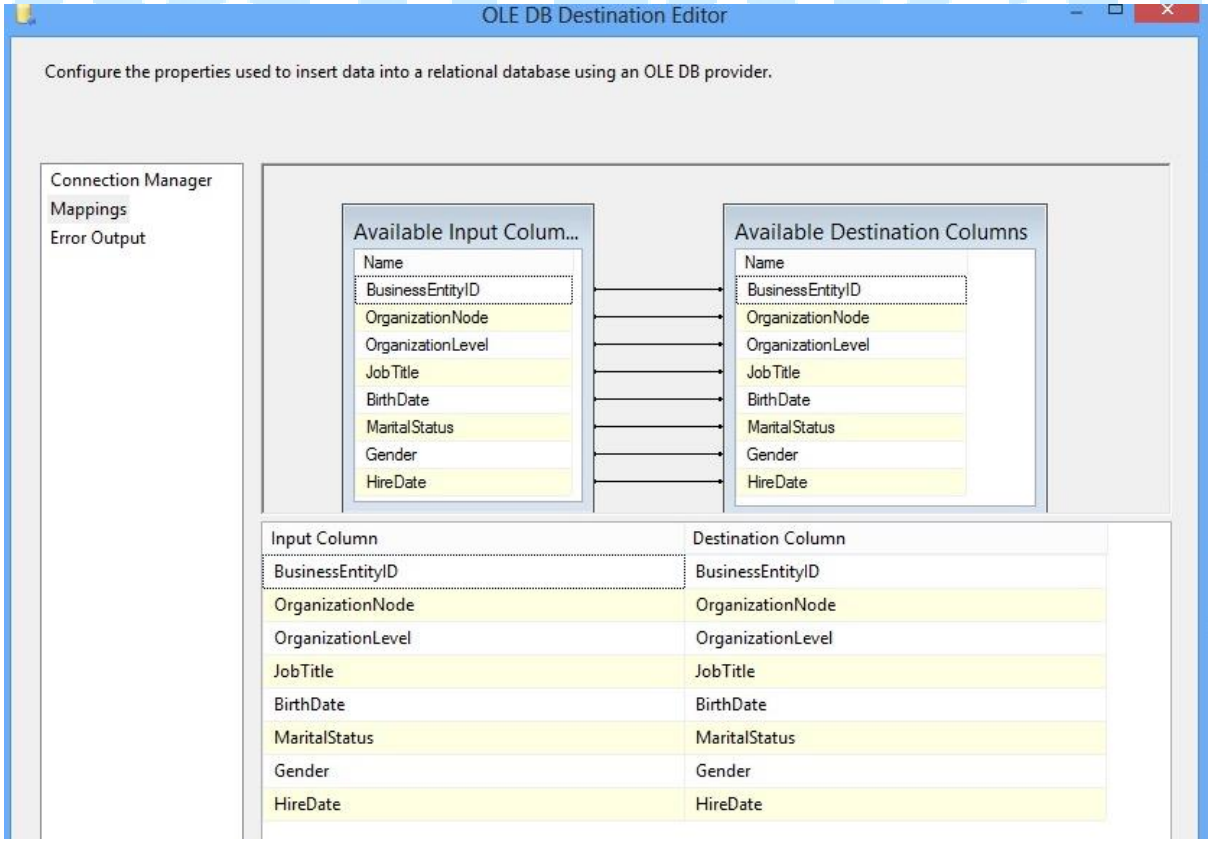



Cache Transformation Editor
Configure the properties used to write data to a cache.
Connection Manager
Mappings
For the Cache connection manager, select an existing connection manager from the list, or create a new connection by clicking New.
Cache connection manager:
Cache Connection Manager
New...
Cache Connection Manager Editor
General
Columns
Connection manager name:
Cache Connection Manager
Description:
Use file cache
File name:
Browse...
Refresh Metadata
The protection level of the package does not apply to the cache file. If the cache file contains sensitive information, use an access control list (ACL) to restrict access to the location or folder in which you store the file.

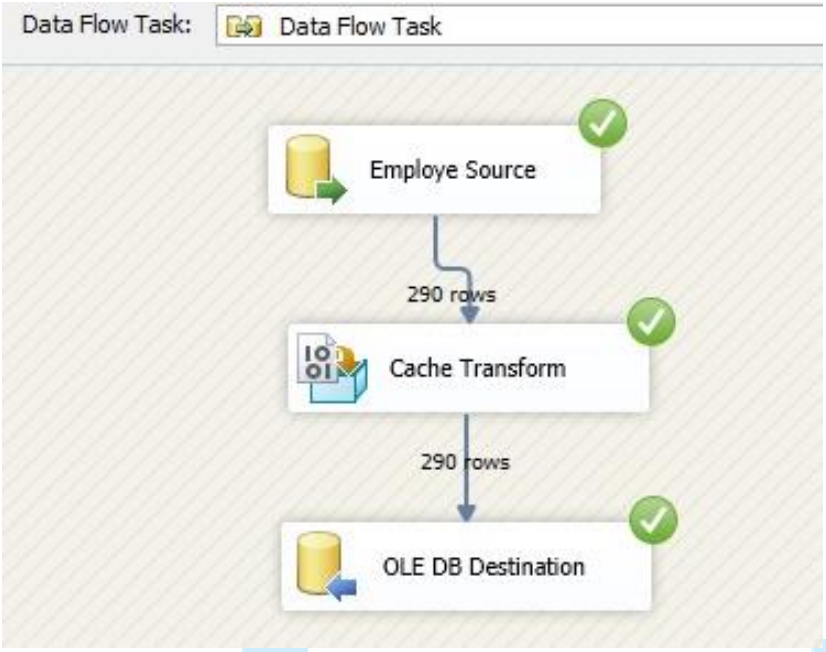

OLE DB Source Editor
Configure the properties used by a data flow to obtain data from any OLE DB provider.
Connection Manager
Columns
Error Output
Available External Colu...
Name
BusinessEntityID
NationalIDNumber
LoginID
OrganizationNode
OrganizationLevel
JobTitle
BirthDate
MaritalStatus
Gender
HireDate
External Column
Output Column
BusinessEntityID
OrganizationNode
OrganizationLevel
JobTitle
BirthDate
MaritalStatus
Gender
HireDate

