

TP2 - Business Intelligence

SSIS - DataFlow

Master : Traitement intelligent des systèmes

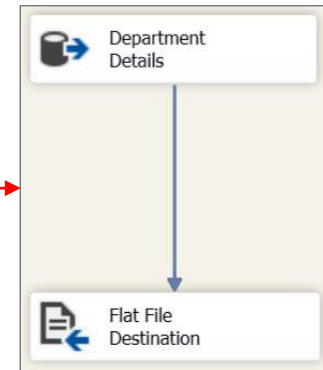
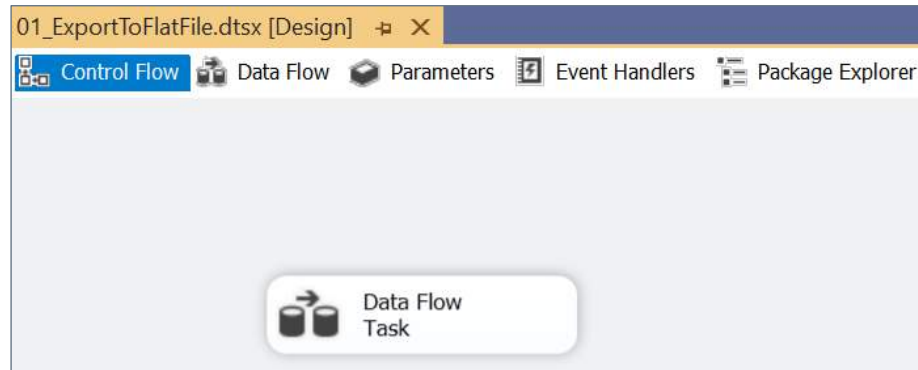
Préparé par: Mme. ELANSARI Khawla

Année Universitaire: 2021/ 2022

Flux de données

| 1. Exporter des données sans transformations

Objectif: Charger les données depuis la table *HumanResources.Departments* dans un fichier Plat. Aucune transformation des données n'est requise



1.1 Configurer la source de données (OLEDB)

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.AdventureWorks2019 New...

Data access mode:
Table or view

Name of the table or the view:
[HumanResources].[Department]

Preview...

OK Cancel Help

Connection Manager

Provider: Native OLE DB\SQL Server Native Client 11.0

Connection
All

Server name:
Refresh

Log on to the server

Authentication: Windows Authentication

User name:

Password:

☐ Save my password

Connect to a database

☒ Select or enter a database name:
AdventureWorks2019

☐ Attach a database file:
 Browse...

Logical name:

Test Connection OK Cancel Help

1.2 Configurer la destination (Fichier Plat)

Flat File Destination Editor

Configure the properties used to connect to and insert data into a text file.

Connection Manager
Mappings

Flat File connection manager:
Flat File Connection Manager

☒ Overwrite data in the file

Header:

Preview...

OK Cancel Help

Flat File Connection Manager Editor

Connection manager name: Flat File Connection Manager

Description:

General
Columns
Advanced
Preview

Select a file and specify the file properties and the file format.

File name: C:\Users\KhawlaElansari\Desktop\Enseigner Browse...

Locale: English (United States) ☐ Unicode

Code page: 1252 (ANSI - Latin I)

Format: Delimited

Text qualifier: <none>

Header row delimiter: {CR}{LF}

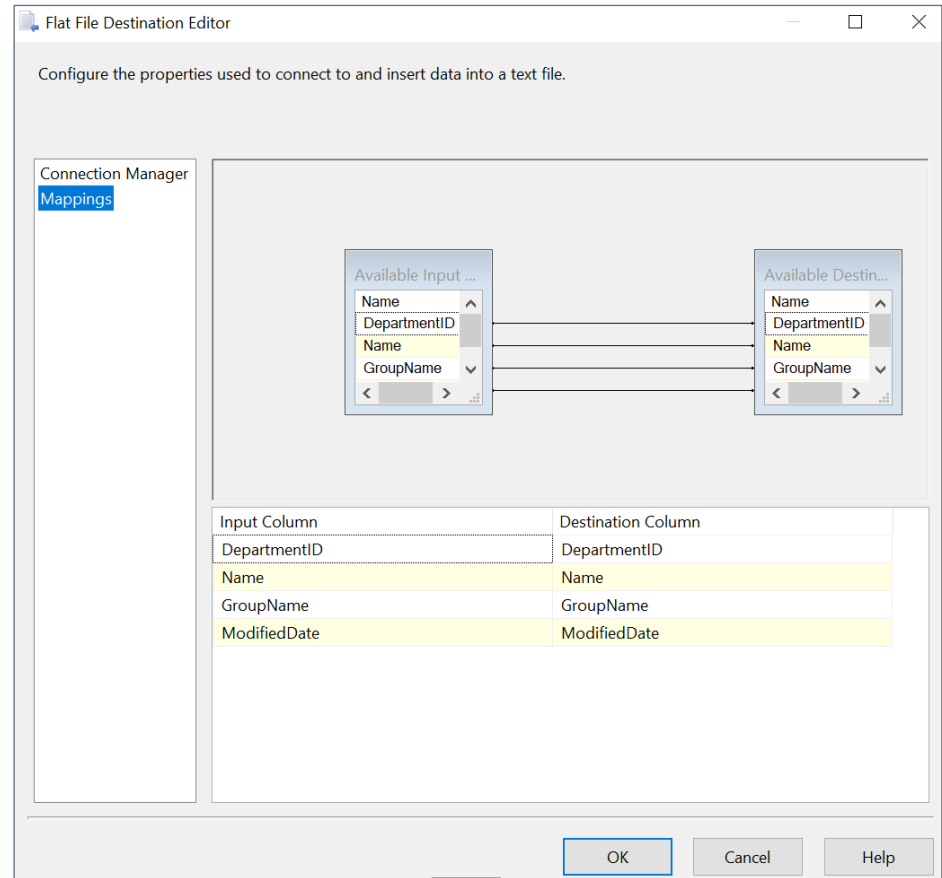
Header rows to skip: 0

☒ Column names in the first data row

OK Cancel Help

1.2 Configurer la destination (Fichier Plat) :

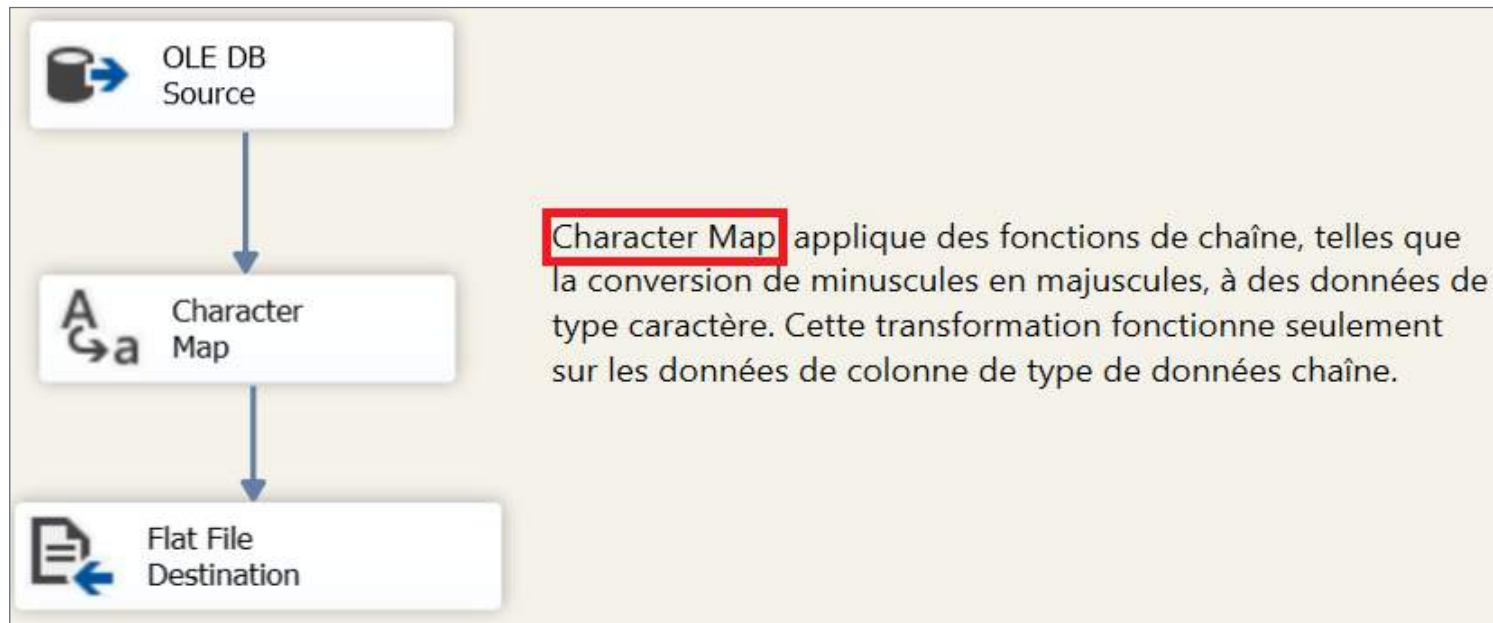
Définir les mappings entre les colonnes sources et les colonnes destination



| 2. Transformation SSIS: Character Map

Objectifs:

- Charger les données de la table **HumanResources.Departments** dans un fichier plat : **CharacterMapOutput.txt**
- Ajouter une nouvelle colonne **DepartmentName** contenant les noms des départements convertis en Majuscule
- Convertir les valeurs du **GroupName** en Minuscule



2. Transformation SSIS: Character Map

Character Map Transformation Editor

Configure the properties used to apply string operations to columns with character data. Place the operation result in a new column or update an existing column.

Available Input Column...

- ☐ Name
- ☐ DepartmentID
- ☒ Name
- ☒ GroupName
- ☐ ModifiedDate

Input Column	Destination	Operation	Output Alias
Name	New column	Uppercase	DepartmentName
GroupName	In-place change	Lowercase	GroupName

Configure Error Output... OK Cancel Help

CharacterMapOutput.txt - Notepad

File Edit Format View Help

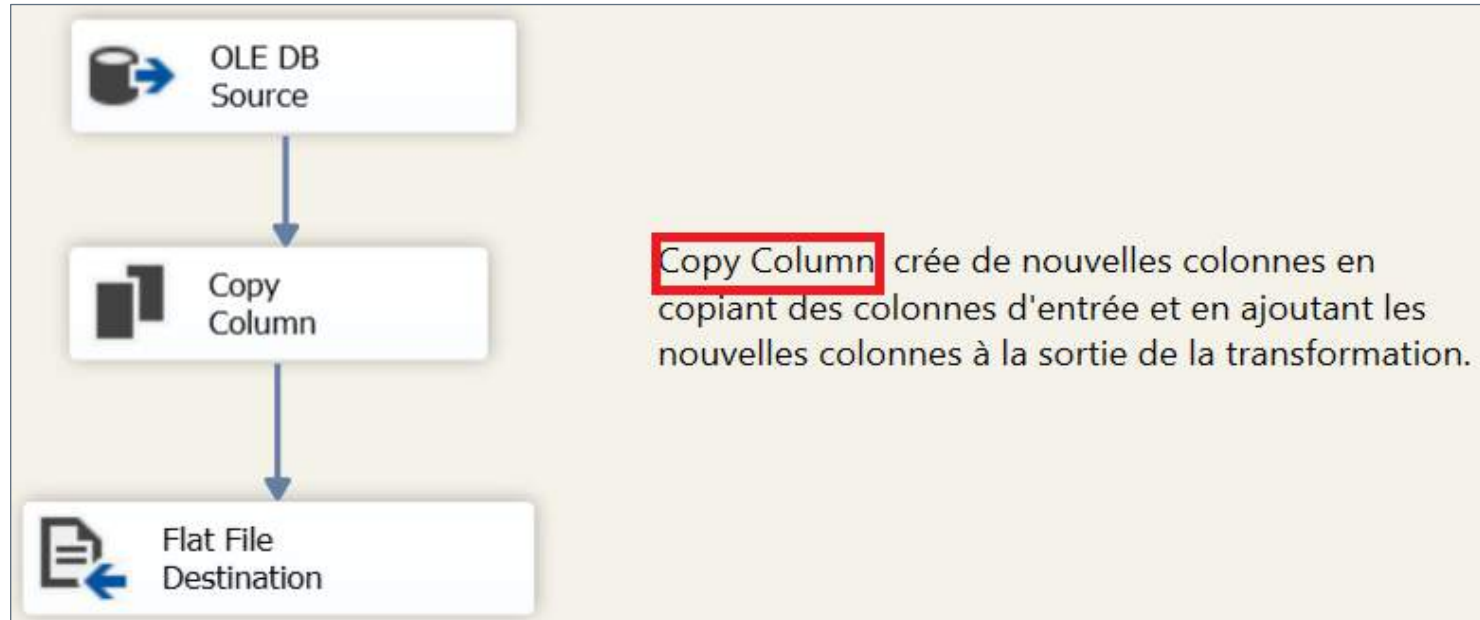
DepartmentID,Name,GroupName,ModifiedDate,DepartmentName

1,Engineering,research and development,2008-04-30 00:00:00,ENGINEERING
2,Tool Design,research and development,2008-04-30 00:00:00,TOOL DESIGN
3,Sales,sales and marketing,2008-04-30 00:00:00,SALES
4,Marketing,sales and marketing,2008-04-30 00:00:00,MARKETING
5,Purchasing,inventory management,2008-04-30 00:00:00,PURCHASING
6,Research and Development,research and development,2008-04-30 00:00:00,RESEARCH AND DEVELOPMENT
7,Production,manufacturing,2008-04-30 00:00:00,PRODUCTION
8,Production Control,manufacturing,2008-04-30 00:00:00,PRODUCTION CONTROL
9,Human Resources,executive general and administration,2008-04-30 00:00:00,HUMAN RESOURCES
10,Finance,executive general and administration,2008-04-30 00:00:00,FINANCE
11,Information Services,executive general and administration,2008-04-30 00:00:00,INFORMATION SERVICES
12,Document Control,quality assurance,2008-04-30 00:00:00,DOCUMENT CONTROL
13,Quality Assurance,quality assurance,2008-04-30 00:00:00,QUALITY ASSURANCE
14,Facilities and Maintenance,executive general and administration,2008-04-30 00:00:00,FACILITIES AND MAINTENANCE
15,Shipping and Receiving,inventory management,2008-04-30 00:00:00,SHIPPING AND RECEIVING
16,Executive,executive general and administration,2008-04-30 00:00:00,EXECUTIVE

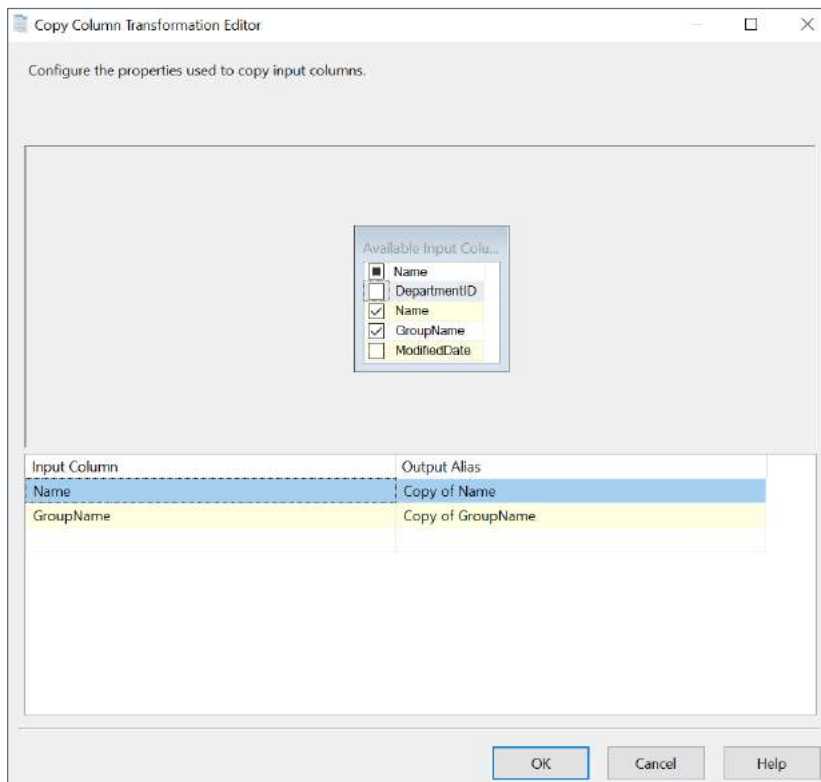
3. Transformation SSIS: Copy Column

Objectifs:

- Charger les données de la table **HumanResources.Departments** dans un fichier plat : **CopyColumnOutput.txt**
- Ajouter 2 nouvelles colonnes copies des colonnes : Name , Groupname. Nommer les nouvelles colonnes: **Copy of Name** , **Copy of Groupname**



