

SQL Server Integration Services (SSIS) Training Kit(Part

Lesson 6:Data Flow Tasks

Data Flow Task

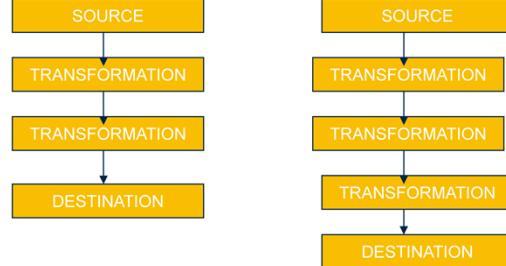
- The Data Flow task encapsulates the data flow engine that moves data between sources and destinations, providing the facility to transform, clean, and modify data as it is moved.
- Addition of a Data Flow task to a package control flow makes it possible for the package to extract, transform, and load data.
- A data flow consists of at least one data flow component, but it is typically a set of connected data flow components: sources that extract data; transformations that modify, route, or summarize data; and destinations that load data.
- Components are connected in the data flow by paths. Each path specifies the two components that are the start and the end of the path.



Copyright © Capgemini 2015. All Rights Reserved 2

Data Flow Task

- A Data Flow task can include multiple data flows. However, the data flow engine determines order of execution when there are multiple data flows within one data flow task. Therefore, when order is important, the package should use multiple Data Flow tasks, each task containing one data flow. You can then apply precedence constraints to control the execution order of the tasks.
- The following diagram shows a Data Flow task with multiple data flows.



Data Flow Task

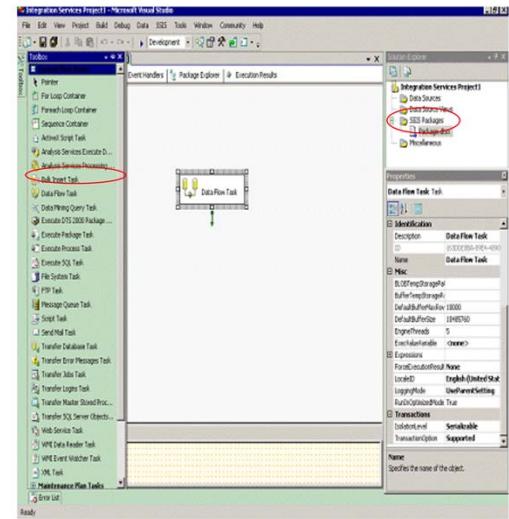
- The Data Flow task also manages error flows. At run time, row-level errors may occur when data flow components convert data, perform a lookup, or evaluate expressions. For example, a data column with a string value cannot be converted to an integer, or an expression tries to divide by zero. Both operations result in errors, and the rows containing errors can be processed separately using an error flow.



Copyright © Capgemini 2015. All Rights Reserved 4

Data Flow Task

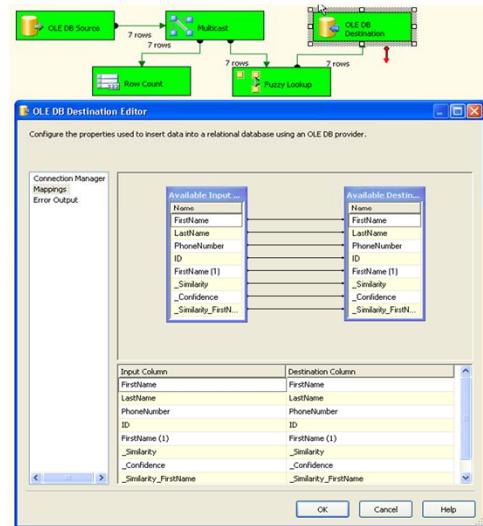
- To add a Data Flow task
- Click the Control Flow tab.
- In the Toolbox, expand Control Flow Items, and drag a Data Flow Task onto the design surface of the Control Flow tab.
- On the Control Flow design surface, right-click the newly added Data Flow Task, click Rename, and change the name to Extract Sample Currency Data.
- It is good practice to provide unique names to all components that you add to a design surface. For ease of use and maintainability, the names should describe the function that each component performs.