8. Drag and drop OLEDB Destination and double click on it to configure to new table of database.

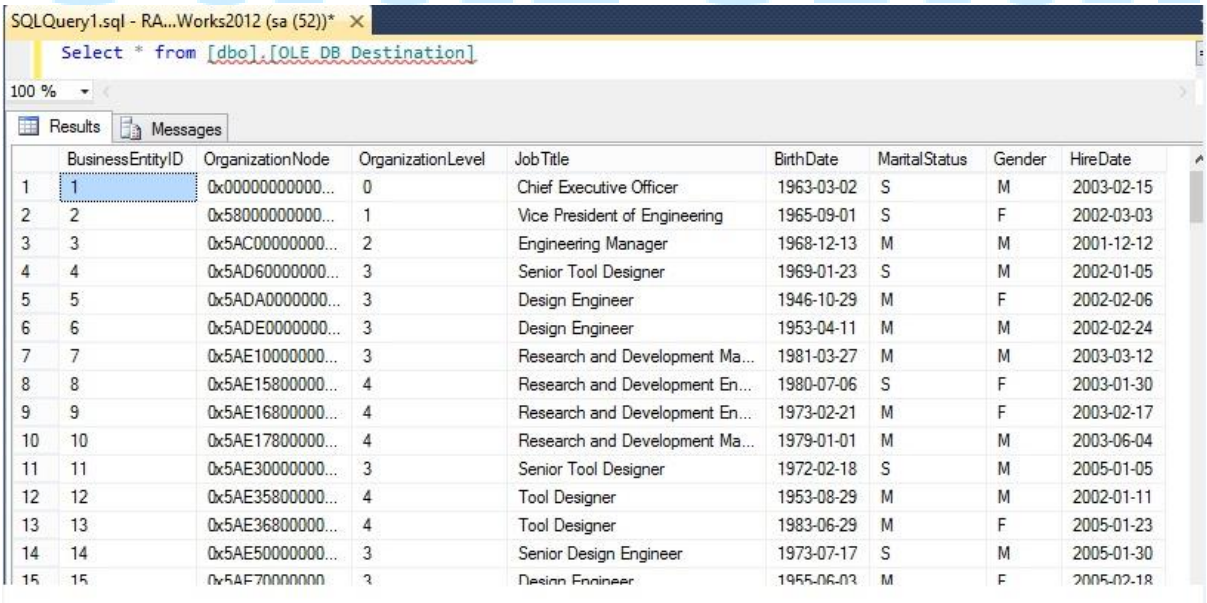




9. Execute the package.



10. Now run the query to check the data in the table.



```
Select * from [dbo].[OLE_DB_Destination]
```

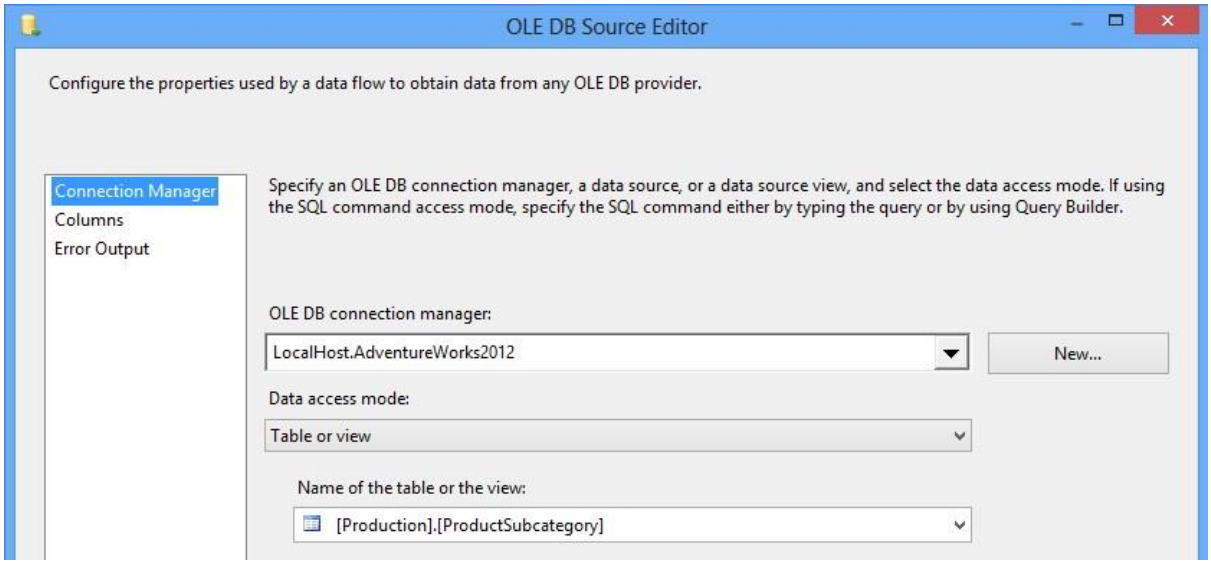
100 %

Results | Messages

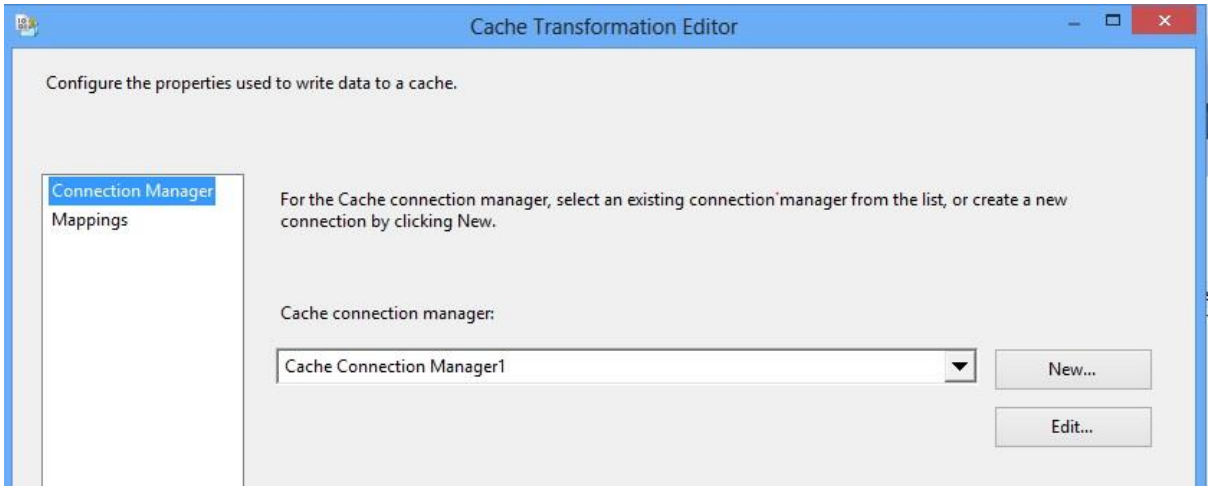
	BusinessEntityID	OrganizationNode	OrganizationLevel	JobTitle	BirthDate	MaritalStatus	Gender	HireDate
1	1	0x00000000000...	0	Chief Executive Officer	1963-03-02	S	M	2003-02-15
2	2	0x58000000000...	1	Vice President of Engineering	1965-09-01	S	F	2002-03-03
3	3	0x5AC00000000...	2	Engineering Manager	1968-12-13	M	M	2001-12-12
4	4	0x5AD60000000...	3	Senior Tool Designer	1969-01-23	S	M	2002-01-05
5	5	0x5ADA0000000...	3	Design Engineer	1946-10-29	M	F	2002-02-06
6	6	0x5ADE0000000...	3	Design Engineer	1953-04-11	M	M	2002-02-24
7	7	0x5AE10000000...	3	Research and Development Ma...	1981-03-27	M	M	2003-03-12