3. Transformation SSIS: Copy Column



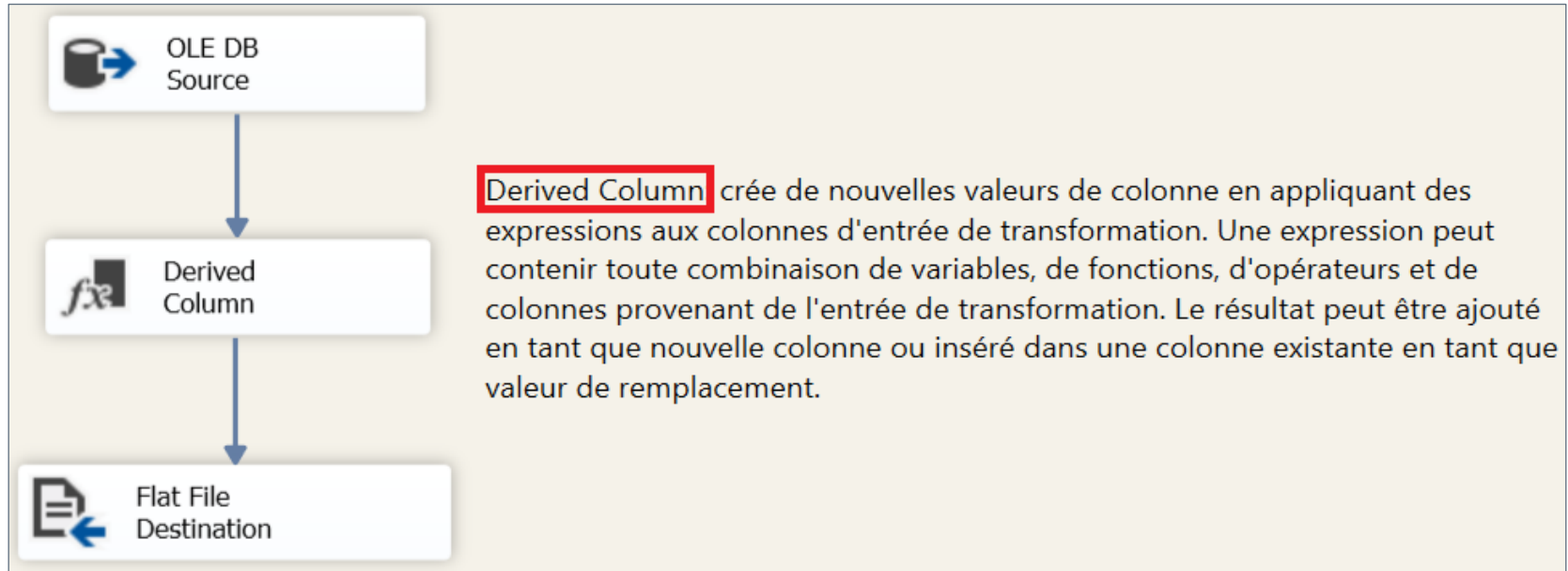
The screenshot shows a Notepad window titled 'CopyColumnOutput.txt - Notepad'. The text inside is a list of 16 rows, each containing a number, a department name, a date, and a description. The first row is highlighted in yellow.

```
DepartmentID,Name,GroupName,ModifiedDate,Copy of Name,Copy of GroupName
1,Engineering,Research and Development,2008-04-30 00:00:00,Engineering,Research and Development
2,Tool Design,Research and Development,2008-04-30 00:00:00,Tool Design,Research and Development
3,Sales,Sales and Marketing,2008-04-30 00:00:00,Sales,Sales and Marketing
4,Marketing,Sales and Marketing,2008-04-30 00:00:00,Marketing,Sales and Marketing
5,Purchasing,Inventory Management,2008-04-30 00:00:00,Purchasing,Inventory Management
6,Research and Development,Research and Development,2008-04-30 00:00:00,Research and Development,Research and Development
7,Production,Manufacturing,2008-04-30 00:00:00,Production,Manufacturing
8,Production Control,Manufacturing,2008-04-30 00:00:00,Production Control,Manufacturing
9,Human Resources,Executive General and Administration,2008-04-30 00:00:00,Human Resources,Executive General and Administration
10,Finance,Executive General and Administration,2008-04-30 00:00:00,Finance,Executive General and Administration
11,Information Services,Executive General and Administration,2008-04-30 00:00:00,Information Services,Executive General and Administration
12,Document Control,Quality Assurance,2008-04-30 00:00:00,Document Control,Quality Assurance
13,Quality Assurance,Quality Assurance,2008-04-30 00:00:00,Quality Assurance,Quality Assurance
14,Facilities and Maintenance,Executive General and Administration,2008-04-30 00:00:00,Facilities and Maintenance,Executive General and Administration
15,Shipping and Receiving,Inventory Management,2008-04-30 00:00:00,Shipping and Receiving,Inventory Management
16,Executive,Executive General and Administration,2008-04-30 00:00:00,Executive,Executive General and Administration
```

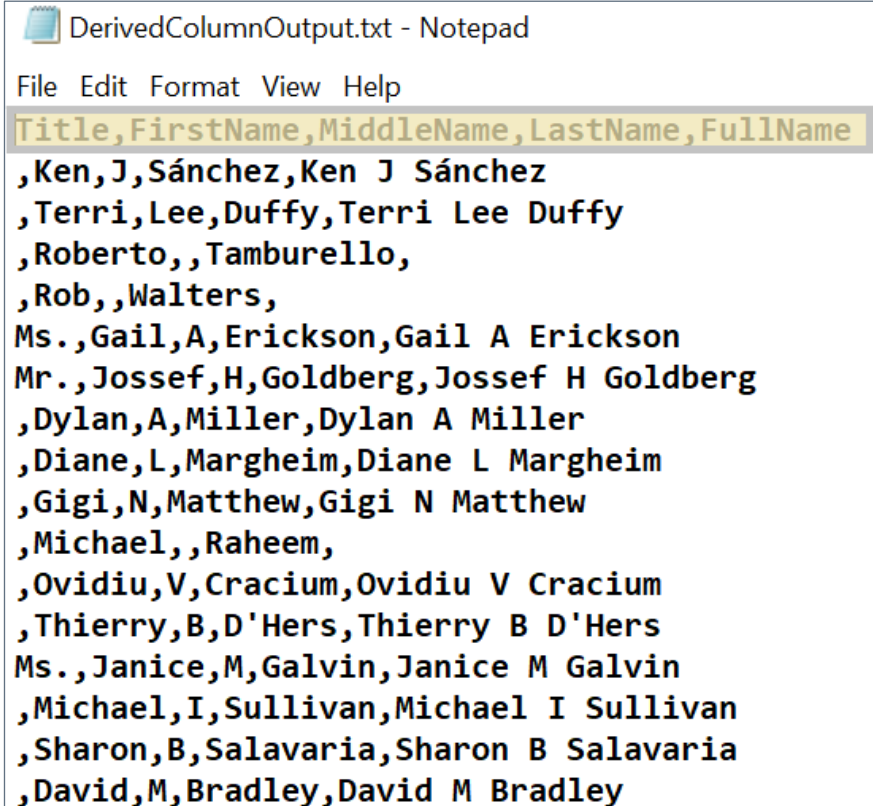
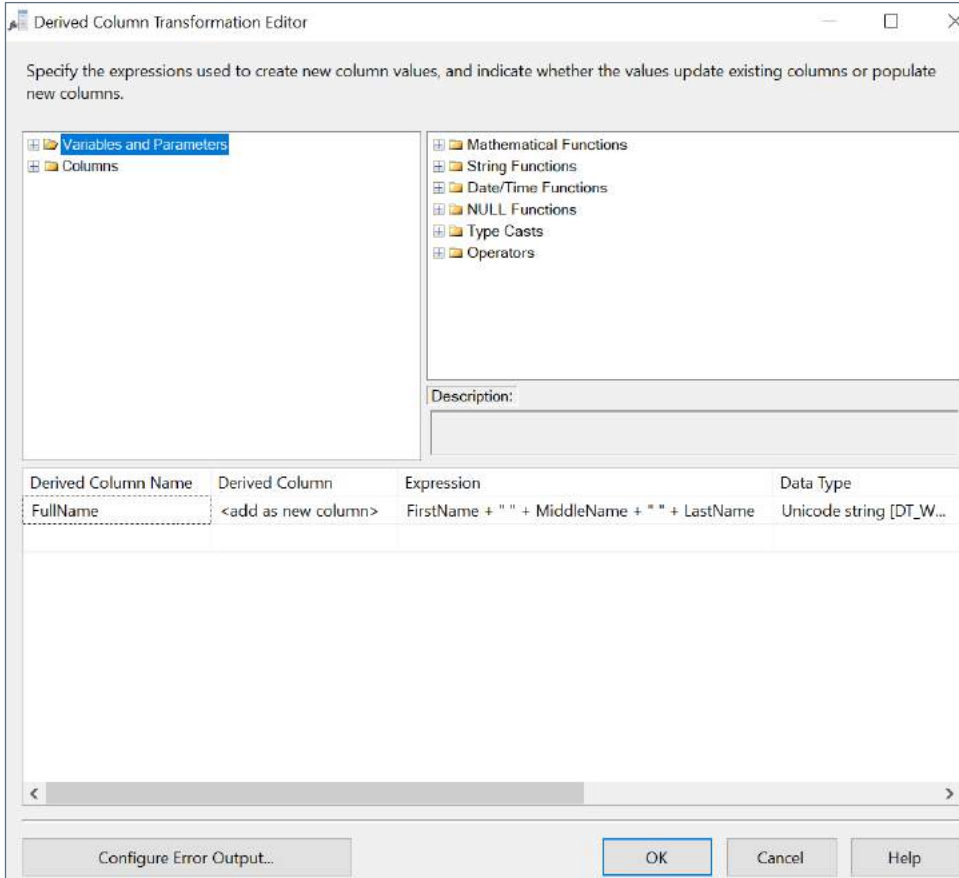
4. Transformation SSIS: Derived Column

Objectifs:

- Charger les données de la table **Person.Person** dans un fichier plat : **DerivedColumnOutput.txt**
- Ajouter une nouvelle colonne **FullName** contenant la concaténation des 3 valeurs: **FirstName**, **MiddleName** et **LastName**



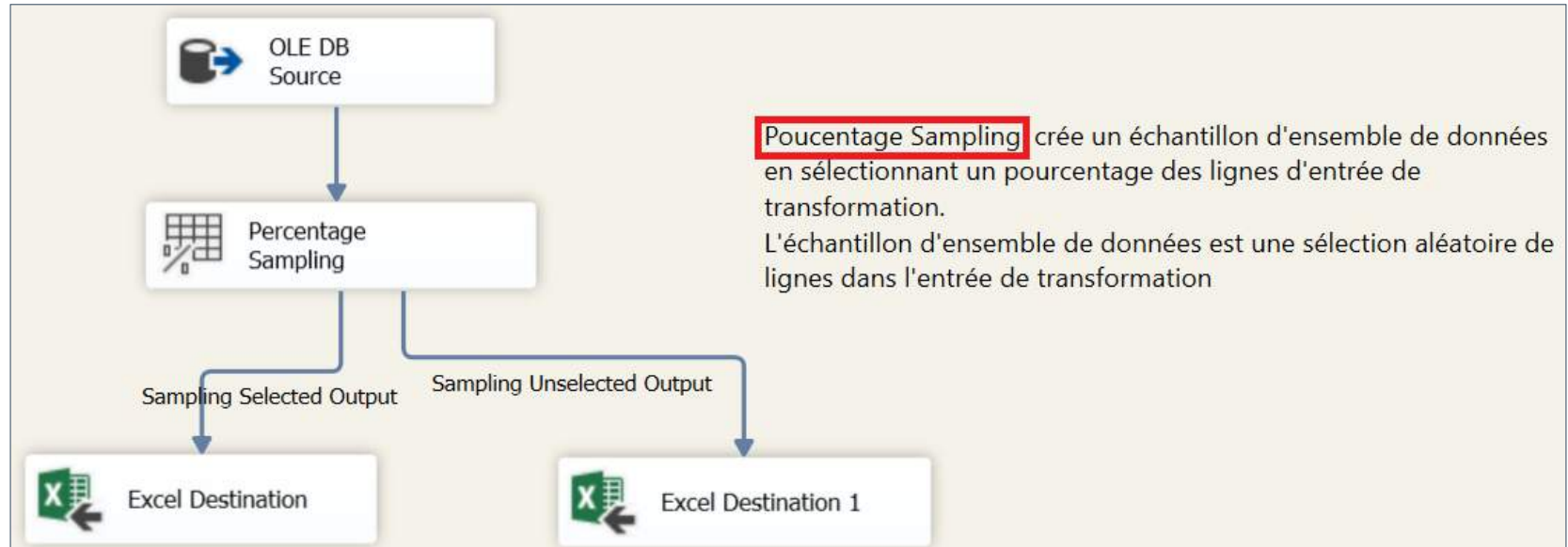
4. Transformation SSIS: Derived Column




5. Transformation SSIS: Percentage Sampling

Objectifs:

- Charger les données de la table **Person.Person** dans un fichier Excel : **PercentageSamplingOutput.xlsx**
- Dans une 1^{ère} feuille du fichier nommée **Selected_Records**, charger 20% des données de la source
- Dans une 2^{ème} feuille du fichier nommée **Unselected_Records**, charger les 80% restantes des données



5. Transformation SSIS: Percentage Sampling

 % Percentage Sampling Transformation Editor

Configure the properties to divide the transformation input into a sample and a leftovers output, using a percentage to specify the sample size.


Percentage of rows:

Sample output name:

Unselected output name:

☐ Use the following random seed:

OK Cancel Help

 Excel Destination Editor

Configure the properties that enable the insertion of data via an Excel provider.

Connection Manager
Mappings
Error Output


Specify a connection manager, data source, or data source view for the Excel destination. Then, select the mode used to access data within the destination. After selecting the data access mode, select from among the additional data access options that appear.

Excel connection manager:
 New...

Data access mode:

Name of the Excel sheet:
 New...

View Existing

 Select a table or view from the list.

OK Cancel Help 14

5. Transformation SSIS: Percentage Sampling

Create Table

```
CREATE TABLE `Selected_Records` (  
  `Title` NVARCHAR(8),  
  `FirstName` NVARCHAR(50),  
  `MiddleName` NVARCHAR(50),  
  `LastName` NVARCHAR(50)  
)
```

OK Cancel

Excel Destination Editor

Configure the properties that enable the insertion of data via an Excel provider.

Connection Manager
Mappings
Error Output

Available Input Columns: Name, Title, **FirstName**, MiddleName

Available Destination Columns: Name, Title, **FirstName**, MiddleName

Input Column	Destination Column
Title	Title
FirstName	FirstName
MiddleName	MiddleName
LastName	LastName

OK Cancel Help

5. Transformation SSIS: Percentage Sampling

Excel Destination Editor

Configure the properties that enable the insertion of data via an Excel provider.

Connection Manager
Mappings
Error Output

Specify a connection manager, data source, or data source view for the Excel destination. Then, select the mode used to access data within the destination. After selecting the data access mode, select from among the additional data access options that appear.

Excel connection manager:
Excel Connection Manager ▼ New...

Data access mode:
Table or view ▼

Name of the Excel sheet:
Selected_Records ▼ New...

View Existing

OK Cancel Help

Excel Destination Editor

Configure the properties that enable the insertion of data via an Excel provider.

Connection Manager
Mappings
Error Output

Specify a connection manager, data source, or data source view for the Excel destination. Then, select the mode used to access data within the destination. After selecting the data access mode, select from among the additional data access options that appear.

Excel connection manager:
Excel Connection Manager ▼ New...

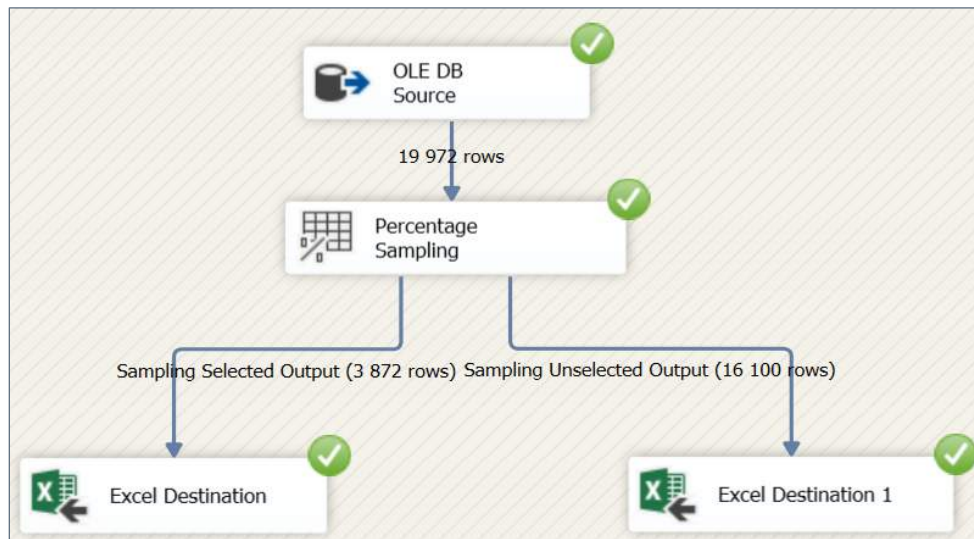
Data access mode:
Table or view ▼

Name of the Excel sheet:
UnSelected_Records ▼ New...

View Existing

OK Cancel Help

6. Transformation SSIS: Percentage Sampling



AutoSave Off PercentageSamplingOutput.xlsx

File Home Insert Draw Page Layout Formulas Data Review

Paste Clipboard Font Alignment Number Styles

Calibri 11 B I U A A % Conditional Formatting Format as Table Cell Styles

A1 Title

	A	B	C	D	E	F	G	H
1	Title	FirstName	MiddleName	LastName				
2		Ken	J	Sánchez				
3		Mary	A	Dempsey				
4		Margie	W	Shoop				
5		Kim	B	Abercrombie				
6		Nancy	A	Anderson				
7		Eugene	O	Kogan				
8		Doris	M	Hartwig				
9		Diane	R	Glimp				
10		Bob	N	Hohman				
11		Jay	G	Adams				
12		Stacy	T	Smith				

UnSelected_Records Selected_Records

Ready

6. Transformation SSIS: Row Sampling

Objectifs:

- Charger 20 enregistrements de la table **Person.Person** dans un fichier plat : **RowSamplingOutput.txt**



6. Transformation SSIS: Row Sampling

Row Sampling Transformation Editor

Configure the properties used to divide the transformation input into a sample and a leftovers output, using a number to specify the sample size.