Copyright © Capgemini 2015. All Rights Reserved 5

Data Flow Task

- Data Flow Sources
 - Eg.Flat file source,OLEDB source
- Data Flow Transformations
 - Eg. Conditional Split ,Aggregate, Lookup
 - May have multiple inputs and outputs.
 - Must have at least one input
 - Must have at least one output
 - Process rows by buffer
 - May output more or less rows than take on input
- Data Flow Destinations
 - Eg.Flat file destinations,Pointer,OLEDB Destinations



Copyright © Capgemini 2015. All Rights Reserved 6

Data Flow Tasks- Source

Lesson 7:

Data Flow Task -Source

ADO NET Source

Excel Source

Flat File Source

OLE DB Source

Raw File Source

SAP BW Source

XML Source



Copyright © Capgemini 2015. All Rights Reserved 8

Data Flow Task –Source....

- ADO NET - The ADO NET source consumes data from a .NET provider and makes the data available to the data flow.
- Excel - The Excel source extracts data from worksheets or ranges in Microsoft Excel workbooks.
- Flat File - The Flat File source reads data from a text file. The text file can be in delimited, fixed width, or mixed format.
 - Delimited format uses column and row delimiters to define columns and rows.
 - Fixed width format uses width to define columns and rows. This format also includes a character for padding fields to their maximum width.
 - Ragged right format uses width to define all columns, except for the last column, which is delimited by the row delimiter.
- OLEDB - The OLE DB source extracts data from a variety of OLE DB-compliant relational databases by using a database table, a view, or an SQL command. For example, the OLE DB source can extract data from tables in Microsoft Office Access or SQL Server databases



Copyright © Capgemini 2015. All Rights Reserved 9

Data Flow Task –Source....

- Raw File - The Raw File source reads raw data from a file. Because the representation of the data is native to the source, the data requires no translation and almost no parsing. This means that the Raw File source can read data more quickly than other sources such as the Flat File and the OLE DB sources.
- SAP BW - The SAP BW source is the source component of the Microsoft Connector 1.1 for SAP BW. Thus, the SAP BW source extracts data from an SAP Netweaver BW version 7 system and makes this data available to the data flow in a Microsoft Integration Services package.
- XML - The XML source reads an XML data file and populates the columns in the source output with the data.
- The data in XML files frequently includes hierarchical relationships. For example, an XML data file can represent catalogs and items in catalogs. Before the data can enter the data flow, the relationship of the elements in XML data file must be determined, and an output must be generated for each element in the file.



Copyright © Capgemini 2015. All Rights Reserved 10

Data Flow Tasks- Transformation

Lesson 8:

Aggregate Transformation

- The Aggregate transformation applies aggregate functions, such as Average, to column values and copies the results to the transformation output. Besides aggregate functions, the transformation provides the GROUP BY clause, which you can use to specify groups to aggregate across.
- The Aggregate transformation is asynchronous, which means that it does not consume and publish data row by row. Instead it consumes the whole row set, performs its groupings and aggregations, and then publishes the results.
- The Aggregate transformation has one input and one or more outputs. It does not support an error output.



Copyright © Capgemini 2015. All Rights Reserved 12

Aggregate Transformation

- The Aggregate transformation supports the following operations :
- Group by :- Divides datasets into groups. Columns of any data type can be used for grouping.
- Sum :- Sums the values in a column. Only columns with numeric data types can be summed.
- Average :- Returns the average of the column values in a column. Only columns with numeric data types can be averaged.
- Count :- Returns the number of items in a group.
- Count distinct :- Returns the number of unique nonnull values in a group.



Copyright © Capgemini 2015. All Rights Reserved 13

Aggregate Transformation

- At the transformation level, you configure the Aggregate transformation for performance by specifying the number of keys and distinct count keys the transformation is expected to handle and the percentage by which memory can be extended during the aggregation. The Aggregate transformation can also be configured to generate a warning instead of failing when the value of a divisor is zero.
- At the output level, you configure the Aggregate transformation for performance by specifying the number of keys the output is expected to contain. The Aggregate transformation supports multiple outputs, and each can be configured differently.
- At the column level, you specify the aggregation that the column performs and the comparison options of the aggregation. You can also configure the Aggregate transformation for performance by specifying the number of keys and distinct count keys that each column contains, and identifying columns as Is Big if a column contains large numeric values or numeric values with high precision.