8	8	0x5AE15800000...	4	Research and Development En...	1980-07-06	S	F	2003-01-30
9	9	0x5AE16800000...	4	Research and Development En...	1973-02-21	M	F	2003-02-17
10	10	0x5AE17800000...	4	Research and Development Ma...	1979-01-01	M	M	2003-06-04
11	11	0x5AE30000000...	3	Senior Tool Designer	1972-02-18	S	M	2005-01-05
12	12	0x5AE35800000...	4	Tool Designer	1953-08-29	M	M	2002-01-11
13	13	0x5AE36800000...	4	Tool Designer	1983-06-29	M	F	2005-01-23
14	14	0x5AE50000000...	3	Senior Design Engineer	1973-07-17	S	M	2005-01-30
15	15	0x5AE70000000...	3	Design Engineer	1955-06-03	M	F	2005-02-18

Solution 2

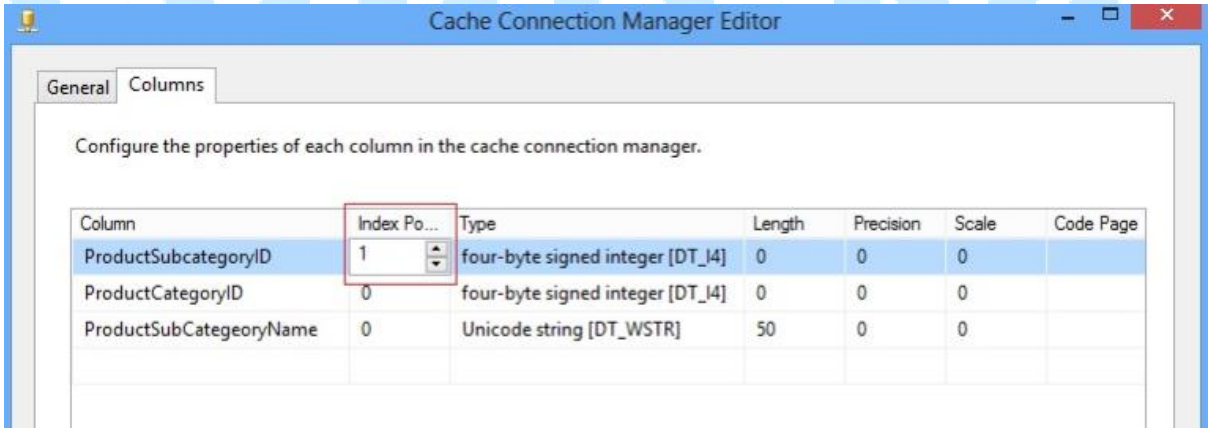
1. Drag and Drop Data Flow Task onto Control Flow and double click on it
2. In Data Flow Tab, drag and drop OLEDB Source and map it to [Production].[ProductSubcategory] table of AdventureWorks2012 database.
3. Go to the Columns section, uncheck the columns which are not required and click OK