Sampling
Columns

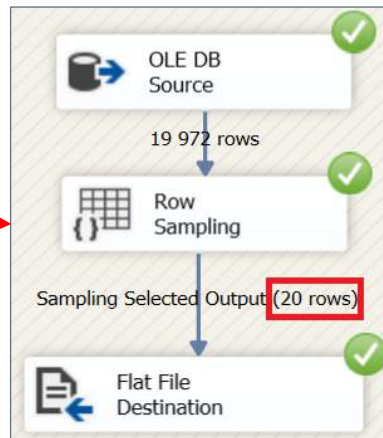
Number of rows: 20

Sample output name: Sampling Selected Output

Unselected output name: Sampling Unselected Output

☐ Use the following random seed: 1

OK Cancel Help



RowSamplingOutput.txt - Notepad

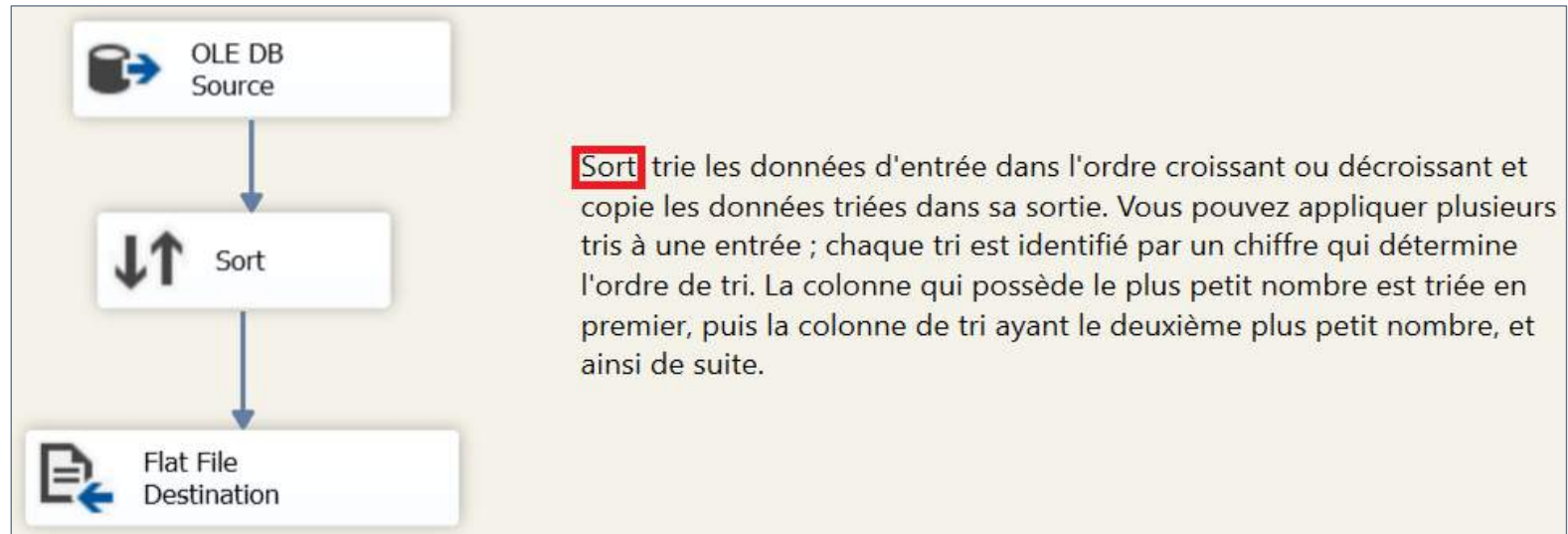
File Edit Format View Help

Title,FirstName,MiddleName,LastName
,Darryl,Y,West
,Joe,,Suri
,Crystal,J,Gao
,Jacob,M,Taylor
,Devin,,Campbell
,Ms.,Esther,K.,Valle
,Mindy,,Jai
,Rachael,,Raman
,Miguel,S,White
,Jason,D,Green
,Clarence,E,Chander
,Hunter,,Jenkins
,Clayton,K,Rai
,Mariah,M,Wood
,Jonathan,,Johnson
,Morgan,,Lewis
,Alex,,Brooks
,Justin,,Flores
,Savannah,E,Gray
,Ms.,Karren,K.,Burkhardt

7. Transformation SSIS: Sort

Objectifs:

- Charger les données de la table **Person.StateProvince** dans un fichier plat : **SortOutput.txt**
- Trier les données selon les valeurs de la colonne *CountryRegionCode* par ordre croissant
- Pour chaque *CountryRegionCode* , trier les données selon les valeurs de la colonne *StateProvinceCode* par ordre décroissant
- Pour chaque *StateProvinceCode* , trier les données selon les valeurs de la colonne *Name* par ordre croissant



7. Transformation SSIS: Sort

Sort Transformation Editor

Specify the columns to sort, and set their sort type and their sort order. All nonselected columns are copied unchanged.

Available Input Columns

<input type="checkbox"/>	Name	Pass T...
<input type="checkbox"/>	StateProvinceID	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	StateProvinceCode	<input type="checkbox"/>
<input checked="" type="checkbox"/>	CountryRegionCode	<input type="checkbox"/>
<input checked="" type="checkbox"/>	Name	<input type="checkbox"/>
<input type="checkbox"/>	TerritoryID	<input checked="" type="checkbox"/>

Input Column	Output Alias	Sort Type	Sort Order
StateProvinceCode	StateProvinceCode	descending	2
CountryRegionCode	CountryRegionCode	ascending	1
Name	Name	ascending	3

☐ Remove rows with duplicate sort values

OK Cancel Help

SortOutput.txt - Notepad

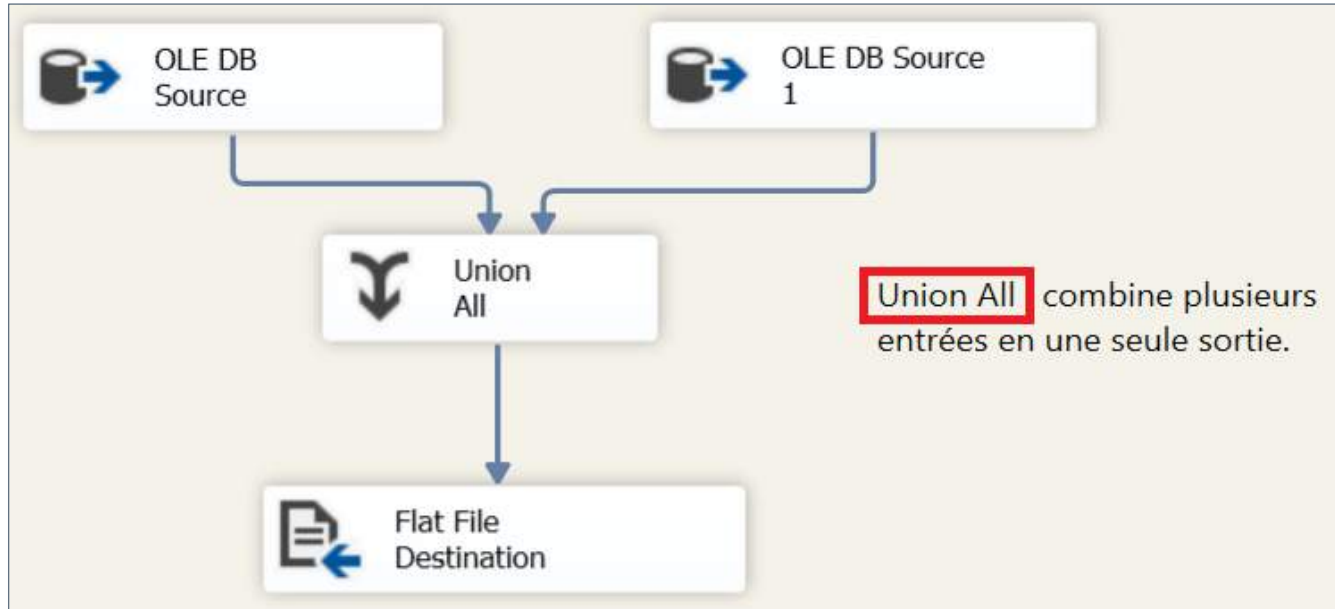
File Edit Format View Help

```
StateProvinceID,StateProvinceCode,CountryRegionCode,Name,TerritoryID
5,AS ,AS,American Samoa,1
77,VIC,AU,Victoria,9
71,TAS,AU,Tasmania,9
66,SA ,AU,South Australia,9
64,QLD,AU,Queensland,9
50,NSW,AU,New South Wales,9
83,YT ,CA,Yukon Territory,6
69,SK ,CA,Saskatchewan,6
63,QC ,CA,Quebec,6
60,PE ,CA,Prince Edward Island,6
```

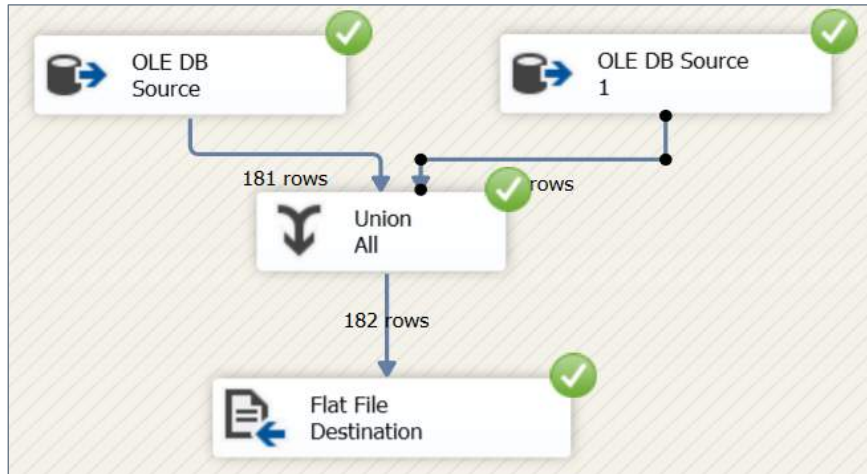
| 8. Transformation SSIS: Union All

Objectifs:

- Charger les données des deux tables **Person.StateProvince**, **Person.NewStateProvinceDetails** dans un fichier plat : **UnionAll.txt**
- N.B:** les deux tables doivent avoir la même structure



8. Transformation SSIS: Union All



UnionAll.txt - Notepad

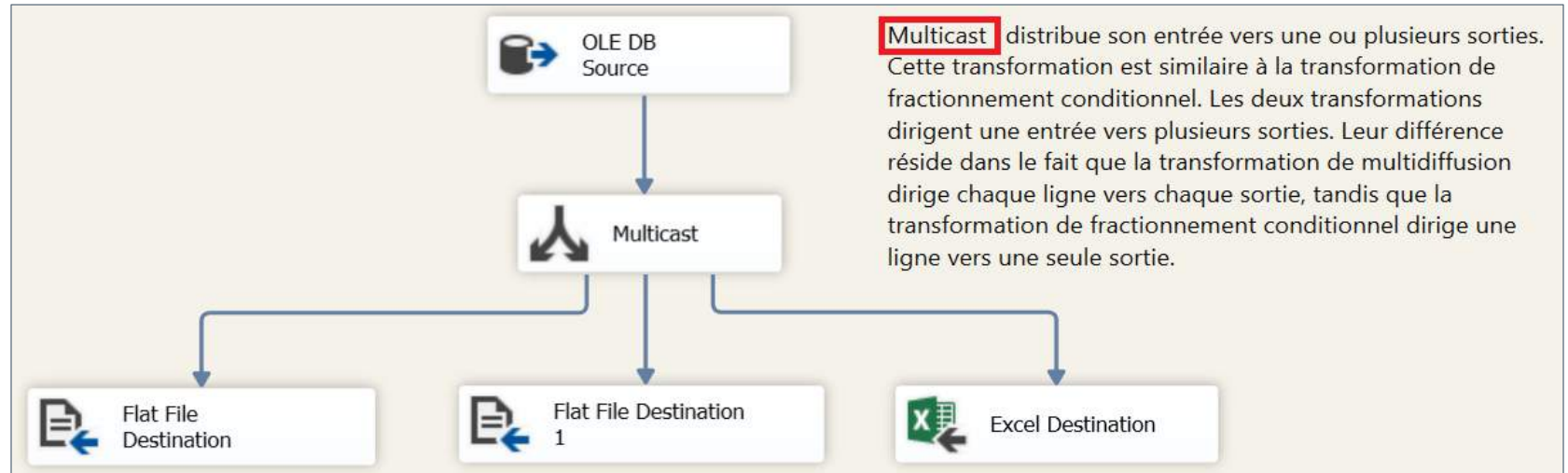
File Edit Format View Help

```
StateProvinceID,StateProvinceCode,CountryRegionCode,Name,TerritoryID
1,AB ,CA,Alberta,6
2,AK ,US,Alaska,1
3,AL ,US,Alabama,5
4,AR ,US,Arkansas,3
5,AS ,AS,American Samoa,1
6,AZ ,US,Arizona,4
7,BC ,CA,British Columbia,6
8,BY ,DE,Bayern,8
9,CA ,US,California,4
10,CO ,US,Colorado,3
11,CT ,US,Connecticut,2
12,DC ,US,District of Columbia,2
13,DE ,US,Delaware,2
14,ENG,GB,England,10
...
179,93 ,FR,Seine Saint Denis,7
180,94 ,FR,Val de Marne,7
181,95 ,FR,Val d'Oise,7
1,87 ,FR,Vienne (Haute),7
```

| 9. Transformation SSIS: Multicast

Objectifs:

- Copier les données de la table **Person.StateProvince** dans 3 fichiers séparés : **Multicast1.txt**, **Multicast2.txt**, **Multicast3.xlsx**



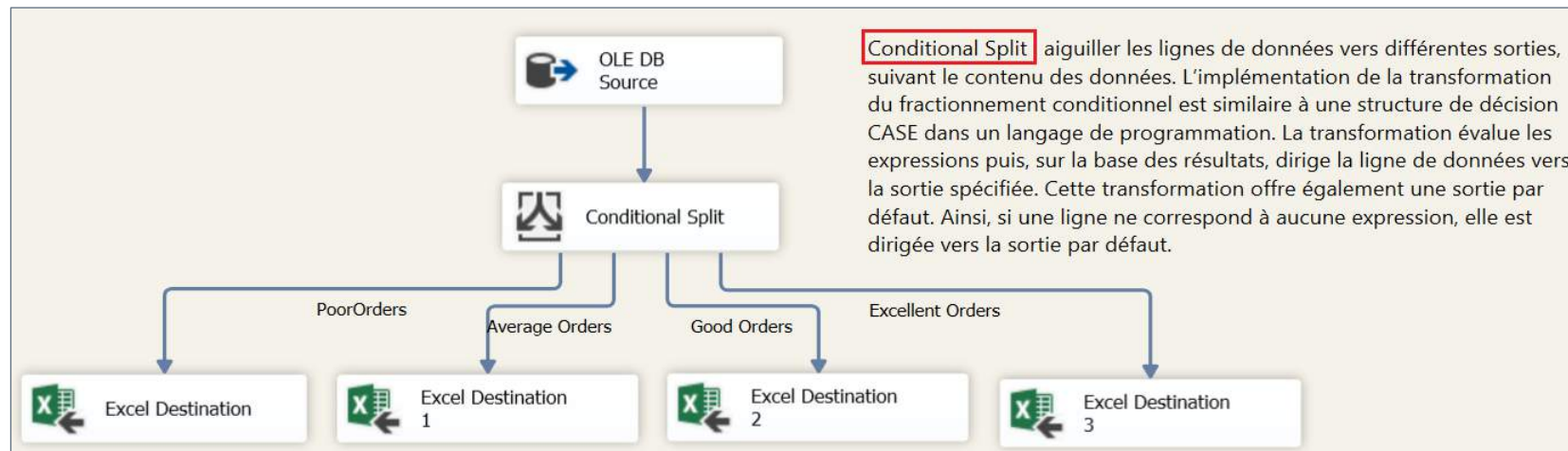
9. Transformation SSIS: Multicast



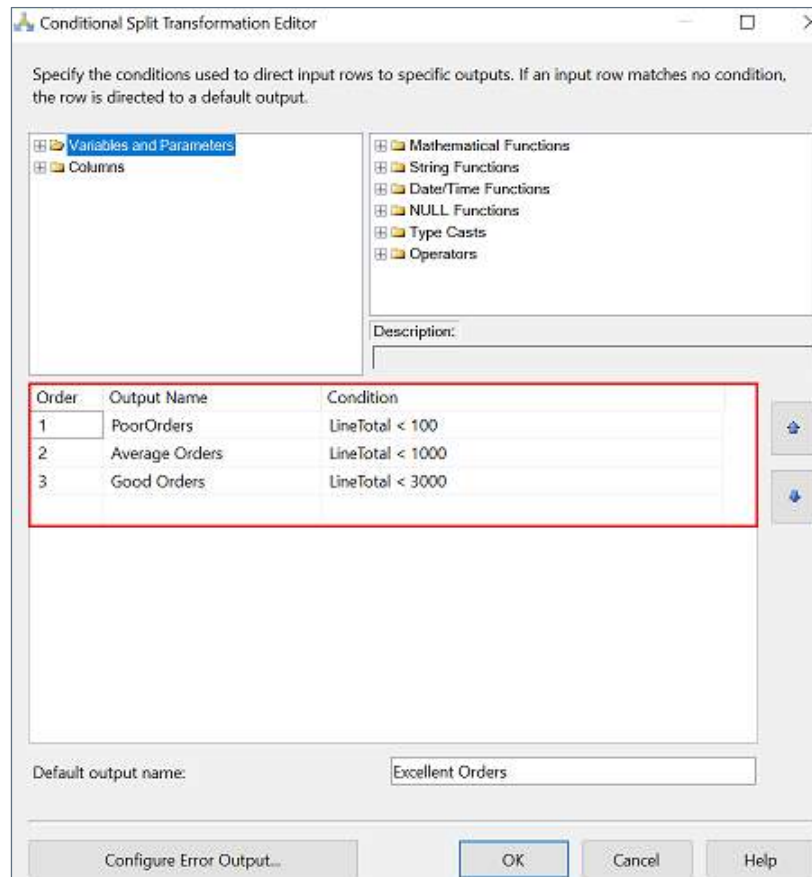
10. Transformation SSIS: Conditional Split

Objectifs:

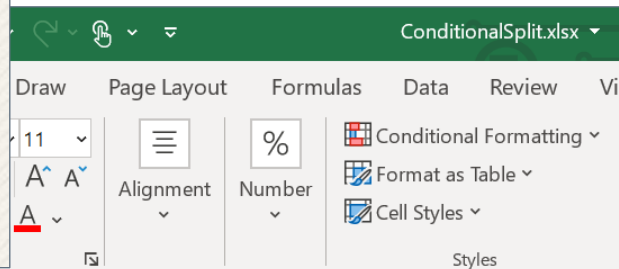
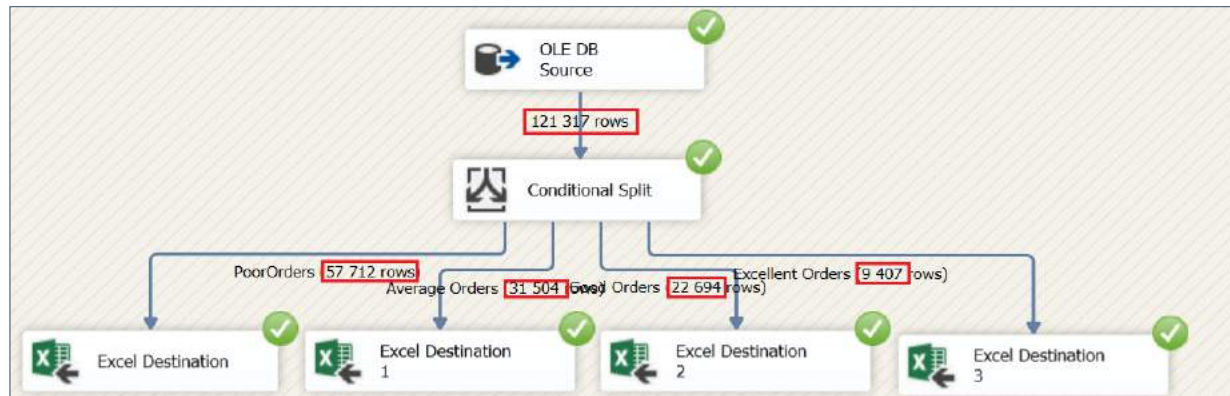
- Charger les données de la table **Sales.SalesOrderDetails** dans 4 feuilles du fichier Excel **ConditionalSplit.xlsx** selon la valeur de la colonne TotalLine :
 - Si **TotalLine<100**: Ajouter la ligne de données à la feuille *PoorOrders*
 - Si **TotalLine<1000**: Ajouter la ligne de données à la feuille *AverageOrders*
 - Si **TotalLine<3000**: Ajouter la ligne de données à la feuille *GoodOrders*
 - Sinon: Ajouter la ligne de données au fichier *ExcellentOrders*