Copyright © Capgemini 2015. All Rights Reserved 14

Aggregate Transformation

Example::

- The following example is showing a data set comprising of 4 rows. Suppose now we have to calculate the sum of the salary all cities and We can achieve this by using Aggregate Transformation.

"Account No"	"First Name"	"Last Name"	"Company"	"Address"	"City"	"Salary"
"10019"	"bobbi"	"Arndt"	"Market Place"	"1000 S Nicolet..."	"Sametown"	1000
"10023"	"bruce"	"Beecher"	"Madson & Hut..."	"1037 W Wisco..."	"Smithville"	1500
"10024"	"bruce"	"Beyer"	"La Salle Clinic"	"108 E Wiscon..."	"Jonestown"	1800
"10025"	"butch"	"Bobbi"	"Town & Count..."	"108 Hilltop Ct"	"Smithville"	3000



Copyright © Capgemini 2015. All Rights Reserved 15

Aggregate Transformation

Output::

- The following is the output i.e total sum of the salary of all cities after Aggregate Transformation is done.

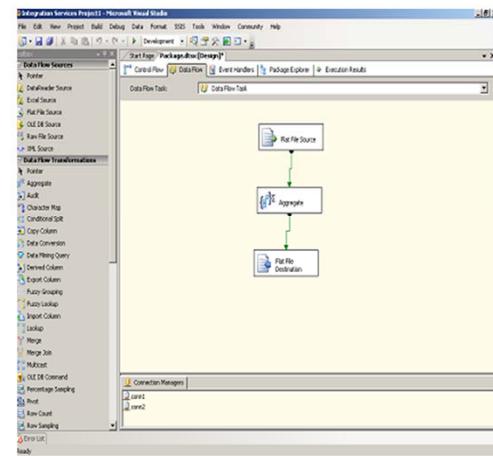
Account No	First Name	Last Name	Company	Address	City	sum of salary
						7300



Copyright © Capgemini 2015. All Rights Reserved 16

Aggregate Transformation

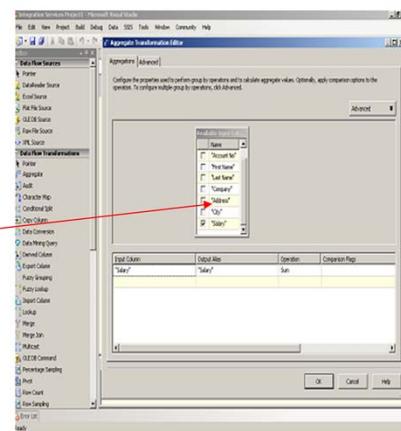
- Click the Data Flow tab, and then, from the Toolbox, drag the Aggregate transformation to the design surface
- Connect the Aggregate transformation to the data flow by dragging a connector—the green or red arrow—from the source or the previous transformation to the Aggregate transformation.



Copyright © Capgemini 2015. All Rights Reserved 17

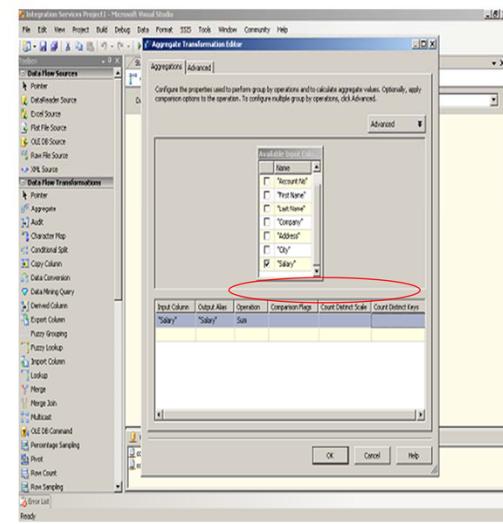
Aggregate Transformation

- Double-click the transformation.
- In the Editor dialog box,
 - click the Aggregations tab.
- In the Available Input Columns list, select the 'Salary' check box by the columns on which we want to aggregate values. The selected columns appear in the table.
- Optionally, you can modify the value in the Output Alias columns.
- Default aggregation operation is Group by. choose the Sum operation. From the drop down
 - Operation list.