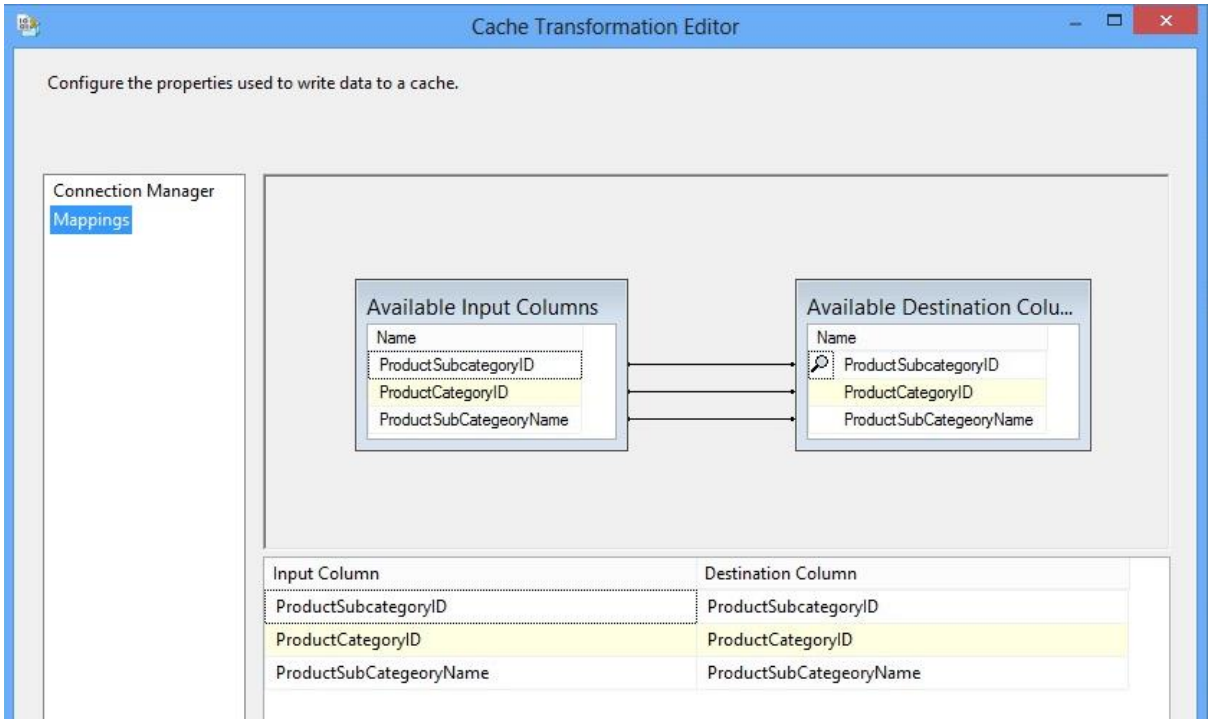


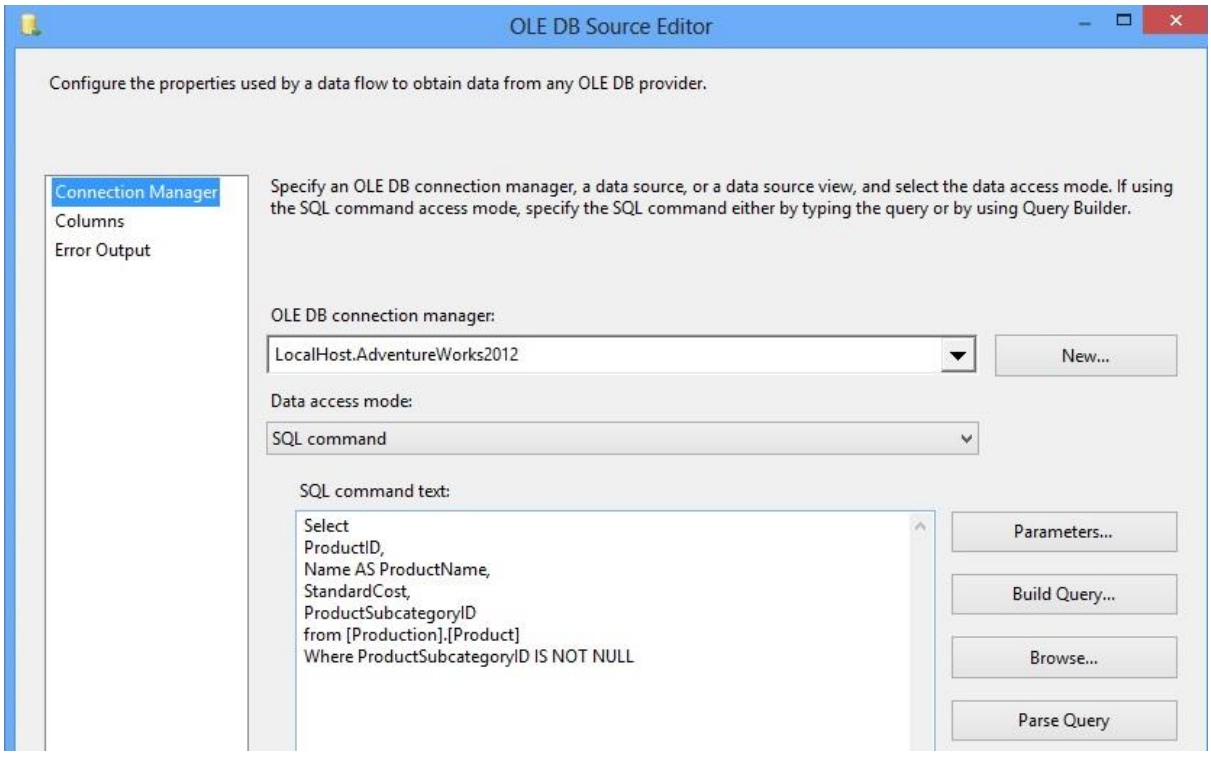
4. Drag and Cache Transform onto DataFlow Task and join to OLEDB Source.
5. Double click on Cache Transform to configure it.
6. Create New Cache Connection Manager.
7. Go to Columns tab, Set the Index Position=1 for Key column.