10. Transformation SSIS: Conditional Split



10. Transformation SSIS: Conditional Split



M8

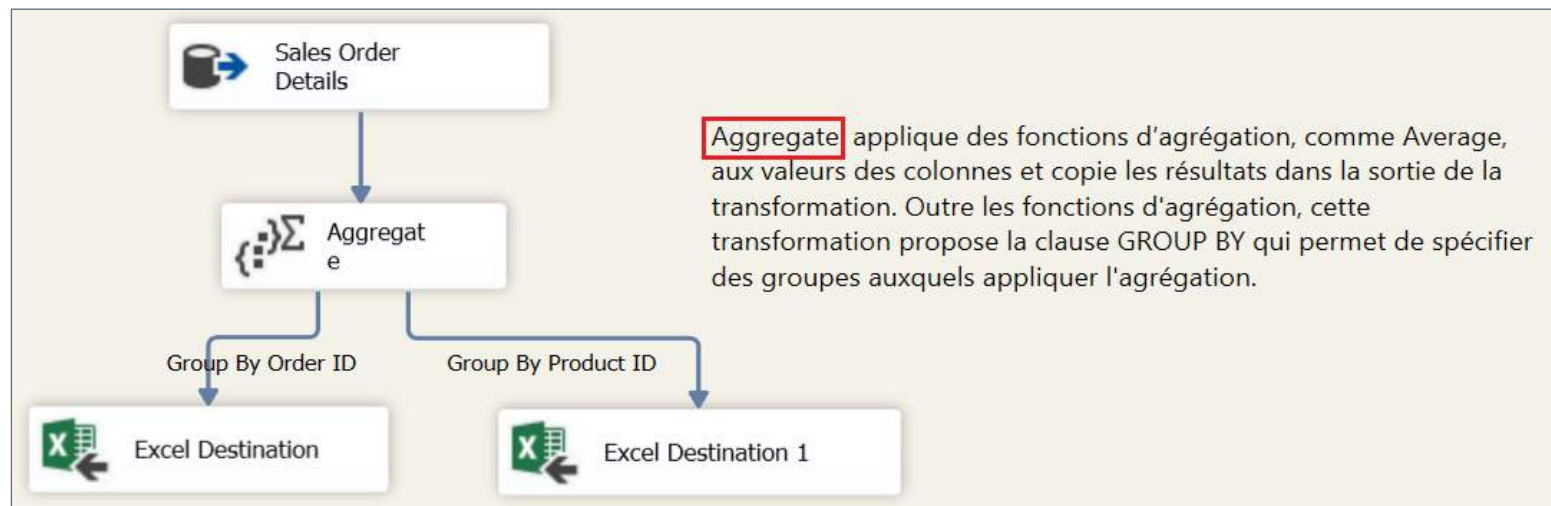
	A	B	C	D	E	F	G	H
1	SalesOrderID	SalesOrderID	OrderQty	ProductID	LineTotal			
2	43659	8	3	714	86.521200			
3	43659	9	1	716	28.840400			
4	43659	10	6	709	34.200000			
5	43659	11	2	712	10.373000			
6	43659	12	4	711	80.746000			
7	43661	18	4	712	20.746000			
8	43661	23	2	711	40.373000			
9	43661	27	2	716	57.680800			
10	43664	55	1	714	28.840400			
11	43664	56	1	716	28.840400			
12	43665	57	2	714	48.373000			

PoorOrders AverageOrders GoodOrders ExcellentOrders

11. Transformation SSIS: Aggregate

Objectifs:

- Charger les données de la table **Sales.SalesOrderDetails** dans 2 feuilles du fichier Excel **AggregateOutput.xlsx** de la manière suivante:
 - La feuille **GrpByOrderID** contient les valeurs SumOfLineTotal, CountLineOrder, AverageLineTotal, MinLineTotal et MaxLineTotal par orderID
 - La feuille **GrpByProdID** contient les valeurs SumOfLineTotal par ProductID.



11. Transformation SSIS: Aggregate

Aggregate Transformation Editor

Aggregations Advanced

Configure the properties used to perform group by operations and to calculate aggregate values. Optionally, apply comparison options to the operation. To configure multiple group by operations, click Advanced.

Aggregation Name	Group By Columns
Group By Order ID	SalesOrderID
Group By Product ID	ProductID

Basic

Available Input Columns

- ☐ Name
- ☐ (*)
- ☒ SalesOrderID
- ☐ SalesOrderDetailID

Input Column	Output Alias	Operation
SalesOrderID	SalesOrderID	Group by
LineTotal	SumOfLineTotal	Sum
LineTotal	Count	Count
LineTotal	AverageLineTotal	Average
LineTotal	MinLineTotal	Minimum
LineTotal	MaxLineTotal	Maximum

OK Cancel Help

Aggregate Transformation Editor

Aggregations Advanced

Configure the properties used to perform group by operations and to calculate aggregate values. Optionally, apply comparison options to the operation. To configure multiple group by operations, click Advanced.

Aggregation Name	Group By Columns
Group By Order ID	SalesOrderID
Group By Product ID	ProductID

Basic

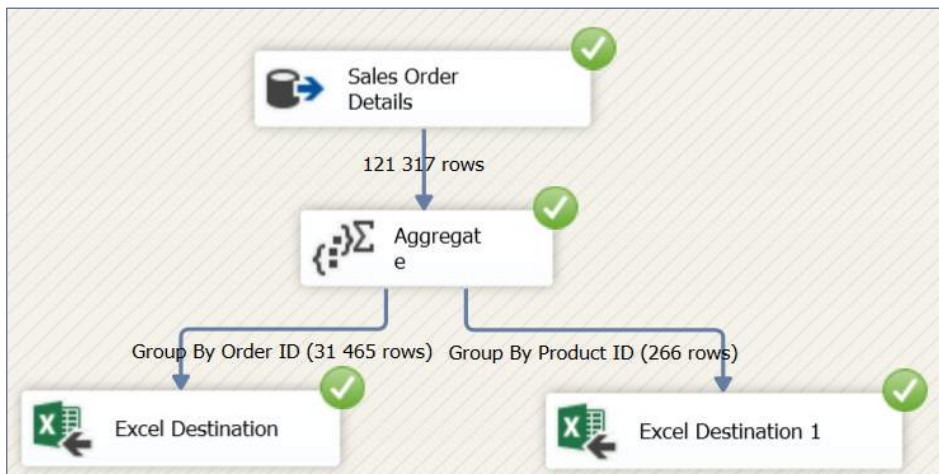
Available Input Columns

- ☐ Name
- ☐ (*)
- ☐ SalesOrderID
- ☐ SalesOrderDetailID
- ☐ OrderQty
- ☒ ProductID
- ☒ LineTotal

Input Column	Output Alias	Operation
ProductID	ProductID	Group by
LineTotal	LineTotal	Sum

OK Cancel Help

11. Transformation SSIS: Aggregate



AutoSave Off AggregateOutput.xlsx

File Home Insert Draw Page Layout Formulas Data Review View

Clipboard Font Alignment Number Styles

H13

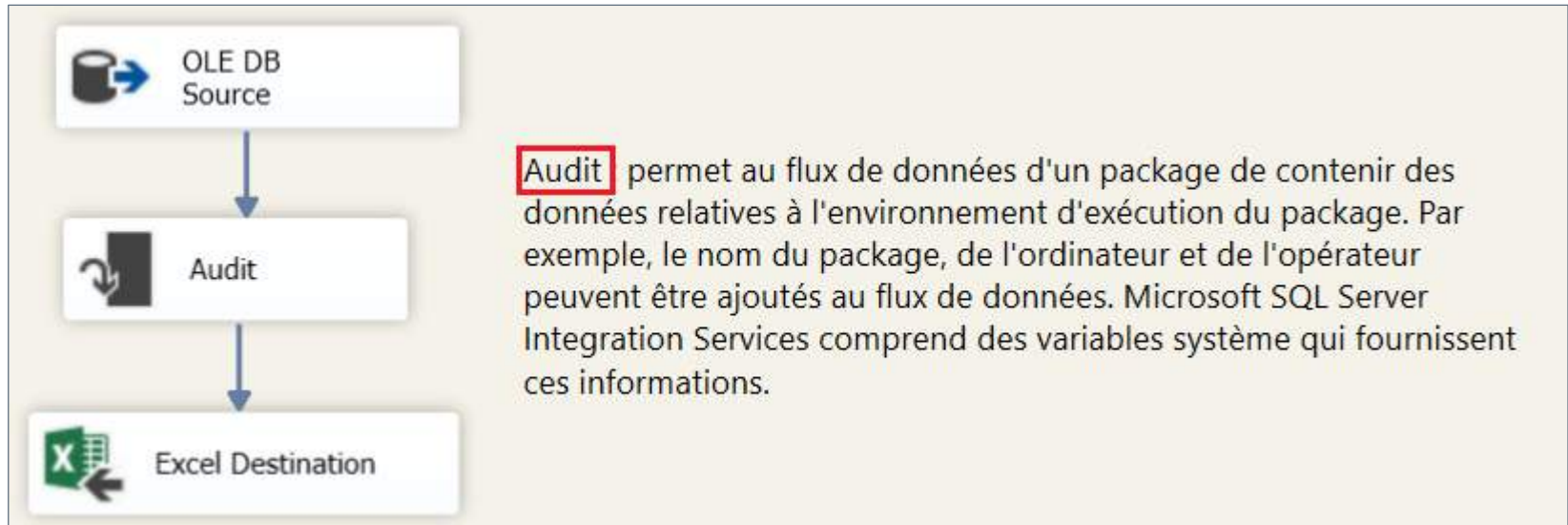
	A	B	C	D	E	F
1	SalesOrder	SumOfLineTotal	Count	AverageLineTotal	MinLineTotal	MaxLineTotal
2	43659	20565.620600	12	1713.801716	10.373000	6074.982000
3	43660	1294.252900	2	647.126450	419.458900	874.794000
4	43661	32726.478600	15	2181.765240	20.746000	8099.976000
5	43662	28832.528900	22	1310.569495	178.580800	5248.764000
6	43663	419.458900	1	419.458900	419.458900	419.458900
7	43664	24432.608800	8	3054.076100	28.840400	8099.976000
8	43665	14352.771300	10	1435.277130	10.373000	4049.988000
9	43666	5056.489600	6	842.748266	356.898000	2146.962000
10	43667	6107.082000	4	1526.770500	17.100000	2039.994000
11	43668	35944.156200	29	1239.453662	20.186500	5248.764000
12	43669	711.781200	1	711.781200	711.781200	711.781200

GrpByOrderID GrpByProdID

| 12. Transformation SSIS: Audit

Objectifs:

- Retourner dans un fichier Excel – **AuditOutput.xlsx** - les données relatives à l'environnement d'exécution du package: **User Name, Task Name, Package Name, Machine Name, Execution Start Time**



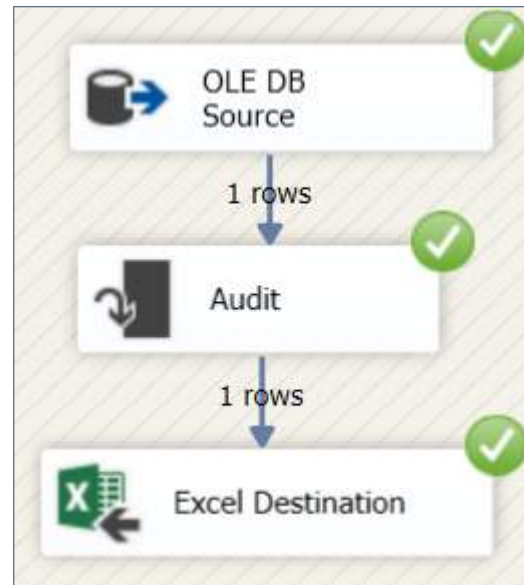
12. Transformation SSIS: Audit

Audit Transformation Editor

Configure the properties used to insert audit information into the data flow.

Output Column Name	Audit Type ▾
User name	User name
Task name	Task name
Package name	Package name
Machine name	Machine name
Execution start time	Execution start time

OK Cancel Help

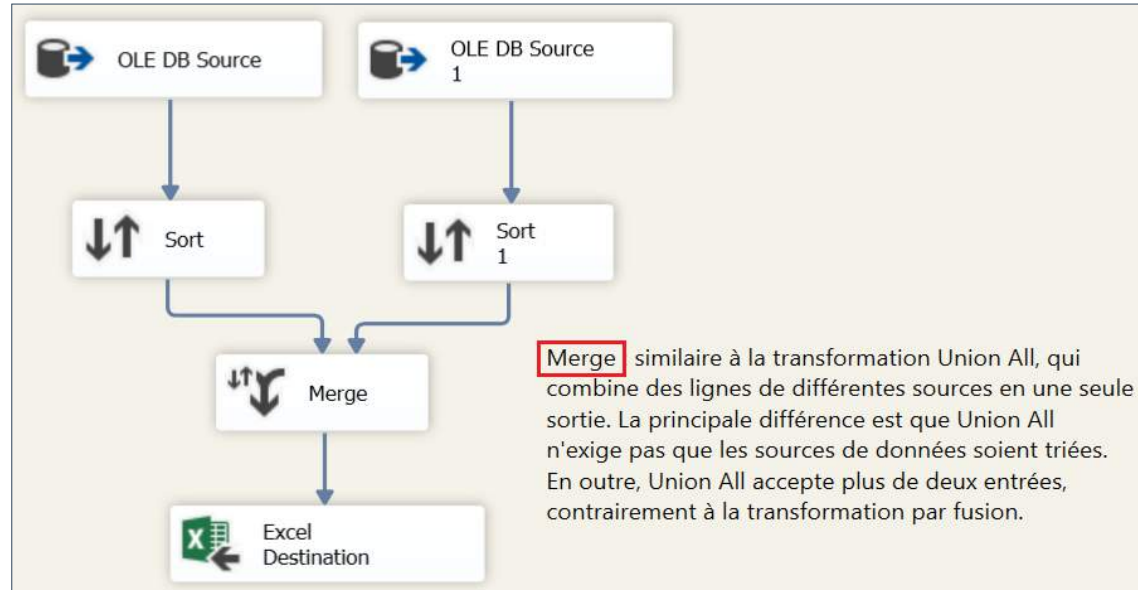


	A	B	C	D	E	F	G
1	CountryRe	StateProvi	Package name	Execution start time	Machine name	User name	Task name
2	FR	87	12_Audit	2021-12-09 22:29:23			Data Flow Task

| 13. Transformation SSIS: Merge

Objectifs:

- Combiner les données triées par *CountryRegionCode* des deux tables **Person.StateProvince** et **Person.NewStateProvinceDetails** dans un fichier Excel trié : **MergeOutput.txt**



13. Transformation SSIS: Merge

Sort Transformation Editor

Specify the columns to sort, and set their sort type and their sort order. All nonselected columns are copied unchanged.

Available Input Columns

<input checked="" type="checkbox"/>	Name	Pass T...
<input type="checkbox"/>	StateProvinceID	<input checked="" type="checkbox"/>
<input type="checkbox"/>	StateProvinceCode	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	CountryRegionCode	<input checked="" type="checkbox"/>
<input type="checkbox"/>	Name	<input checked="" type="checkbox"/>
<input type="checkbox"/>	TerritoryID	<input checked="" type="checkbox"/>

Input Column	Output Alias	Sort Type	Sort Order	C
CountryRegionCode	CountryRegionCode	ascending	1	

☐ Remove rows with duplicate sort values

OK Cancel Help

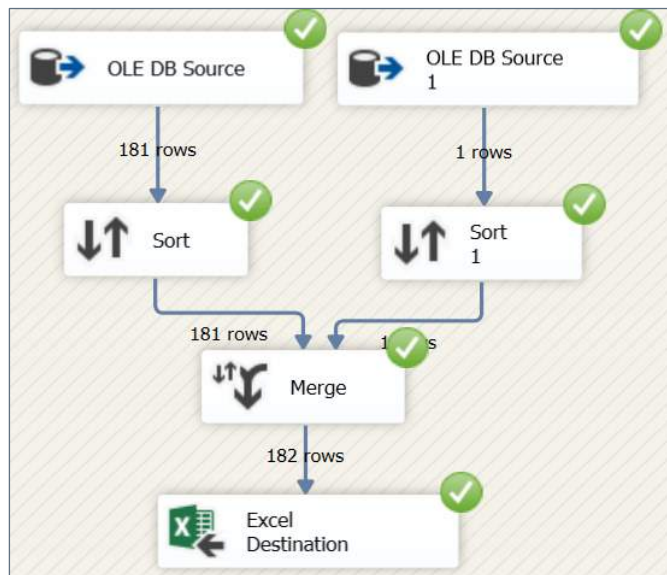
Merge Transformation Editor

Configure the properties used to merge two sorted inputs into one output by creating mappings between columns.