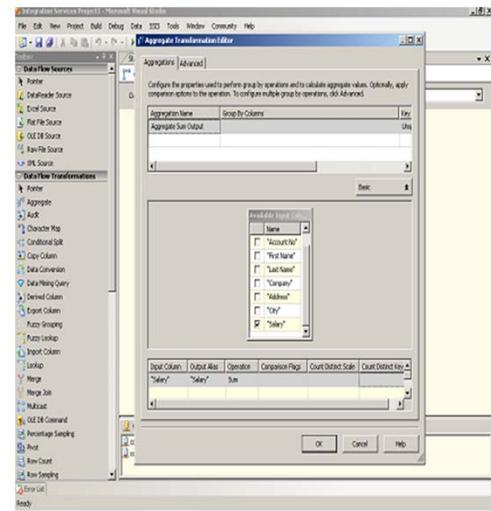
Aggregate Transformation

- To change the default comparison, select the comparison flags to use when defining the groups to aggregate in the Comparison Flags column.
- Optionally, for the Count distinct aggregation, specify an exact count of distinct values in the Count Distinct Keys column, or select an approximate count in the Count Distinct Scale column.



Aggregate Transformation

- Optionally, click Advanced and update the name of the Aggregate transformation output. If the aggregations include a group by operation, you can select an approximate count of grouping key values in the Keys Scale column, or specify an exact number of grouping key values in the Keys column.
- Optionally, click the **Advanced** tab and set the attributes that apply to all the operations that the Aggregate transformation performs.



Copyright © Capgemini 2015. All Rights Reserved 20

Aggregate Transformation

- The Aggregate transformation includes a set of properties that you can set to enhance the performance of the transformation.
- Set the Keys and Keys Scale properties of the component and the component outputs. Using Keys, you can specify the exact number of keys the transformation is expected to handle, and using Keys Scale, you can specify an approximate number of keys. When you specify a value for Keys, which is the value the transformation will receive when the package runs, the transformation avoids reorganizing cached totals, improving performance.
- Set the Count Distinct Keys and Count Distinct Scale properties of the component. Using Count Distinct Keys, you can specify the exact number of keys the transformation is expected to handle for a count distinct operation. Using Count Distinct Scale, you can specify an approximate number of keys for a count distinct operation. When you specify a value for Count Distinct Scale, which is the value the transformation will receive when the package runs, the transformation also avoids reorganizing cached totals, improving performance.



Copyright © Capgemini 2015. All Rights Reserved 21

Audit Transformation

Lesson 9:

Audit Transformation

- “The Audit transformation enables the data flow in a package to include data about the environment in which the package runs. For example, the name of the package, computer, and operator can be added to the data flow. Microsoft SQL Server 2008 Integration Services (SSIS) includes system variables that provide this information”.
- “You configure the Audit transformation by providing the name of a new output column to add to the transformation output, and then mapping the system variable to the output column. You can map a single system variable to multiple columns”.
- “This transformation has one input and one output. It does not support an “error output””.



Copyright © Capgemini 2015. All Rights Reserved 23

Audit Transformation

- EXAMPLE: The audit Fields are used to describe general information such as, Package Name, Machine Name, User Name, Package Id, Version Id etc.
- The output shown here includes few of the audit fields after Audit Field Transformation.

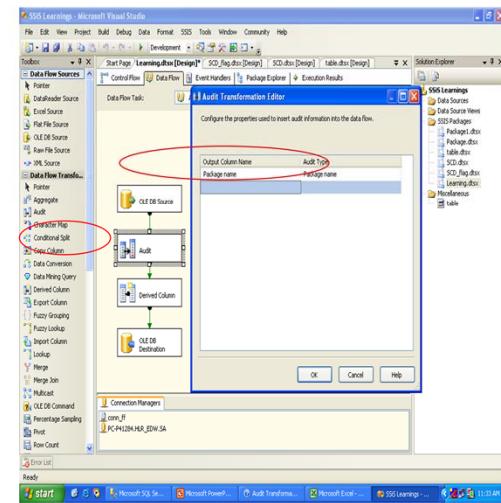
Employee Id	First Name	Package name	Execution insta...	Machine name	User name
10019	Bobbi	Package1	{74ED1227-94...	PC-P41213	PATNI\goretush
10023	Bruce	Package1	{74ED1227-94...	PC-P41213	PATNI\goretush
10024	Bruce	Package1	{74ED1227-94...	PC-P41213	PATNI\goretush
10025	Butch	Package1	{74ED1227-94...	PC-P41213	PATNI\goretush
10026	Calla	Package1	{74ED1227-94...	PC-P41213	PATNI\goretush
10027	Carol	Package1	{74ED1227-94...	PC-P41213	PATNI\goretush
10028	Carol	Package1	{74ED1227-94...	PC-P41213	PATNI\goretush
10029	Cheri	Package1	{74ED1227-94...	PC-P41213	PATNI\goretush
10030	Chuck	Package1	{74ED1227-94...	PC-P41213	PATNI\goretush
10031	Chuck	Package1	{74ED1227-94...	PC-P41213	PATNI\goretush
10032	Chuck	Package1	{74ED1227-94...	PC-P41213	PATNI\goretush
10034	Colleen	Package1	{74ED1227-94...	PC-P41213	PATNI\goretush
10035	Connie	Package1	{74ED1227-94...	PC-P41213	PATNI\goretush
10036	Connie	Package1	{74ED1227-94...	PC-P41213	PATNI\goretush