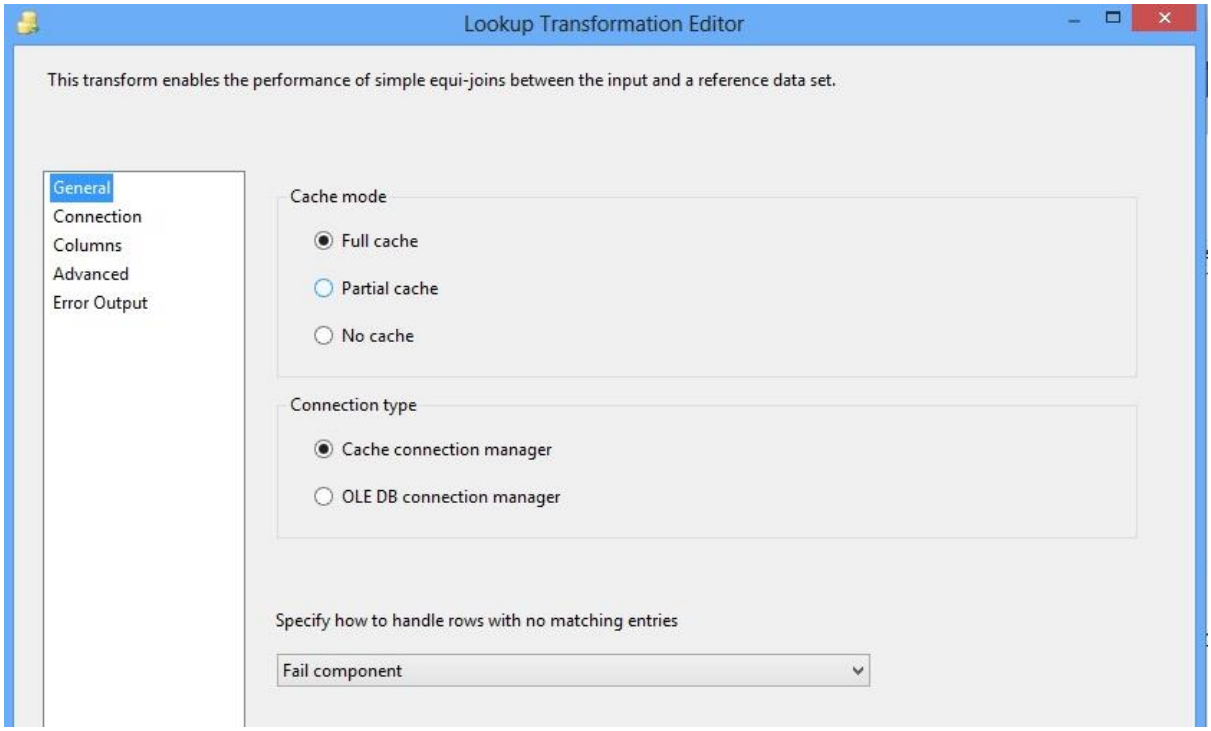


8. Drag and Drop Data Flow Task onto Control Flow, join it to first Data flow task.
9. Double click on 2nd Data Flow Task.
10. In Data Flow Tab, drag and drop OLEDB Source and map it to [Production].[Product] table of AdventureWorks2012 database
11. Drag and drop Lookup Transformation and join with OLEDB Source.
12. Double click the Lookup Transformation to configure it.
13. Select: Cache Model, Connection Type

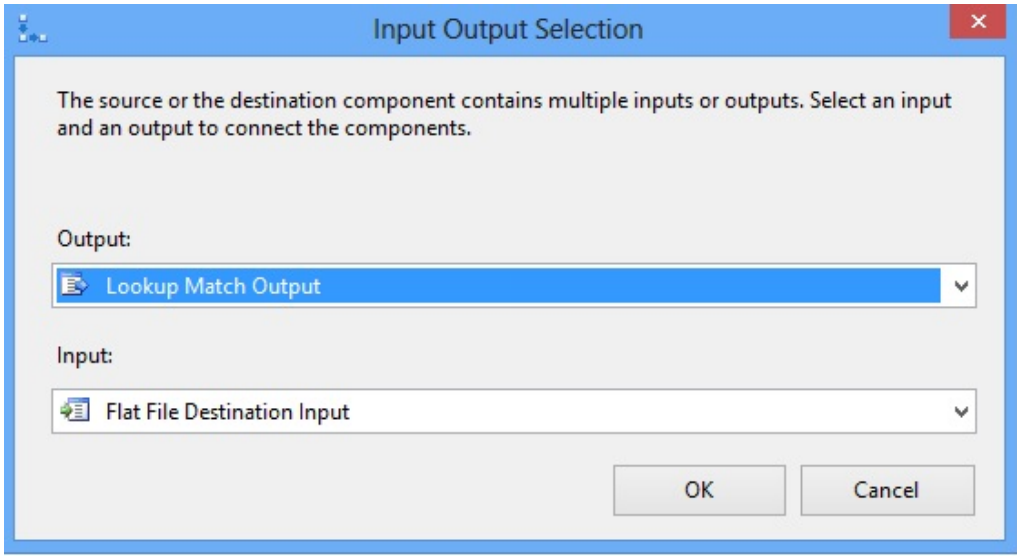



Cache Transformation Editor
Configure the properties used to write data to a cache.
Connection Manager
Mappings
Available Input Columns
Name
ProductSubcategoryID
ProductCategoryID
ProductSubCategeoryName
Available Destination Colu...
Name
ProductSubcategoryID
ProductCategoryID
ProductSubCategeoryName
Input Column
Destination Column
ProductSubcategoryID
ProductSubcategoryID
ProductCategoryID
ProductCategoryID
ProductSubCategeoryName
ProductSubCategeoryName