Output Column Name	Merge Input 1	Merge Input 2
StateProvinceID	StateProvinceID	StateProvinceID
StateProvinceCode	StateProvinceCode	StateProvinceCode
CountryRegionCode (Sort key: 1)	CountryRegionCode (Sort key: 1)	CountryRegionCode (Sort key: 1)
Name	Name	Name
TerritoryID	TerritoryID	TerritoryID

OK Cancel Help

13. Transformation SSIS: Merge



AutoSave Off MergeOutput.xlsx

File Home Insert Draw Page Layout Formulas Data Review

Clipboard Font Alignment Number Styles

K19

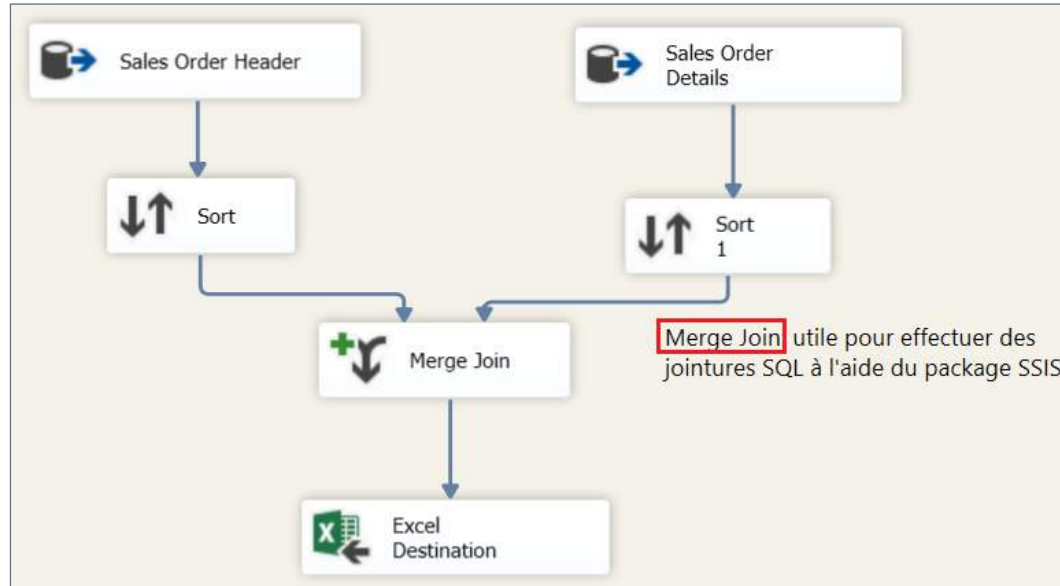
	A	B	C	D	E
1	StateProvi	StateProvi	CountryRegionCode	Name	TerritoryID
2	5	AS	AS	American Samoa	1
3	66	SA	AU	South Australia	9
4	64	QLD	AU	Queensland	9
5	77	VIC	AU	Victoria	9
6	50	NSW	AU	New South Wales	9
7	71	TAS	AU	Tasmania	9
8	31	MB	CA	Manitoba	6
9	41	NB	CA	Brunswick	6
10	69	SK	CA	Saskatchewan	6
11	1	AB	CA	Alberta	6
12	57	ON	CA	Ontario	6

Merge_Output


| 14. Transformation SSIS: MergeJoin

Objectifs:

- Charger le résultat de la jointure des deux tables: **SalesOrderHeader** et **SalesOrderDetail** dans un fichier Excel trié : **MergeJoinOutput.txt**



14. Transformation SSIS: MergeJoin

 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: Inner join Swap Inputs

Sort

<input checked="" type="checkbox"/>	Name	Or...	Joi...
<input checked="" type="checkbox"/>	SalesOrderID	1	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	OrderDate	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	DueDate	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	ShipDate	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	SalesOrderNumber	0	<input type="checkbox"/>

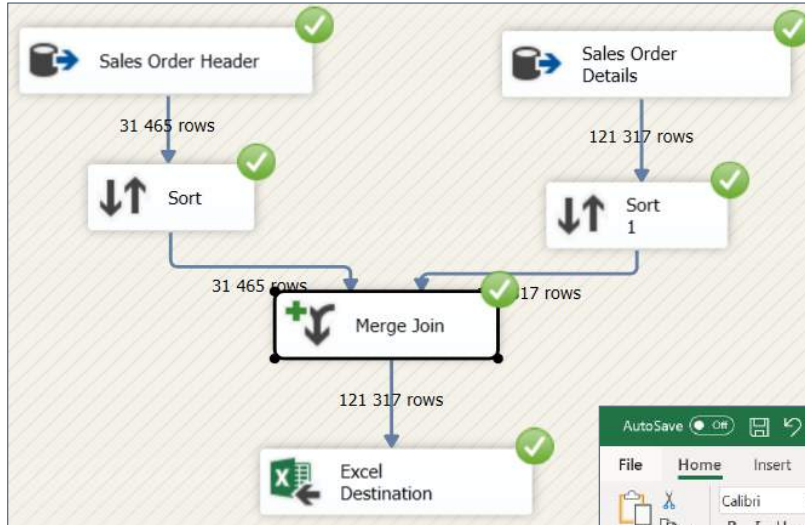
Sort 1

<input type="checkbox"/>	Name	Or...	Joi...
<input type="checkbox"/>	SalesOrderID	1	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	SalesOrderDetailID	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	OrderQty	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	ProductID	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	LineTotal	0	<input type="checkbox"/>

Input	Input Column	Output Alias
Sort	SalesOrderID	SalesOrderID
Sort	OrderDate	OrderDate
Sort	DueDate	DueDate
Sort	ShipDate	ShipDate
Sort	SalesOrderNumber	SalesOrderNumber
Sort 1	SalesOrderDetailID	SalesOrderDetailID
Sort 1	OrderQty	OrderQty
Sort 1	ProductID	ProductID

OK Cancel Help

14. Transformation SSIS: MergeJoin



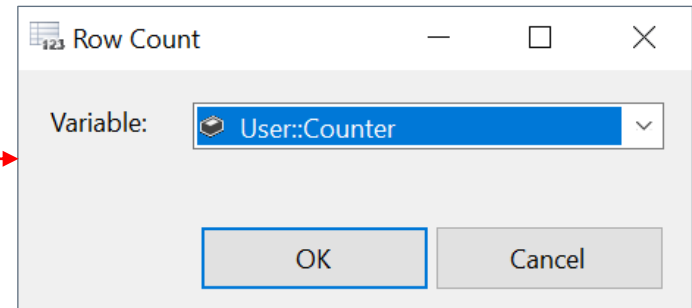
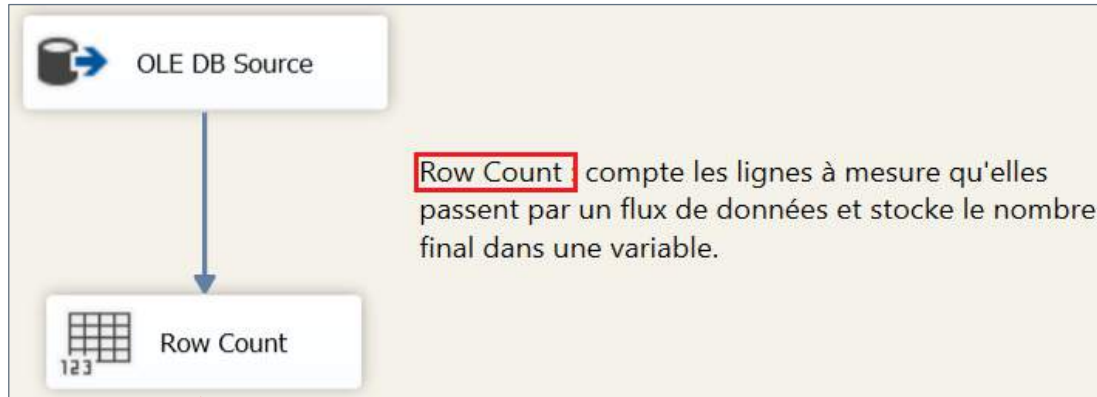
The screenshot shows an Excel spreadsheet titled 'MergeJoinOutput.xlsx'. The data is organized into columns: SalesOrderID, OrderDate, DueDate, ShipDate, SalesOrderNumber, SalesOrderDetailID, OrderQty, ProductID, and LineTotal. The data is sorted by SalesOrderID, with the first 11 rows showing multiple detail records for SalesOrderID 43659.

	A	B	C	D	E	F	G	H	I	J
	SalesOrderID	OrderDate	DueDate	ShipDate	SalesOrderNumber	SalesOrderDetailID	OrderQty	ProductID	LineTotal	
1	43659	2011-05-31 00:00:00	2011-06-12 00:00:00	2011-06-07 00:00:00	SO43659	1	1	776	2024.994000	
2	43659	2011-05-31 00:00:00	2011-06-12 00:00:00	2011-06-07 00:00:00	SO43659	12	4	711	80.746000	
3	43659	2011-05-31 00:00:00	2011-06-12 00:00:00	2011-06-07 00:00:00	SO43659	10	6	709	34.200000	
4	43659	2011-05-31 00:00:00	2011-06-12 00:00:00	2011-06-07 00:00:00	SO43659	9	1	716	28.840400	
5	43659	2011-05-31 00:00:00	2011-06-12 00:00:00	2011-06-07 00:00:00	SO43659	8	3	714	86.521200	
6	43659	2011-05-31 00:00:00	2011-06-12 00:00:00	2011-06-07 00:00:00	SO43659	7	1	774	2039.994000	
7	43659	2011-05-31 00:00:00	2011-06-12 00:00:00	2011-06-07 00:00:00	SO43659	11	2	712	10.373000	
8	43659	2011-05-31 00:00:00	2011-06-12 00:00:00	2011-06-07 00:00:00	SO43659	5	1	772	2039.994000	
9	43659	2011-05-31 00:00:00	2011-06-12 00:00:00	2011-06-07 00:00:00	SO43659	4	1	771	2039.994000	
10	43659	2011-05-31 00:00:00	2011-06-12 00:00:00	2011-06-07 00:00:00	SO43659	3	1	778	2024.994000	
11	43659	2011-05-31 00:00:00	2011-06-12 00:00:00	2011-06-07 00:00:00	SO43659	6	1	773	1678.000000	

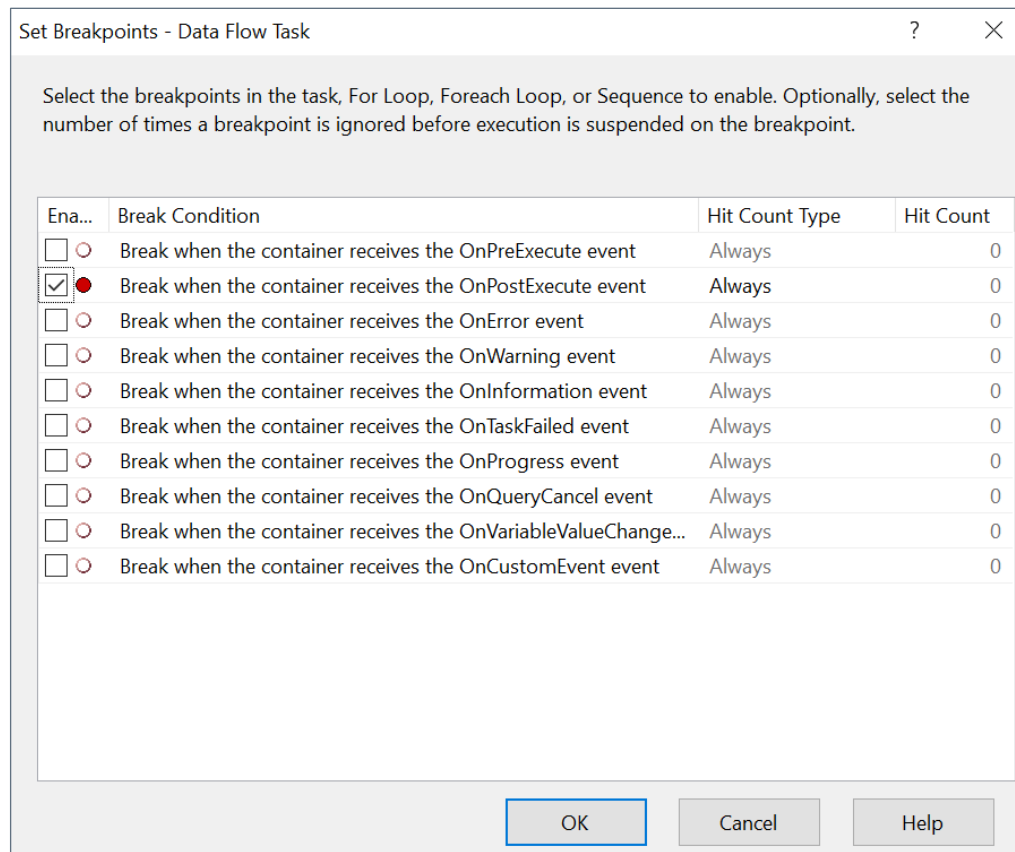
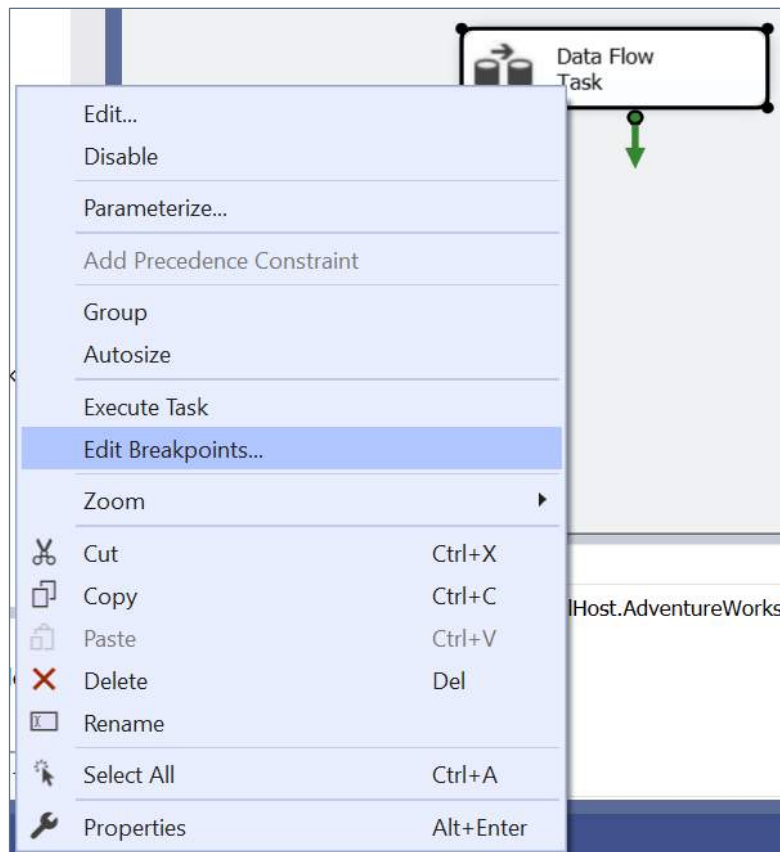
15. Transformation SSIS: Row Count

Objectifs:

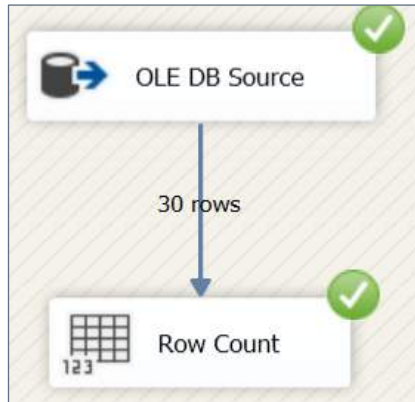
- Enregistrer le nombre de lignes de la table **Occupation** dans une variable **Counter**



15. Transformation SSIS: Row Count



15. Transformation SSIS: Row Count



The main window shows a 'Data Flow Task' with a green arrow pointing down to the 'Locals' window. The 'Locals' window displays the following information:

Connection Managers

- Flat File Connection Manager
- LocalHost.AdventureWorks2019
- Excel Connection Manager

Locals

Search (Ctrl+E) 🔍 Search Depth: ▾

Name	Value
▶ User::Counter	{30}

16. Transformation SSIS: Lookup

SalesOrderDetail

Available External Columns

<input type="checkbox"/>	Name
<input checked="" type="checkbox"/>	SalesOrderID
<input checked="" type="checkbox"/>	SalesOrderDetailID
<input type="checkbox"/>	CarrierTrackingNumber
<input checked="" type="checkbox"/>	OrderQty
<input checked="" type="checkbox"/>	ProductID
<input type="checkbox"/>	SpecialOfferID
<input type="checkbox"/>	UnitPrice
<input type="checkbox"/>	UnitPriceDiscount
<input checked="" type="checkbox"/>	LineTotal
<input type="checkbox"/>	rowguid
<input type="checkbox"/>	ModifiedDate

LookupOutput.xlsx

File Home Insert Draw Page Layout Formulas Data Review View Developer Help Inquire Power Pivot

Clipboard Font Alignment Number Styles Cells Editing Ideas Sensitivity Share Upload

	A	B	C	D	E	F	G	H	I	J	K
1	SalesOrderID	SalesOrderDetailID	OrderQty	ProductID	LineTotal	Name	ProductNumber	Color	ListPrice		
2	43659	1	1	776	2024.994000	Mountain-100 Black, 42	BK-M82B-42	Black	3374.99		
3	43659	2	3	777	6074.982000	Mountain-100 Black, 44	BK-M82B-44	Black	3374.99		
4	43659	3	1	778	2024.994000	Mountain-100 Black, 48	BK-M82B-48	Black	3374.99		
5	43659	4	1	771	2039.994000	Mountain-100 Silver, 38	BK-M82S-38	Silver	3399.99		
6	43659	5	1	772	2039.994000	Mountain-100 Silver, 42	BK-M82S-42	Silver	3399.99		
7	43659	6	2	773	4079.988000	Mountain-100 Silver, 44	BK-M82S-44	Silver	3399.99		
8	43659	7	1	774	2039.994000	Mountain-100 Silver, 48	BK-M82S-48	Silver	3399.99		
9	43659	8	3	714	86.521200	Long-Sleeve Logo Jersey, M	LJ-0192-M	Multi	49.99		
10	43659	9	1	716	28.840400	Long-Sleeve Logo Jersey, XL	LJ-0192-X	Multi	49.99		
11	43659	10	6	709	34.200000	Mountain Bike Socks, M	SO-B909-M	White	9.5		