Copyright © Capgemini 2015. All Rights Reserved 24

Audit Transformation

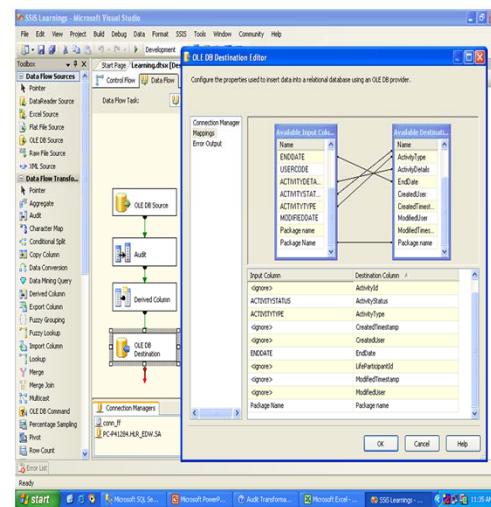
- Drag a 'OLE DB Source' from Data Flow Task Navigator to Data Flow Task Window.
- Set the 'Connection Manager' by giving appropriate information, e.g. Connection Manager Name, Table or Query, Table Name.
- Drag the Audit transformation from Data Flow Transformation Navigation Toolbox.
- Select the output column name and audit field from and click 'OK' eg. Package Name.



Copyright © Capgemini 2015. All Rights Reserved 25

Audit Transformation

- Drag the 'OLE DB Destination' select the proper connection and map the columns properly.
- The Audit column – package name will also be mapped.
- Derived Column transformation is used between Audit and OL DB destination transformation to cast the Package Name, Data type – Unicode String(DT_WSTR) to Data type – Non-Unicode, as mentioned in the target table structure.



Copyright © Capgemini 2015. All Rights Reserved 26

Audit Transformation

- SYSTEM VARIABLES:
- ExecutionInstanceGUID:
 - The GUID that identifies the execution instance of the package.
- PackageID:
 - The unique identifier of the package.
- PackageName:
 - The package name.
- VersionID:
 - The version of the package.
- ExecutionStartTime:
 - The time the package started to run.



Copyright © Capgemini 2015. All Rights Reserved 27

Audit Transformation

- **MachineName:**
The computer name.
- **UserName**
The login name of the person who started the package.
- **TaskName**
The name of the Data Flow task with which the Audit transformation is associated.
- **TaskId**
The unique identifier of the Data Flow task.



Copyright © Capgemini 2015. All Rights Reserved 28

Cache Transformation

Lesson 10:

Cache Transformation

- “The Cache transformation is a brand new feature in SQL 2008, allowing you to cache the data used in the Lookup transform. The Lookup transform can then utilize this cached data to perform the lookup operation ”
- We can configure the component to cache the lookup dataset, rather than retrieving the data on a per row basis. In SSIS 2008, your caching options for performing lookup operations have been extended through the Cache transformation and Cache connection manager .