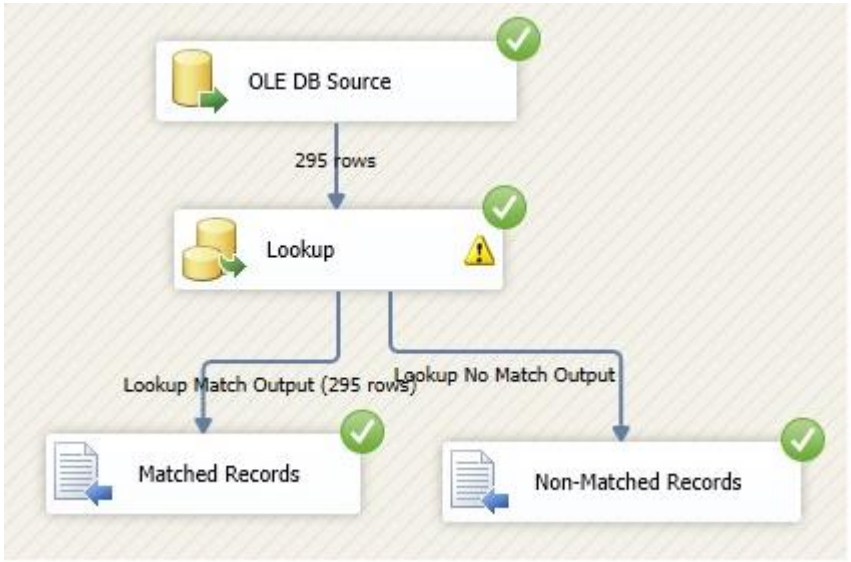

OLE DB Source Editor
Configure the properties used by a data flow to obtain data from any OLE DB provider.
Connection Manager
Columns
Error Output
Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.
OLE DB connection manager:
LocalHost.AdventureWorks2012
New...
Data access mode:
SQL command
SQL command text:
Select
ProductID,
Name AS ProductName,
StandardCost,
ProductSubcategoryID
from [Production].[Product]
Where ProductSubcategoryID IS NOT NULL
Parameters...
Build Query...
Browse...
Parse Query



14. Drag and drop 2 Flat file Destinations onto Data Flow Tab and configure.
15. Rename first Flat File Destination to Matched Records and 2nd Flat file destination to Non-Matched Records.
16. Connect the Lookup Transformation with first Flat File Destination and Select Output.
17. Connect the Lookup Transformation with second Flat File Destination



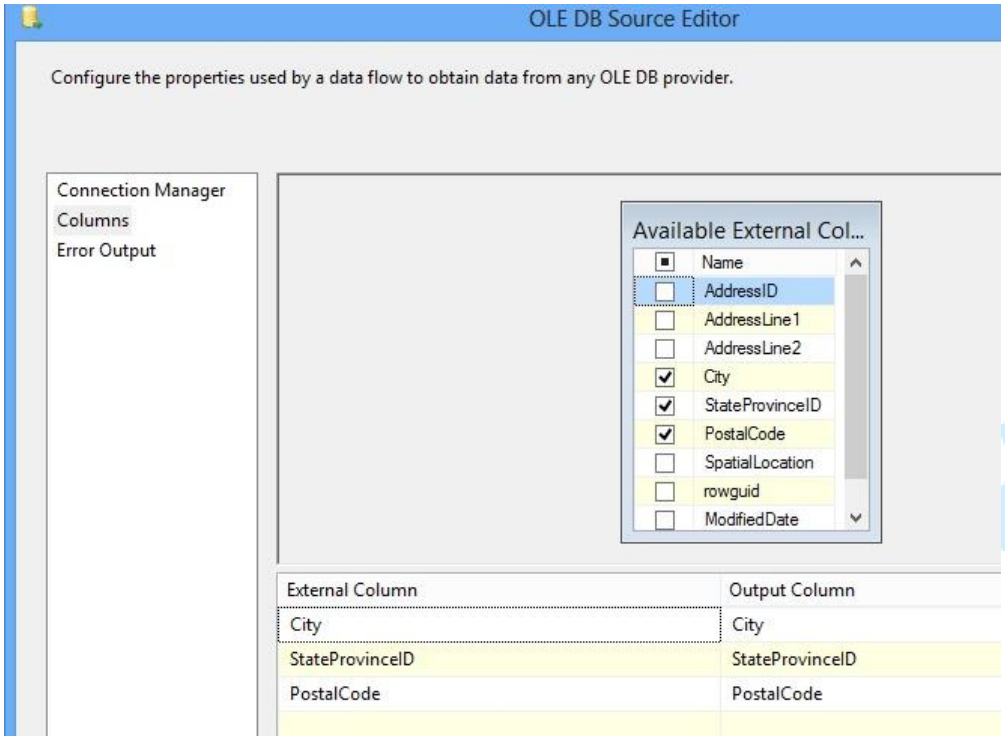
18. Execute the package.