Lookup_Output Fail

Product

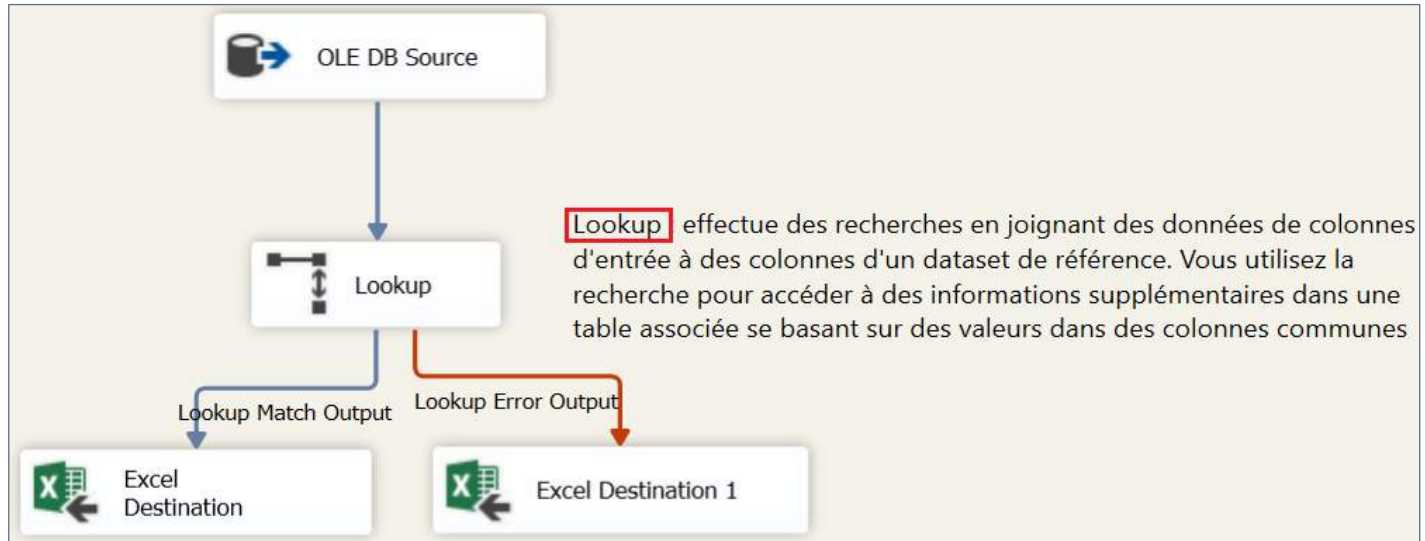
Available Lookup Columns

<input type="checkbox"/>	Name	I...
<input type="checkbox"/>	ProductID	
<input checked="" type="checkbox"/>	Name	
<input checked="" type="checkbox"/>	ProductNumber	
<input type="checkbox"/>	MakeFlag	
<input type="checkbox"/>	FinishedGood...	
<input checked="" type="checkbox"/>	Color	
<input type="checkbox"/>	SafetyStockL...	
<input type="checkbox"/>	ReorderPoint	
<input type="checkbox"/>	StandardCost	
<input checked="" type="checkbox"/>	ListPrice	
<input type="checkbox"/>	Size	

| 16. Transformation SSIS: Lookup

Objectifs:

- Pour chaque order de ta table **SalesOrderDetail**, on veut retourner aussi les details du produit commandé depuis la table **Products**. Le résultat doit être chargé dans un fichier Excel **LookupOutput**



16. Transformation SSIS: Lookup

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.AdventureWorks2019] New...

Data access mode:
Table or view

Name of the table or the view:
[Sales].[SalesOrderDetail]

Preview...

OK Cancel Help

Lookup Transformation Editor

This transform enables the performance of simple equi-joins between the input and a reference data set.

General
Connection
Columns
Advanced
Error Output

Cache mode
☒ Full cache
☐ Partial cache
☐ No cache

Connection type
☐ Cache connection manager
☒ OLE DB connection manager

Specify how to handle rows with no matching entries
Redirect rows to error output

OK Cancel Help

| 16. Transformation SSIS: Lookup

Cache modes

- **Full Cache** : Si les données de référence sont très importantes, nous pouvons utiliser Full Cache. Pour qu'il n'y ait pas de déplacements vers la base de données pour obtenir les données de SQL vers SSIS. Des données entières sont capturées avant d'exécuter le Lookup. Ce mode est le mode préféré.

A ne pas utiliser quand:

- *l'ensemble de données de recherche est très volumineux*
- *OU lorsque le nombre d'enregistrements d'entrée est vraiment petit par rapport à l'ensemble de données de recherche.*

Exemple :

- Données Source: 2 enregistrements.
- Données de référence: 200 millions enregistrements.

=> SSIS va charger les 200 millions enregistrements en mémoire, puis va faire le lookup. => Consommation de la mémoire => Ne pas utiliser le Full cache

| 16. Transformation SSIS: Lookup

Cache modes

- **No cache**: SSIS ne conserve aucun cache et envoie une requête à la base de données de référence pour chaque enregistrement qu'il doit rechercher.

Quand l'utiliser?

- *Le nombre d'enregistrements dans l'ensemble de données source est très petit (et le restera)*
- *ET l'ensemble de données de référence est très volumineux (sinon, utilisez simplement le mode cache complet)*

Exemple :

- Données Source: 2 enregistrements.
- Données de référence: 200 millions enregistrements.

=> SSIS va envoyer la requête à la BD 2 fois uniquement => No Cache est la meilleure option

| 16. Transformation SSIS: Lookup

Cache modes

- **Partial cache**: si une correspondance est trouvée, la clé et les valeurs de recherche sont ajoutées au cache. Si cette même clé entre à nouveau dans le composant de recherche, elle peut récupérer la valeur correspondante du cache local au lieu de la table de référence

Quand l'utiliser?

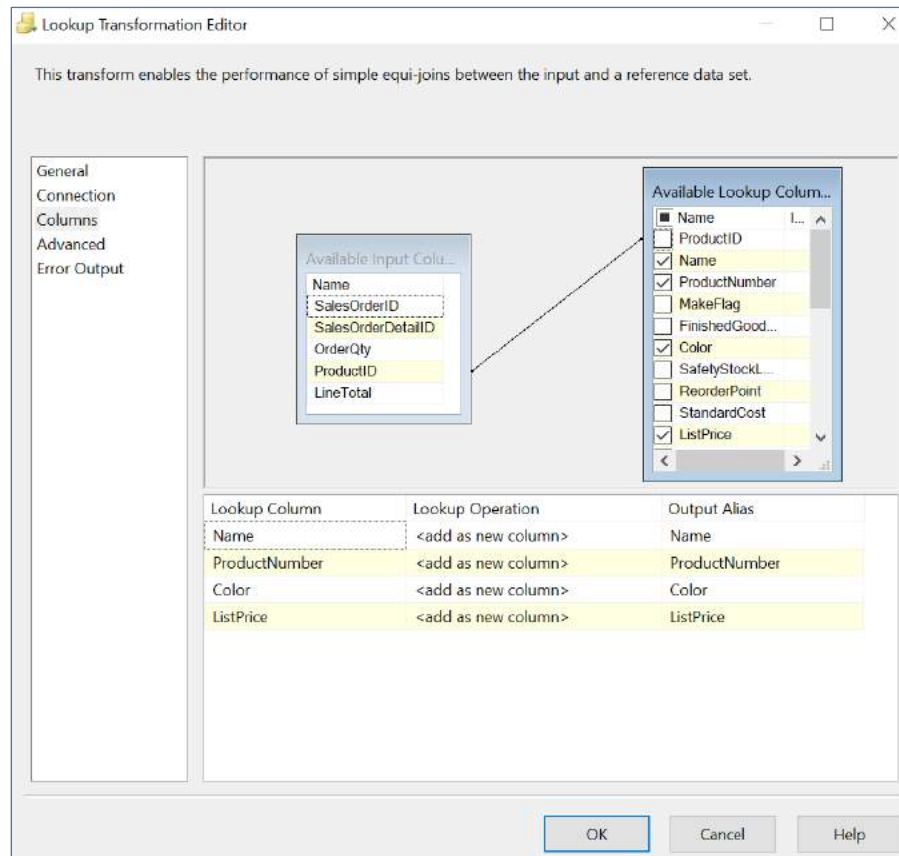
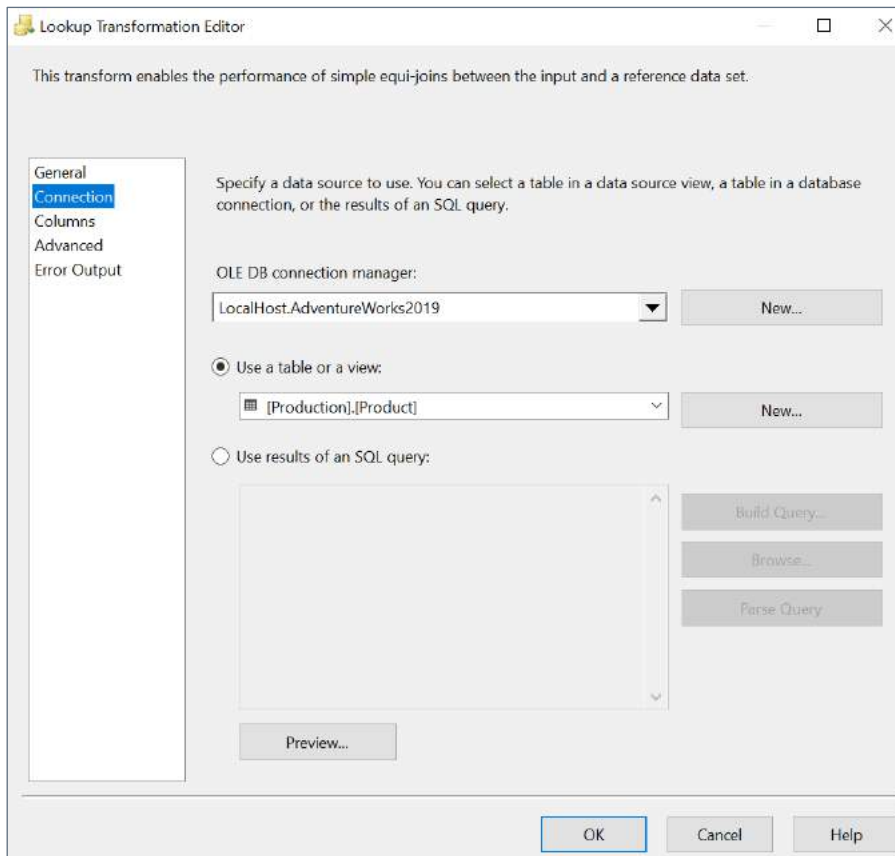
- *Le nombre de combinaisons de recherche uniques dans l'ensemble de données source est faible*
- *ET l'ensemble de données de référence est très volumineux (sinon, utilisez simplement le mode cache complet)*

Exemple :

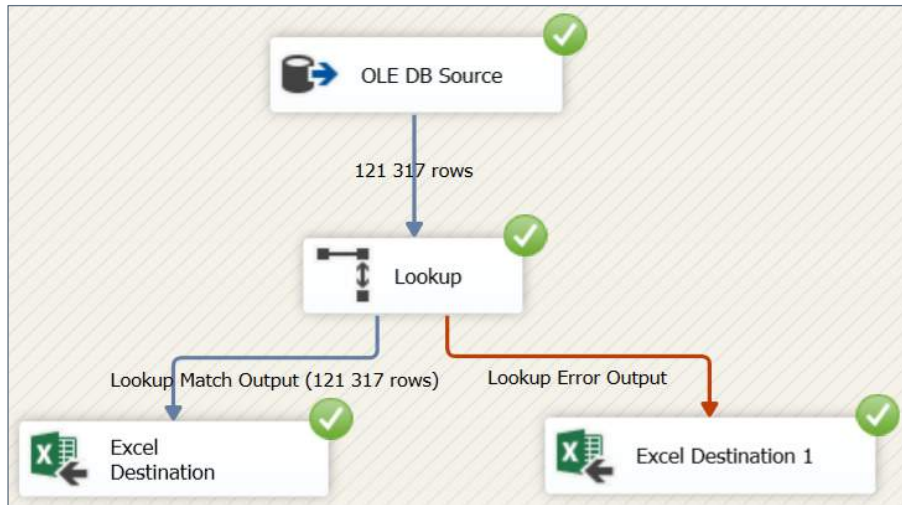
- Données Source: 10.000 enregistrements avec 100 valeurs uniques seulement
- Données de référence: 200 millions enregistrements.

=> Partial cache est la meilleure option

16. Transformation SSIS: Lookup



16. Transformation SSIS: Lookup



LookupOutput.xlsx

Khawla Elansari

General

Conditional Formatting

Format as Table

Cell Styles

Clipboard

Font

Alignment

Number

Styles

Cells

Editing

Ideas

Sensitivity

Share

Upload

K17

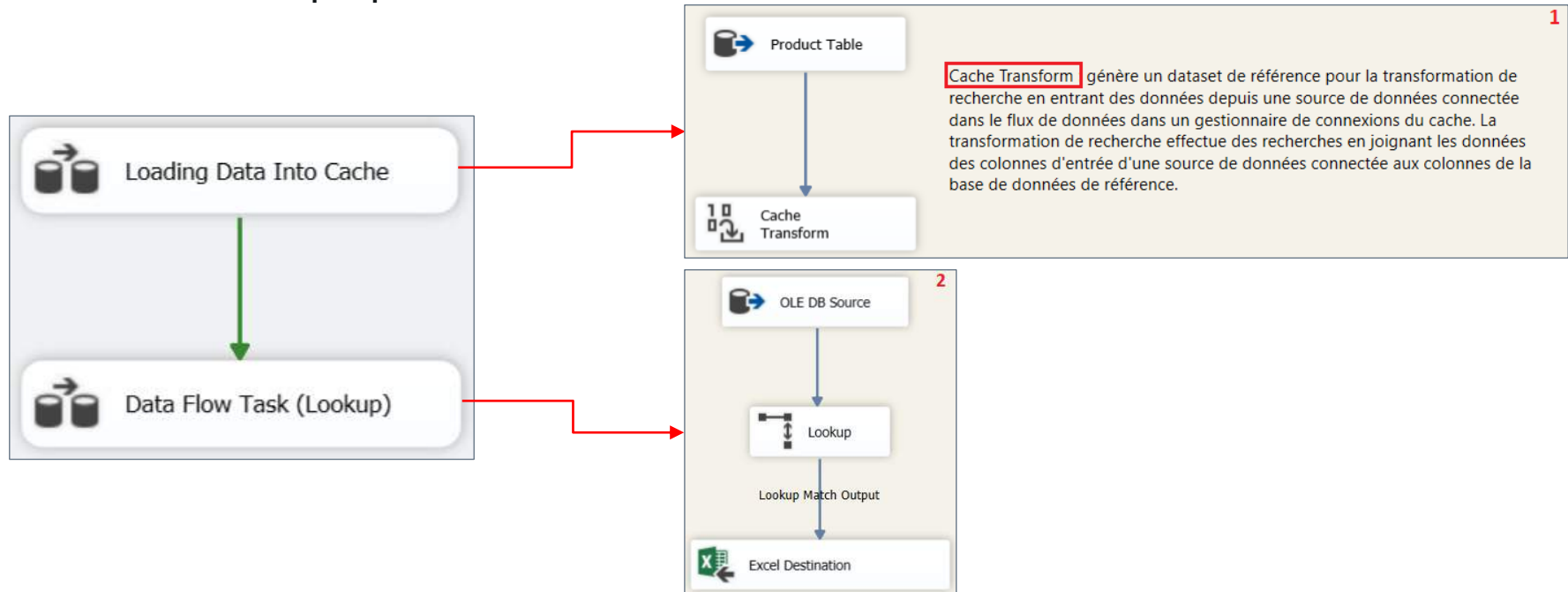
	A	B	C	D	E	F	G	H	I	J	K
1	SalesOrderID	SalesOrderDetailID	OrderQty	ProductID	LineTotal	Name	ProductNumber	Color	ListPrice		
2	43659	1	1	776	2024.994000	Mountain-100 Black, 42	BK-M82B-42	Black	3374.99		
3	43659	2	3	777	6074.982000	Mountain-100 Black, 44	BK-M82B-44	Black	3374.99		
4	43659	3	1	778	2024.994000	Mountain-100 Black, 48	BK-M82B-48	Black	3374.99		
5	43659	4	1	771	2039.994000	Mountain-100 Silver, 38	BK-M82S-38	Silver	3399.99		
6	43659	5	1	772	2039.994000	Mountain-100 Silver, 42	BK-M82S-42	Silver	3399.99		
7	43659	6	2	773	4079.988000	Mountain-100 Silver, 44	BK-M82S-44	Silver	3399.99		
8	43659	7	1	774	2039.994000	Mountain-100 Silver, 48	BK-M82S-48	Silver	3399.99		
9	43659	8	3	714	86.521200	Long-Sleeve Logo Jersey, M	LJ-0192-M	Multi	49.99		
10	43659	9	1	716	28.840400	Long-Sleeve Logo Jersey, XL	LJ-0192-X	Multi	49.99		
11	43659	10	6	709	34.200000	Mountain Bike Socks, M	SO-B909-M	White	9.5		

Lookup_Output Fail

17. Transformation SSIS: Cache Transform

Objectifs:

- Charger les données de la table **SalesOrderDetail** dans un cache
- Utiliser un lookup pour charger les détails du produit commandé depuis la table **Products**. Le résultat doit être chargé dans un fichier Excel **LookupOutput**



17. Transformation SSIS: Cache Transform

1- Loading data into Cache

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.AdventureWorks2019 New...

Data access mode:
Table or view ▼

Name of the table or the view:
⌵ [Production].[Product] ▼

Preview...

OK Cancel Help

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.