Copyright © Capgemini 2015. All Rights Reserved 30

Cache Transformation

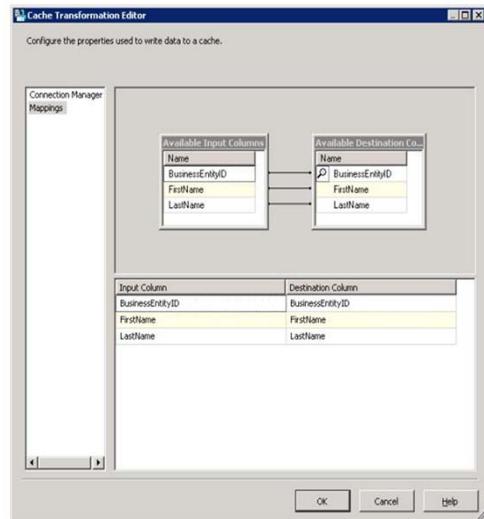
- By using the new transformation and connection manager, you can cache lookup data from any type of data source (not only an OLE DB source), persist the data to the memory cache or into a cache file on your hard drive, and use the data in multiple data flows or packages
- The primary purpose of the Cache transformation is to persist data through a Cache connection manager. When configuring the transformation, you must, in addition to specifying the connection manager, define the column mappings.



Copyright © Capgemini 2015. All Rights Reserved 31

Cache Transformation

- The following figure shows the Mappings page of the Cache Transformation Editor.

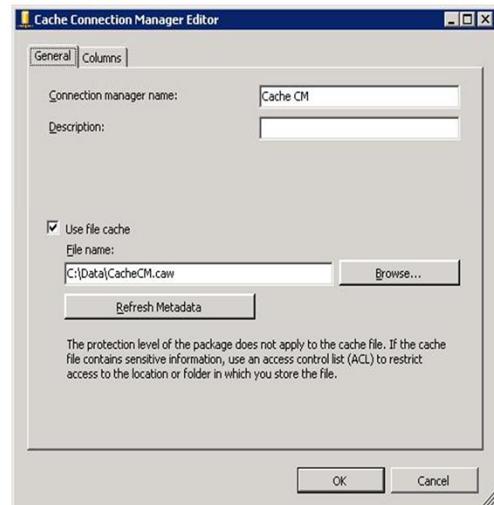


Copyright © Capgemini 2015. All Rights Reserved 32

Cache Transformation

Cache connection manager

- To support cached lookups, you must, along with configuring the Cache transformation, configure the Cache connection manager. The following figure shows the General tab of the Connection Manager Editor.



Cache Transformation

Best practices

- Reuse the cache to reduce database load
- Share the cache between lookups to reduce memory usage
- Using the CCM is not always faster than OLEDB - the cost of disk access can out weight the benefits of pre-creating the cache
- The cache is essentially clear text - do not store sensitive data inside of the cache
- In terms of Cache Modes(Full,Partial,No Cache) and the best practices that surround them, using a cache connection manager is equivalent to using a Full Cache mode

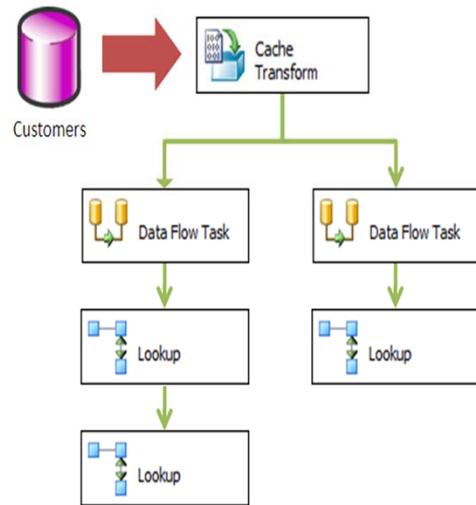


Copyright © Capgemini 2015. All Rights Reserved 34

Cache Transformation

Reducing database and memory usage

- If your reference database is remote, or under heavy load, consider using the Cache Connection Manager instead of an OLEDB connection.
- Once a cache is used (or created) in an SSIS package, it will be kept in memory until the package has finished executing. The cache can be reused across multiple data flows, and shared between multiple lookups in the same data flow. It can also be persisted to disk, and reused across package executions.



Character Map Transformation

Lesson 11:

Character Map Transformation

- “The Character Map transformation applies string functions, such as conversion from lowercase to uppercase, to character data. This transformation operates only on column data with a string data type.”
- This is a passive transformation.
- You configure the Character Map transformation in the following ways
 - Specify the columns to convert.
 - Specify the operations to apply to each column.
- The Character Map transformation can convert column data in place or add a column to the transformation output and put the converted data in the new column. You can apply different sets of mapping operations to the same input column and put the results in different columns.