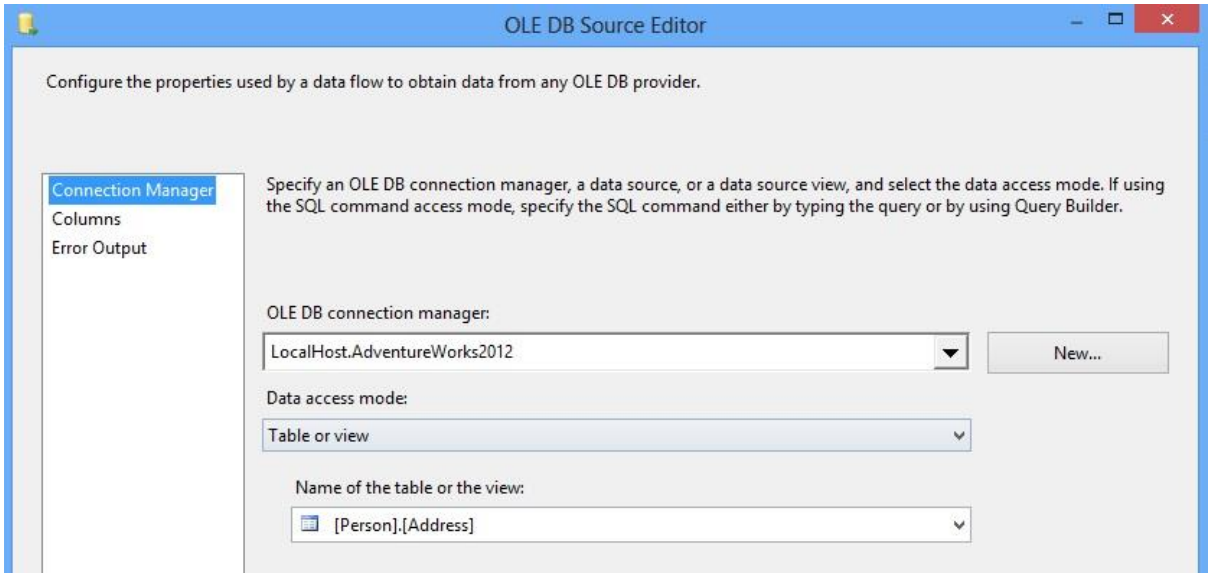


Solution 3

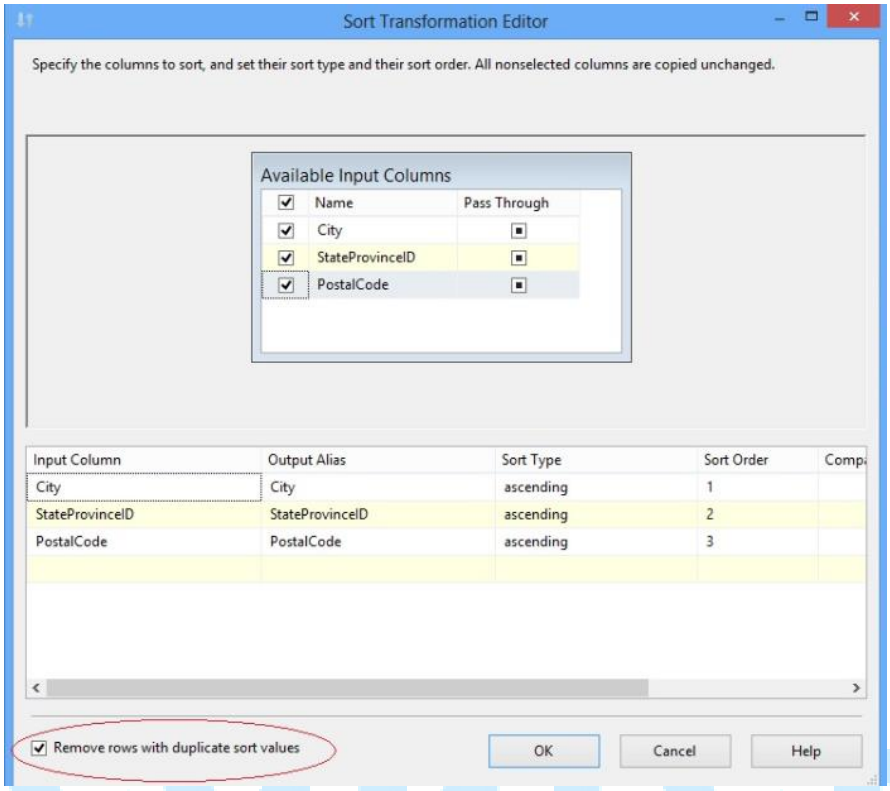
Remove the duplicates using SORT Transformation:

1. Drag and Drop Data Flow Task onto Control Flow and double click on it.
2. In Data Flow Tab, drag and drop OLEDB Source and map it to [Person].[Person] table of AdventureWorks2012 database.
3. Go to the Columns section, uncheck the columns which are not required and click OK.