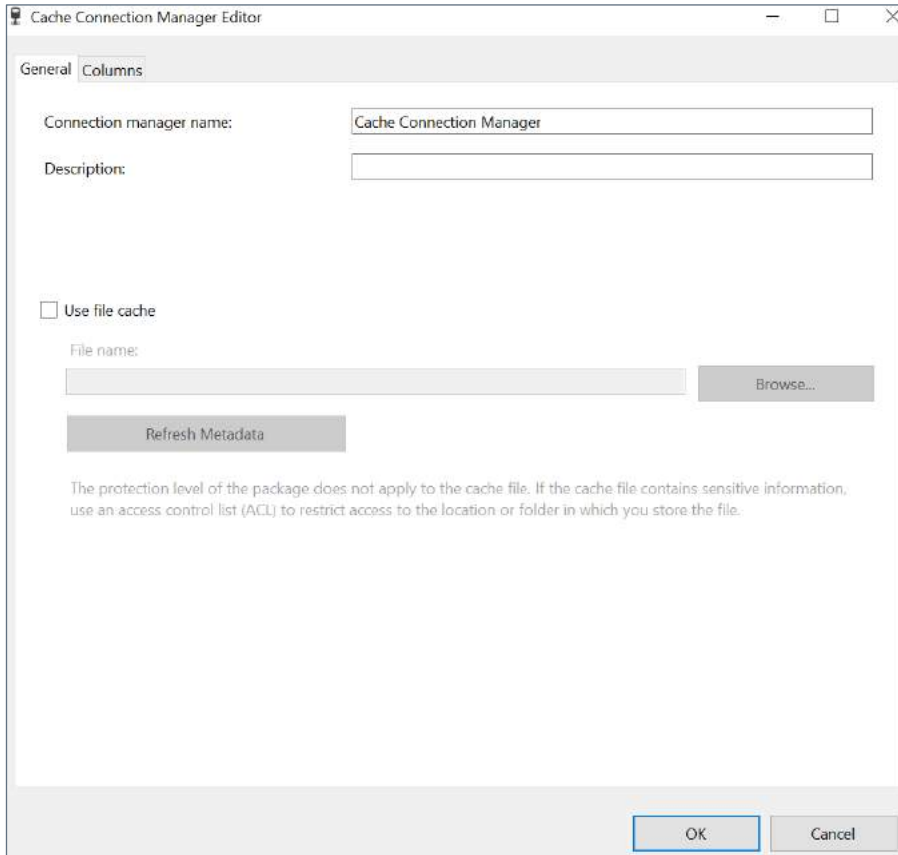
Cache connection manager:
Cache Connection Manager New...

Edit...

OK Cancel Help

17. Transformation SSIS: Cache Transform

1- Loading data into Cache



Cache Connection Manager Editor

General Columns

Connection manager name: Cache Connection Manager

Description:

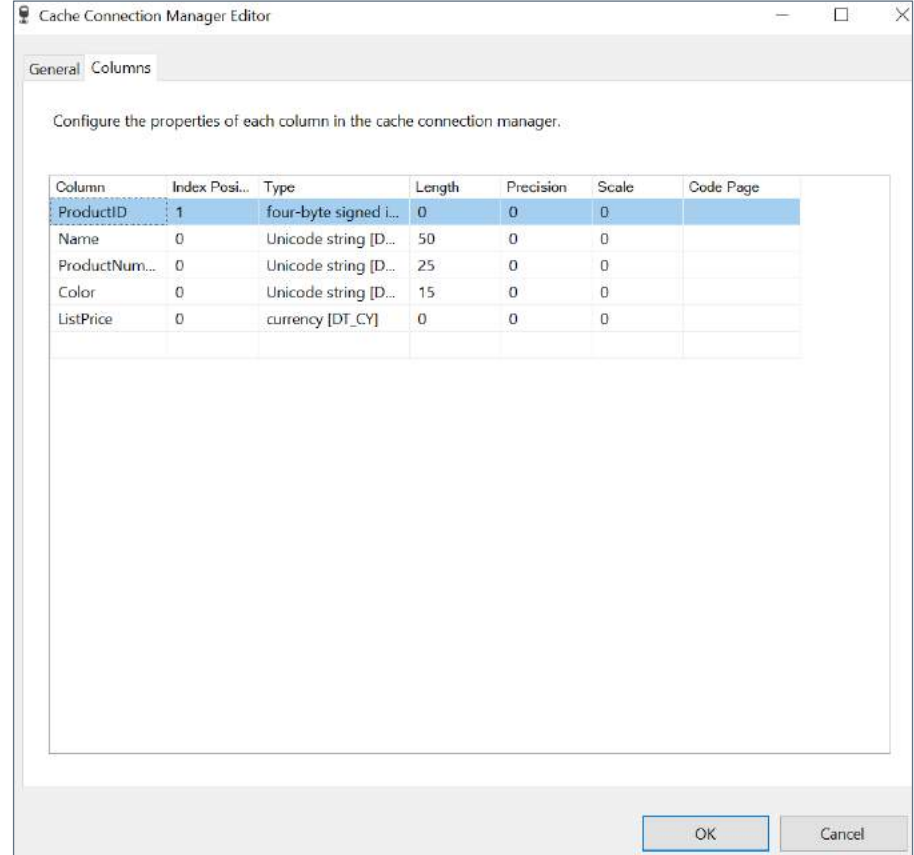
☐ Use file cache

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.

OK Cancel



Cache Connection Manager Editor

General Columns

Configure the properties of each column in the cache connection manager.

Column	Index Posi...	Type	Length	Precision	Scale	Code Page
ProductID	1	four-byte signed i...	0	0	0	
Name	0	Unicode string [D...	50	0	0	
ProductNum...	0	Unicode string [D...	25	0	0	
Color	0	Unicode string [D...	15	0	0	
ListPrice	0	currency [DT_CY]	0	0	0	

OK Cancel

17. Transformation SSIS: Cache Transform

2- Data Flow Task (Lookup)

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:
LocalHostAdventureWorks2019 New...

Data access mode:
Table or view

Name of the table or the view:
[Sales].[SalesOrderDetail]

Preview...

OK Cancel Help

Lookup Transformation Editor

This transform enables the performance of simple equi-joins between the input and a reference data set.

General
Connection
Columns
Advanced
Error Output

Cache mode
☒ Full cache
☐ Partial cache
☐ No cache

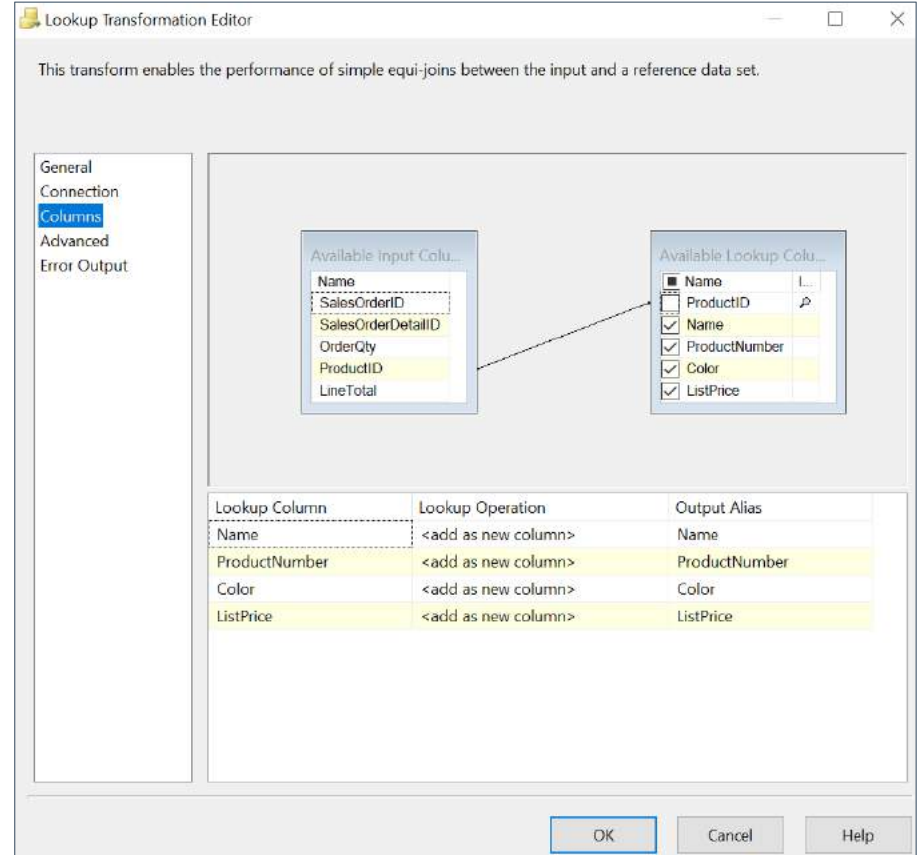
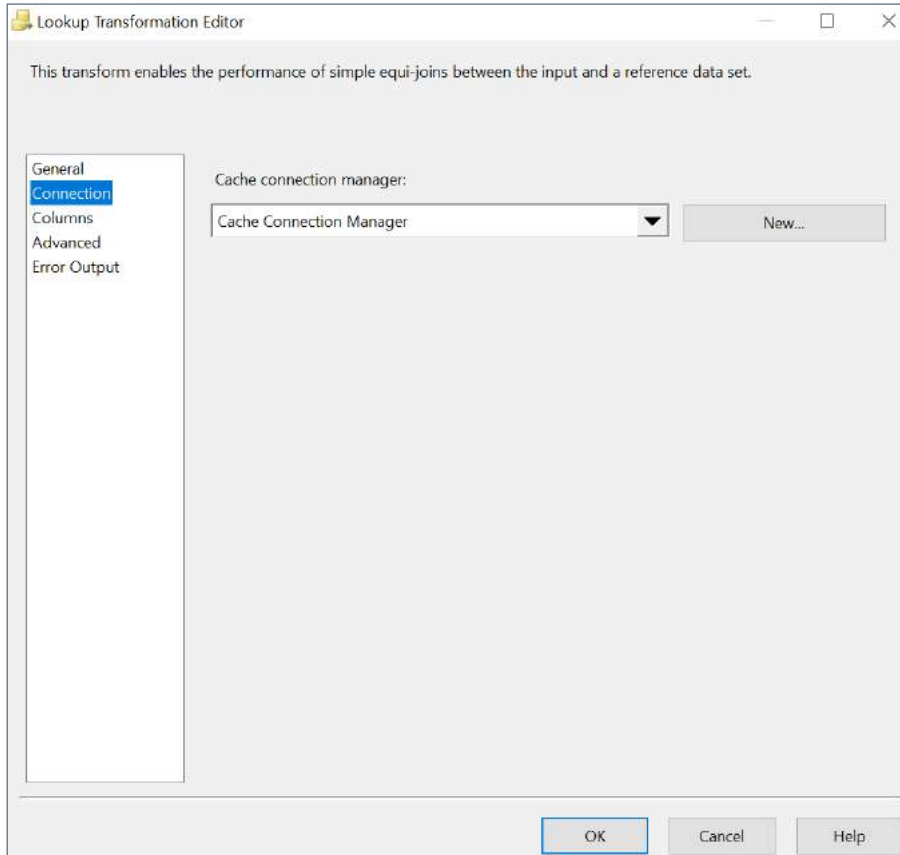
Connection type
☒ Cache connection manager
☐ OLE DB connection manager

Specify how to handle rows with no matching entries
Redirect rows to no match output

OK Cancel Help

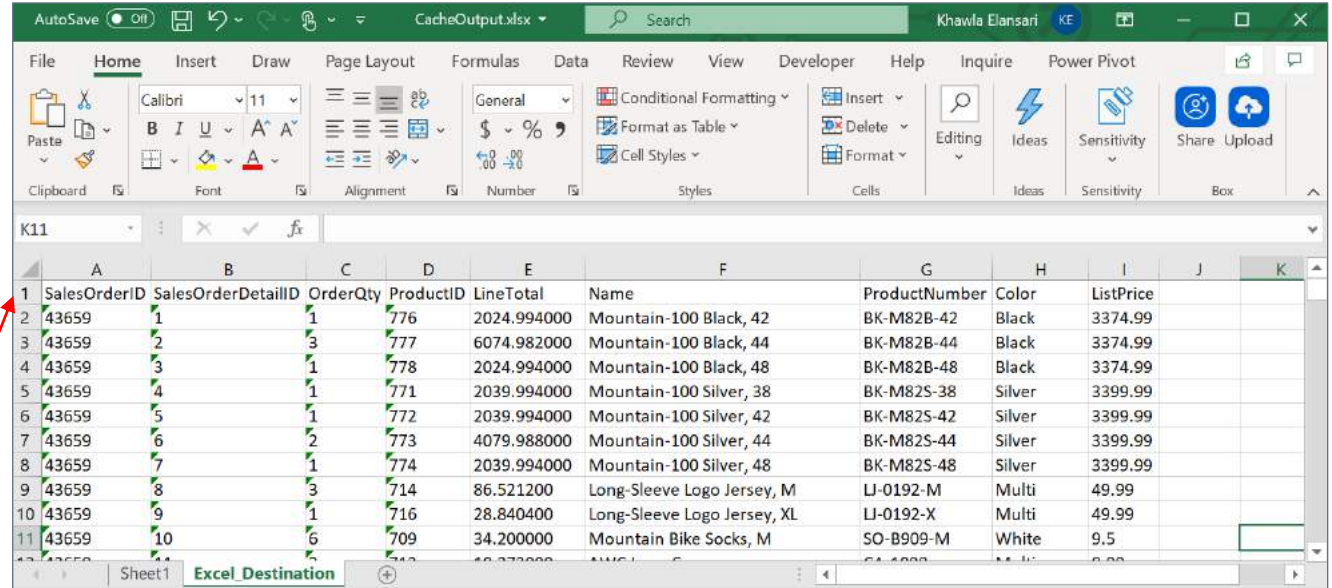
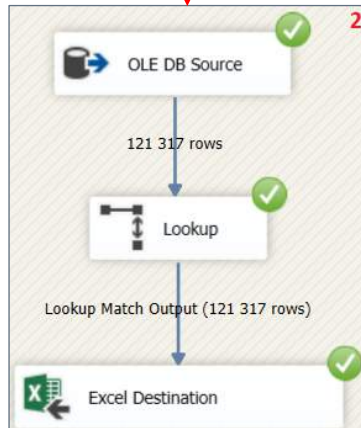
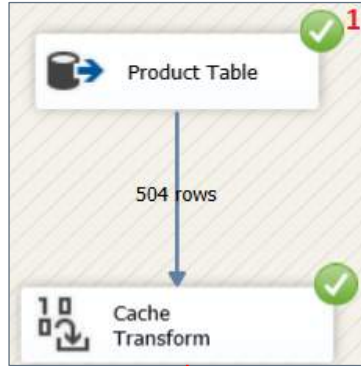
17. Transformation SSIS: Cache Transform

2- Data Flow Task (Lookup)



17. Transformation SSIS: Cache Transform

2- Data Flow Task (Lookup)



The screenshot shows the Microsoft Excel interface with the file 'CacheOutput.xlsx' open. The data is displayed in a table with columns A through K. A red arrow points from the 'Lookup' task in Diagram 2 to the first row of data in the Excel table.

	A	B	C	D	E	F	G	H	I	J	K
1	SalesOrderID	SalesOrderDetailID	OrderQty	ProductID	LineTotal	Name	ProductNumber	Color	ListPrice		
2	43659	1	1	776	2024.994000	Mountain-100 Black, 42	BK-M82B-42	Black	3374.99		
3	43659	2	3	777	6074.982000	Mountain-100 Black, 44	BK-M82B-44	Black	3374.99		
4	43659	3	1	778	2024.994000	Mountain-100 Black, 48	BK-M82B-48	Black	3374.99		
5	43659	4	1	771	2039.994000	Mountain-100 Silver, 38	BK-M82S-38	Silver	3399.99		
6	43659	5	1	772	2039.994000	Mountain-100 Silver, 42	BK-M82S-42	Silver	3399.99		
7	43659	6	2	773	4079.988000	Mountain-100 Silver, 44	BK-M82S-44	Silver	3399.99		
8	43659	7	1	774	2039.994000	Mountain-100 Silver, 48	BK-M82S-48	Silver	3399.99		
9	43659	8	3	714	86.521200	Long-Sleeve Logo Jersey, M	LJ-0192-M	Multi	49.99		
10	43659	9	1	716	28.840400	Long-Sleeve Logo Jersey, XL	LJ-0192-X	Multi	49.99		
11	43659	10	6	709	34.200000	Mountain Bike Socks, M	SO-B909-M	White	9.5		

18. Transformation SSIS: FuzzyLookup

Objectifs:

- Pour chaque employée du fichier **FuzzyExample**, on veut retourner aussi les détails de l'occupation depuis la table **Occupations**. Le résultat doit être chargé dans un fichier Excel **FuzzyLookupOutput**

OccupationID	OccupationTitle
1	CUSTOMER SERVICE REPRESENTATIVE
2	SHIFT LEADER
3	ASSISTANT MANAGER
4	STORE MANAGER
5	DISTRICT MANAGER
6	REGIONAL MANAGER
102	CUSTOMER SERVICE REPRESENTATIVE
103	SHIFT LEADER
104	ASSISTANT MANAGER
105	STORE MANAGER
106	DISTRICT MANAGER
107	REGIONAL MANAGER