Copyright © Capgemini 2015. All Rights Reserved 37

Character Map Transformation

- The following table describes the mapping operations that the Character Map transformation supports.

Operation	Description
Byte reversal	Reverses byte order.
Full width	Maps half-width characters to full-width characters.
Half width	Maps full-width characters to half-width characters.
Hiragana	Maps katakana characters to hiragana characters.
Katakana	Maps hiragana characters to katakana characters.
Linguistic casing	Applies linguistic casing instead of the system rules. It refers to functionality provided by the Win32 API for Unicode simple case mapping of Turkic and other locales.
Lowercase	Converts characters to lowercase.
Simplified Chinese	Maps traditional Chinese characters to simplified Chinese characters.
Traditional Chinese	Maps simplified Chinese characters to traditional Chinese characters.
Uppercase	Converts characters to uppercase.



Copyright © Capgemini 2015. All Rights Reserved 38

Character Map Transformation

- More than one operation can be performed in a transformation. However, some mapping operations are mutually exclusive. The following table lists restrictions that apply when you use multiple operations on the same column. Operations in the columns Operation A and Operation B are mutually exclusive.

Operation A	Operation B
Lowercase	Uppercase
Hiragana	Katakana
Half width	Full width
Traditional Chinese	Simplified Chinese
Lowercase	Hiragana, katakana, half width, full width
Uppercase	Hiragana, katakana, half width, full width

- Use the Character Map Transformation Editor dialog box to select the string functions to apply to column data and to specify whether mapping is an in-place change or added as a new column.



Copyright © Capgemini 2015. All Rights Reserved 39

Character Map Transformation

Example:

- The following example is showing a data set comprising of 10 rows. Suppose now requirement is that all characters of the column 'First Name' should be in Uppercase. For that we have to use Character Map Transformation to convert the characters to uppercase.

Account No	First Name	Last Name	Company	Address	City
"10019"	"bobbi"	"Andt"	"Market Place"	"1000 S Nicolet..."	"Sametown","T...
"10023"	"bruce"	"Beecher"	"Madson & Hutz..."	"1037 W Wisco..."	"Smithville","A...
"10024"	"bruce"	"Beyer"	"La Salle Clinic"	"108 E Wiscon..."	"Jonestown","..."
"10025"	"butch"	"Bobbi"	"Town & Count..."	"108 Hilltop Ct"	"Smithville","A...
"10026"	"calla"	"Boshers"	"Saturn of App..."	"110 Fox River..."	"Smithville","A...
"10027"	"carol"	"Brauer"	"Bemiss Corp."	"110 W North ..."	"Jonestown","..."
"10028"	"carol"	"Braun"	"AAL Member ..."	"1115 E Glend..."	"Smithville","A...
"10029"	"cheri"	"Buksyk"	"Office Support"	"1122 Milwauk..."	"Jonestown","..."
"10030"	"chuck"	"Buss"	"EAA"	"1134 S Franklin..."	"Overton","AK..."
"10031"	"chuck"	"Carpenter"	"Kurz Electric"	"115 S Drew St"	"Smithville","A...



Character Map Transformation

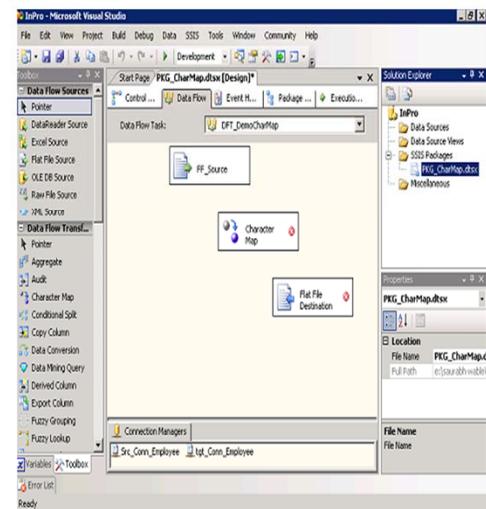
- All the characters of the column 'First Name' is converted from lowercase
- Output::
to uppercase in the O/P data set after Character Map Transformation is done.