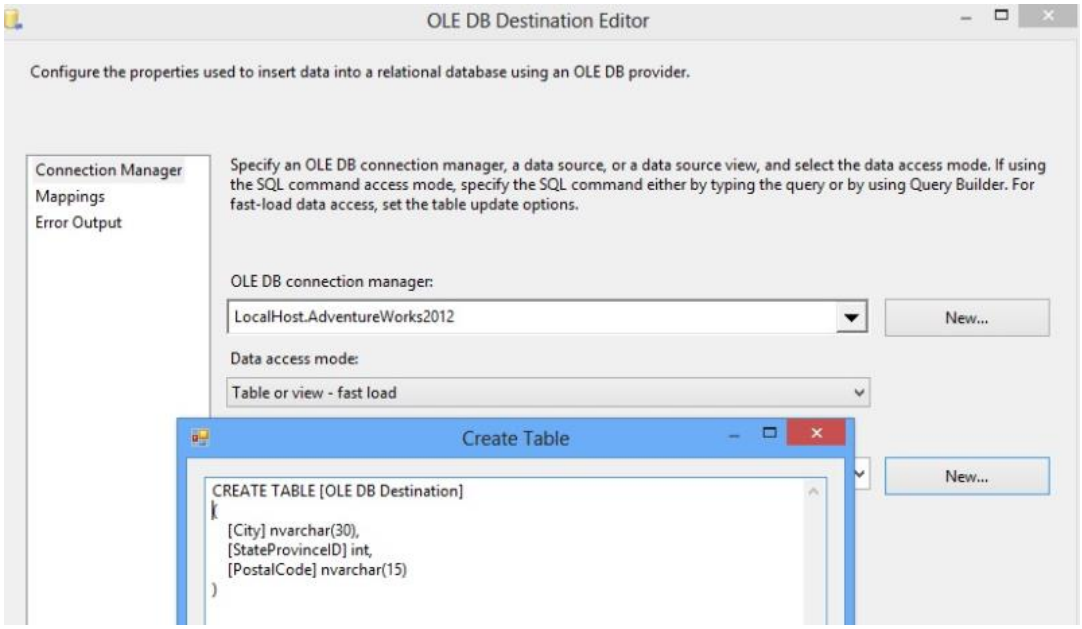


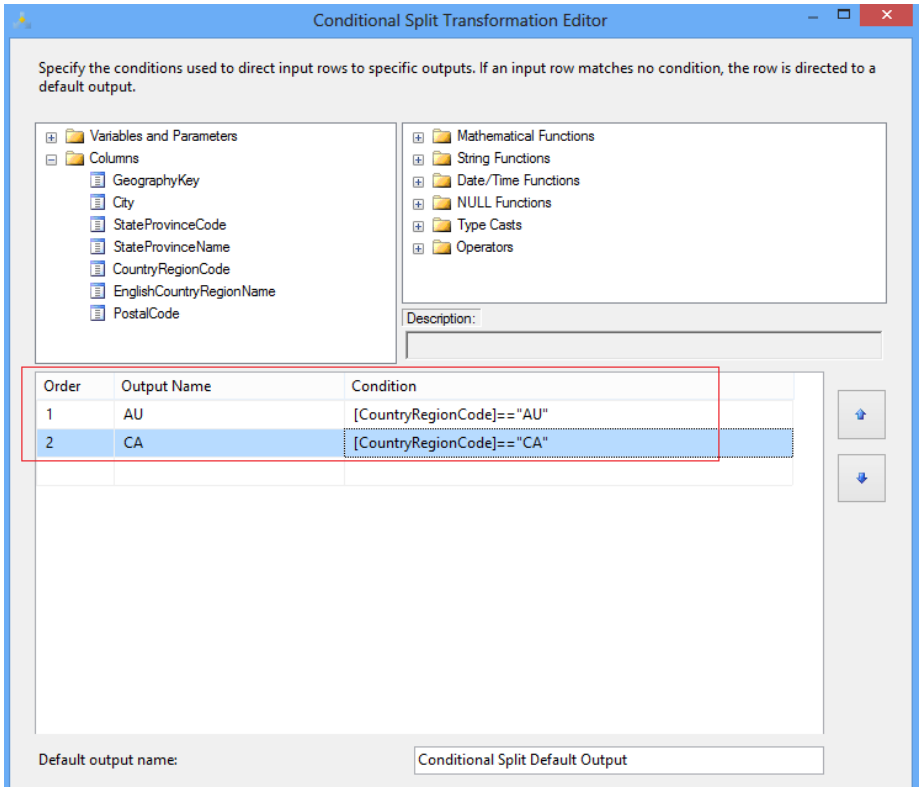


4. Drag and drop SORT Transformation and double click on it selected columns for sorting.
5. Check the radio button: Remove rows with duplicate sort values



6. Drag and drop OLEDB Destination and double click on it configure it a table in Database.
7. Go to Mapping section and click OK.
8. Execute the package now.

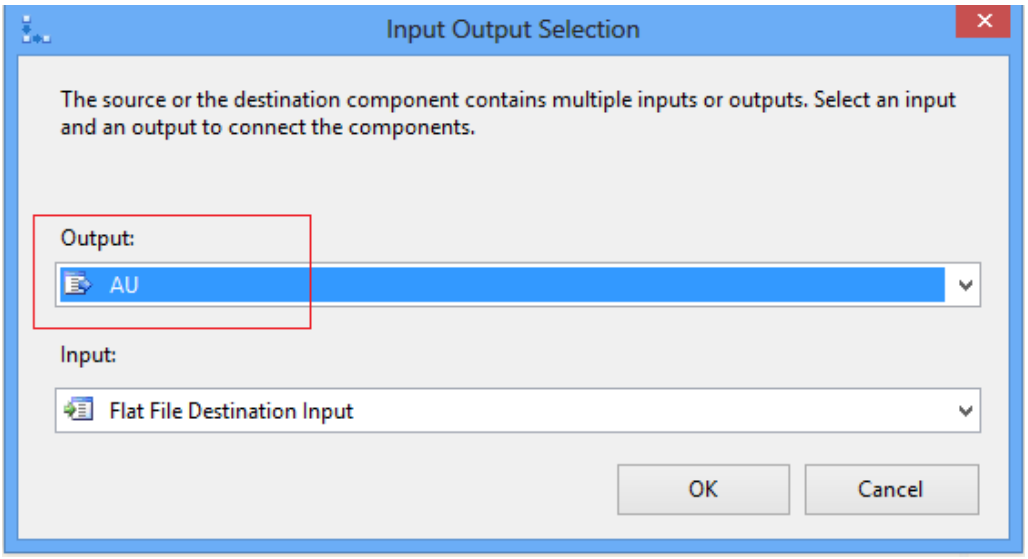




Specify the 2 conditions in the Conditional Split Transformation.

Drag and drop Flat File Destination, Excel Destination and OLEDB Destination.

Connect the outputs of Conditional Split Transformation to each destination and Output Condition.



9. The package removes the duplicate rows.