FuzzyExample.txt - Notepad

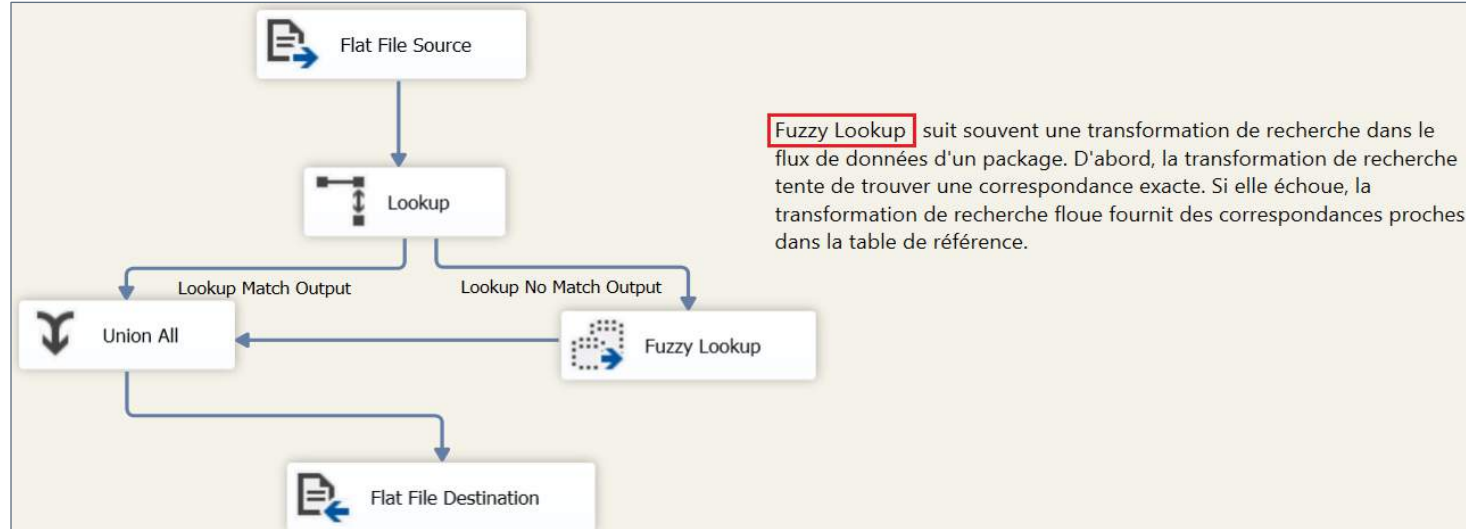
File Edit Format View Help

LASTNAME,FIRSTNAME,TITLE
HENDERSON,HARRY,CUST SERVICE REP.
WALKER,JEREMY,CUSTOMER SERVICE REPRESENTATIVE
ALBREKTSON,CHRISTOPHER,SHIFT LEAD
MCFLY,MARTY,REGION MANAGER
OWENS,BENJAMIN,ASSISTANT MANAGER
SCOTT,MICHAEL,REGIONAL MANAGER
HALPERT,JIM,STORE MGR.
FOSTER,DERRICK,CUSTOMER SERVICE REP
SCHRUTE,DWIGHT,CUSTOMER SERVICE REPRESENTATIVE
PAIGE,CARRINGTON,STORE MANAGER
SCHACHT,BRADLEY,DISTRICT MANAGER
BEESLY,PAM,ASSIST. MANAGER
LANDEN,COLLIN,CUSTOMER SERVICE REPRESENTATIVE
SHOWALTER,CHRISTOPHER,SHIFT LEADER
MCNAMARA,SEAN,REGIONAL MANAGER
COLBERT,STEPHEN,STORE MANAGER
RYAN,DUSTIN,DISTRICT MGR.
TROY,CHRISTIAN,REGIONAL MGR
ISHEE,SAM,SHIFT LEADER
WOOD,JOHN,CUSTOMER SERVICE REP
ENGLE,DANIEL,ASSISTANT MANAGER
STEWART,JON,DISTRICT MANAGER

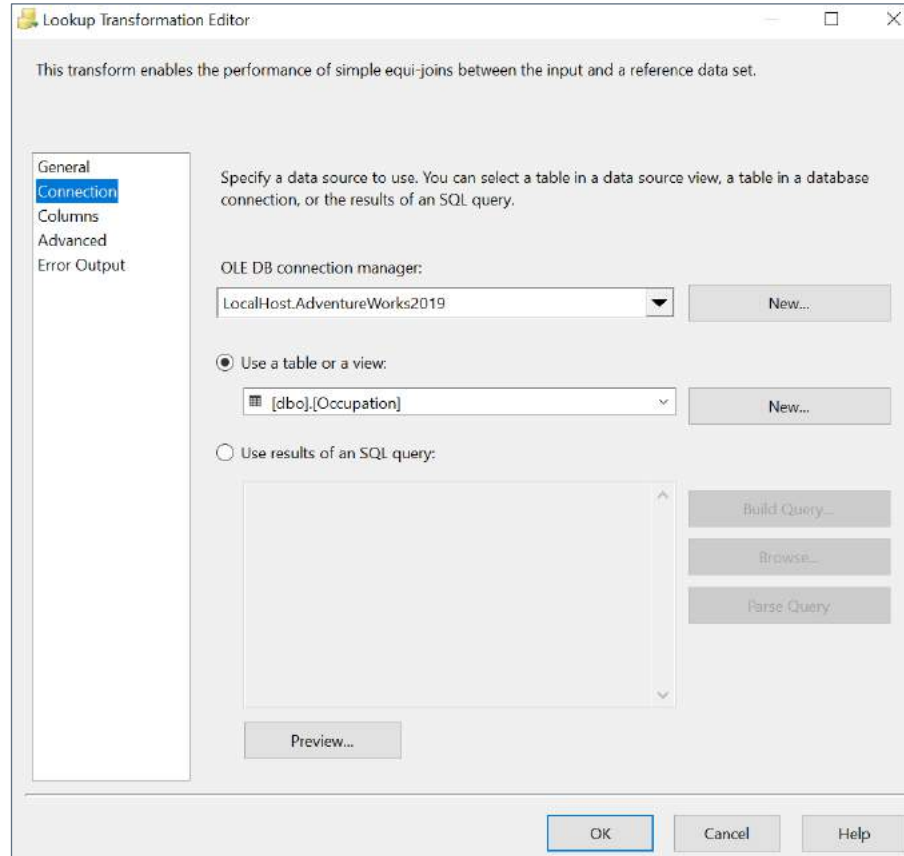
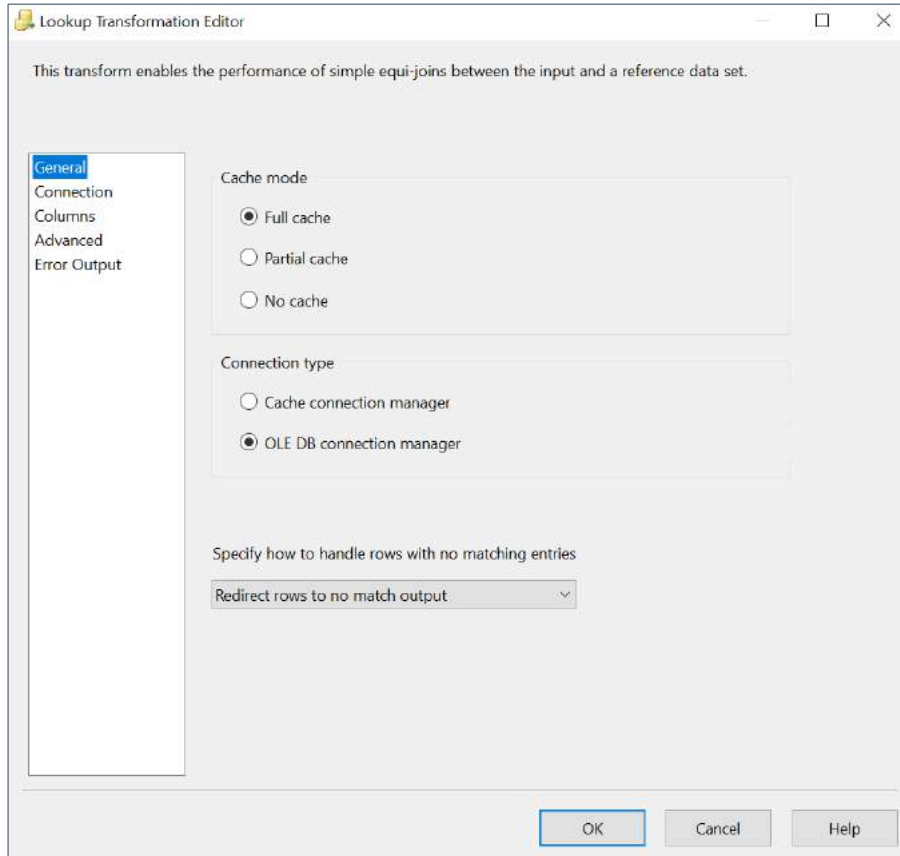
| 18. Transformation SSIS: FuzzyLookup

Objectifs:

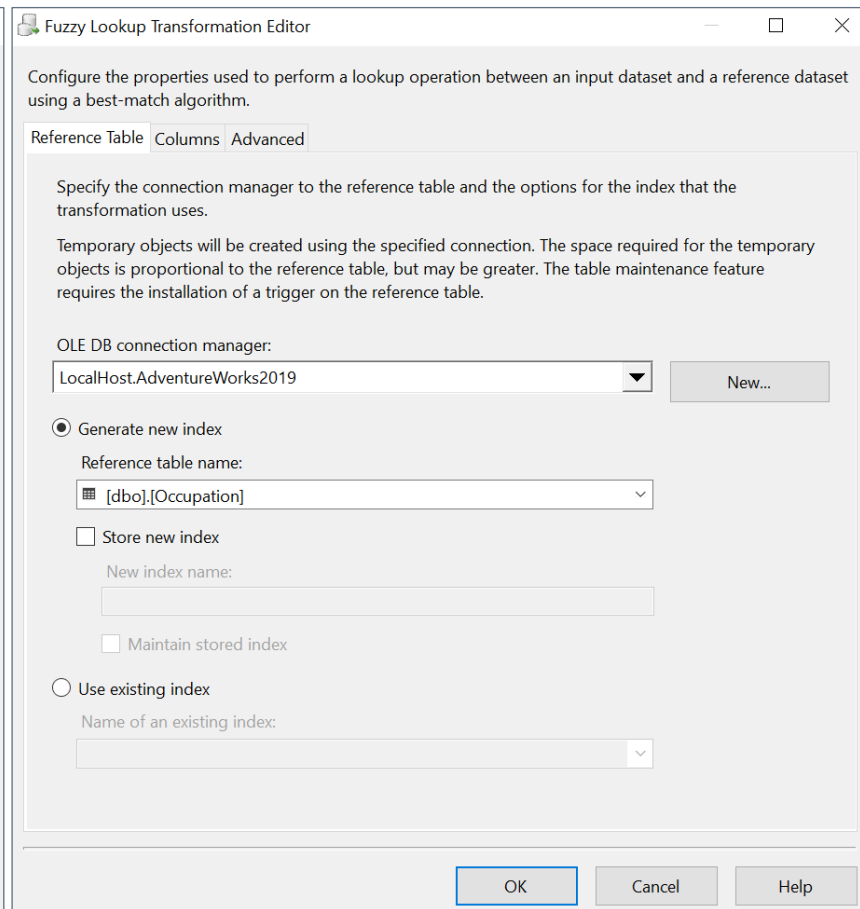
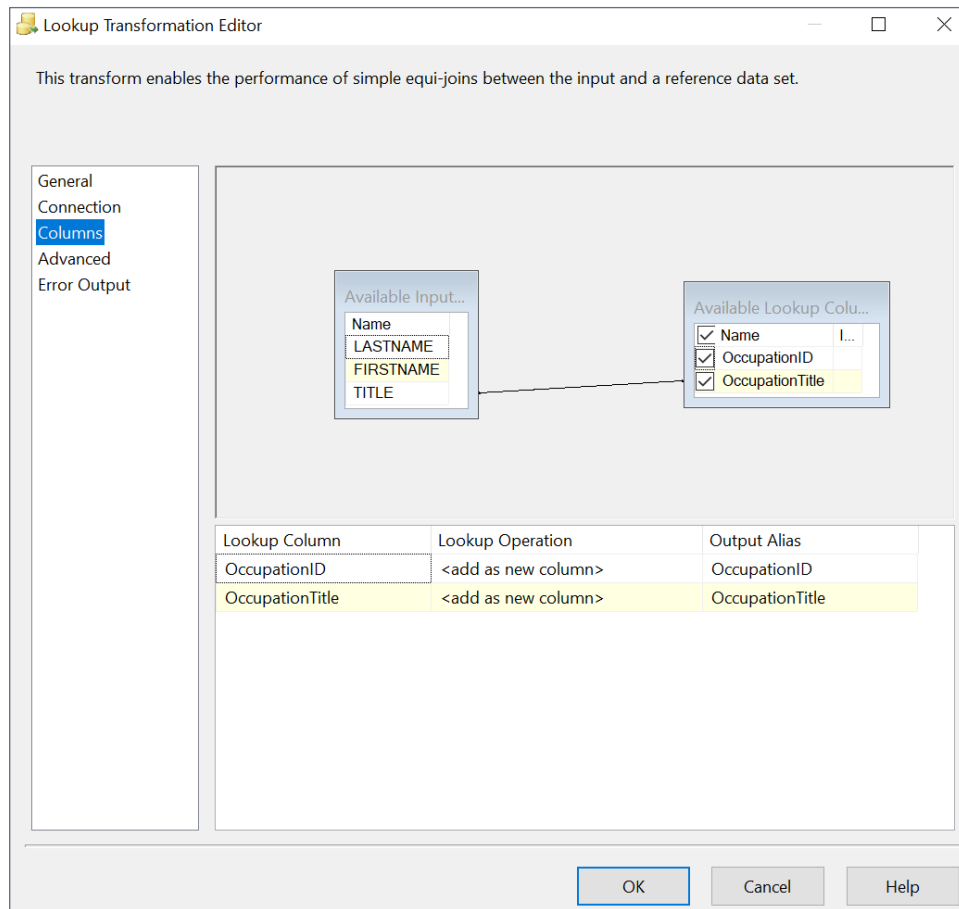
- Pour chaque employee du fichier **FuzzyExample**, on veut retourner aussi les details du de l'occupation depuis la table **Occupations**. Le résultat doit être chargé dans un fichier Excel **FuzzyLookupOutput**



18. Transformation SSIS: FuzzyLookup



18. Transformation SSIS: FuzzyLookup



18. Transformation SSIS: FuzzyLookup

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

Name	Pass T...
LASTNAME	<input checked="" type="checkbox"/>
FIRSTNAME	<input checked="" type="checkbox"/>
TITLE	<input checked="" type="checkbox"/>

Available Lookup Colu...

<input checked="" type="checkbox"/> Name
<input checked="" type="checkbox"/> OccupationID
<input checked="" type="checkbox"/> OccupationTitle

Lookup Column	Output Alias
OccupationID	OccupationID
OccupationTitle	OccupationTitle

OK Cancel Help

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

Maximum number of matches to output per lookup: 1

Similarity threshold: 0,50

Token delimiters

☒ Space ☐ Tab

☐ Carriage return ☐ Line feed

Additional delimiters:

OK Cancel Help

18. Transformation SSIS: FuzzyLookup

Union All Transformation Editor

Configure the properties used to merge multiple inputs into one output by creating mappings between columns.

Output Column Name	Union All Input 1	Union All Input 2
LASTNAME	LASTNAME	LASTNAME
FIRSTNAME	FIRSTNAME	FIRSTNAME
TITLE	TITLE	TITLE
OccupationID	OccupationID	OccupationID
OccupationTitle	OccupationTitle	OccupationTitle

OK Cancel Help

Flat File Connection Manager Editor

Connection manager name: Flat File Connection Manager 1

Description:

General Columns Advanced Preview

Select a file and specify the file properties and the file format.

File name: C:\Users\KhawlaElansari\Desktop\Enseigner Browse...

Locale: French (Canada) Unicode

Code page: 1252 (ANSI - Latin I)

Format: Delimited

Text qualifier: <none>

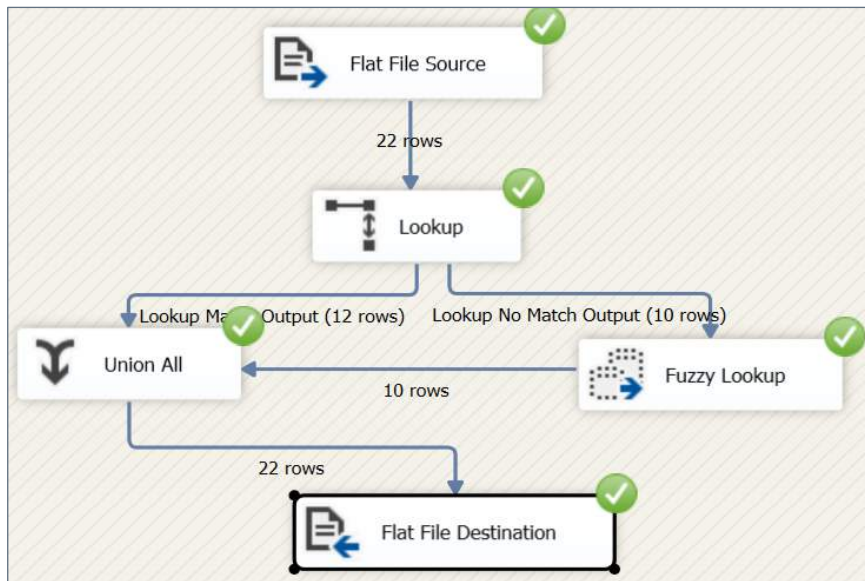
Header row delimiter: {CR}{LF}

Header rows to skip: 0

☐ Column names in the first data row

OK Cancel Help

18. Transformation SSIS: FuzzyLookup



LASTNAME	FIRSTNAME	TITLE	OccupationID	OccupationTitle
WALKER	JEREMY	CUSTOMER SERVICE REPRESENTATIVE	1	CUSTOMER SERVICE REPRESENTATIVE
OWENS	BENJAMIN	ASSISTANT MANAGER	3	ASSISTANT MANAGER
SCOTT	MICHAEL	REGIONAL MANAGER	6	REGIONAL MANAGER
SCHRUTE	DWIGHT	CUSTOMER SERVICE REPRESENTATIVE	1	CUSTOMER SERVICE REPRESENTATIVE
PAIGE	CARRINGTON	STORE MANAGER	4	STORE MANAGER
SCHACHT	BRADLEY	DISTRICT MANAGER	5	DISTRICT MANAGER
SHOWALTER	CHRISTOPHER	SHIFT LEADER	2	SHIFT LEADER
MCNAMARA	SEAN	REGIONAL MANAGER	6	REGIONAL MANAGER
COLBERT	STEPHEN	STORE MANAGER	4	STORE MANAGER
ISHEE	SAM	SHIFT LEADER	2	SHIFT LEADER
ENGLE	DANIEL	ASSISTANT MANAGER	3	ASSISTANT MANAGER
STEWART	JON	DISTRICT MANAGER	5	DISTRICT MANAGER
HENDERSON	HARRY	CUST SERVICE REP.	102	CUSTOMER SERVICE REPRESENTATIVE
ALBREKTSON	CHRISTOPHER	SHIFT LEAD	103	SHIFT LEADER
MCFLY	MARTY	REGION MANAGER	107	REGIONAL MANAGER
HALPERT	JIM	STORE MGR.	105	STORE MANAGER
FOSTER	DERRICK	CUSTOMER SERVICE REP	102	CUSTOMER SERVICE REPRESENTATIVE
BEESLY	PAM	ASSIST. MANAGER	104	ASSISTANT MANAGER
LANDEN	COLLIN	CUTSOMER SERVICE REPRESENTATIVE	102	CUSTOMER SERVICE REPRESENTATIVE
RYAN	DUSTIN	DISTRICT MGR.	106	DISTRICT MANAGER
TROY	CHRISTIAN	REGIONAL MGR	107	REGIONAL MANAGER
WOOD	JOHN	CUSTOMER SERVICE REP	102	CUSTOMER SERVICE REPRESENTATIVE



Questions ?