Account No	First Name	Last Name	Company	Address	City
"10019"	"BOBBI"	"Arndt"	"Market Place"	"1000 S Nicolet..."	"Sametown", "T...
"10023"	"BRUCE"	"Beecher"	"Madson & Hut..."	"1037 W Wisco..."	"Smithville", "A...
"10024"	"BRUCE"	"Beyer"	"La Salle Clinic"	"108 E Wiscon..."	"Jonestown", "...
"10025"	"BUTCH"	"Bobbi"	"Town & Count..."	"108 Hilltop Ct"	"Smithville", "A...
"10026"	"CALLA"	"Boschers"	"Saturn of App..."	"110 Fox River..."	"Smithville", "A...
"10027"	"CAROL"	"Brauer"	"Bemiss Corp."	"110 W North ..."	"Jonestown", "...
"10028"	"CAROL"	"Braun"	"AAL Member ..."	"1115 E Glend..."	"Smithville", "A...
"10029"	"CHERI"	"Bulsky"	"Office Support"	"1122 Milwaukee..."	"Jonestown", "...
"10030"	"CHUCK"	"Bus"	"EAA"	"1134 S Franklin..."	"Overton", "AK..."
"10031"	"CHUCK"	"Carpenter"	"Kurz Electric"	"115 S Drew St"	"Smithville", "A...



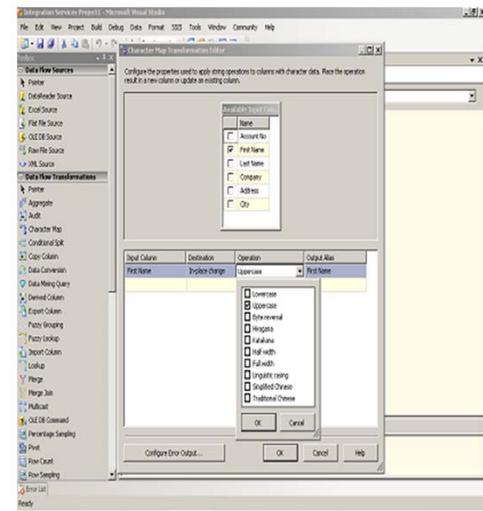
Character Map Transformation

- Drag flat file Source, character map Transformation and flat file Destination as shown
- Configure the FF_Source with Src_Conn_Employee connection.
- Link output of FF_Source to character map transformation.



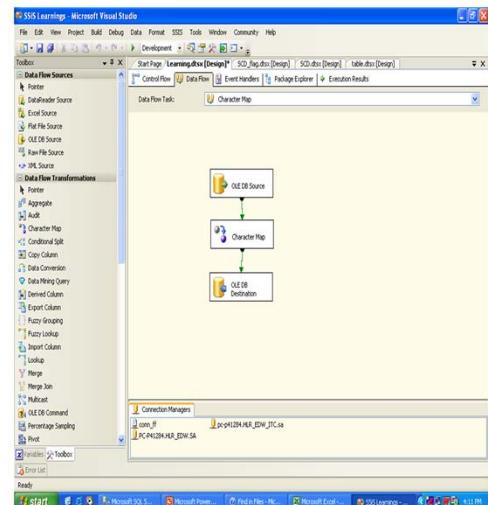
Character Map Transformation

- Double click to open transformation editor as shown.
- Select columns, selected column are automatically added below
- Select the operation you want to be performed.
- Input Column: for selecting column.
- Destination : either a new column or In place change.
- Operation: you can specify operation/s on column.
- Output Alias: for supplying either new column name or
- Alias name to input column.
- Other column redirected as it is.



Character Map Transformation

- Click ok, and connect the output of transformation to Flat File Destination.
- Double Click to configure the Flat file destination.
- Add the connection.
- Map the columns of source and destination.
- Run the data flow and check the output



Copyright © Capgemini 2015. All Rights Reserved 44

Conditional Split Transformation

Lesson 12:

Conditional Split Transformation

- The Conditional Split transformation evaluates expressions, and based on the results, directs the data row to the specified output.
- This transformation also provides a default output, so that if a row matches no expression it is directed to the default output.
- Provide an expression that evaluates to a Boolean for each condition you want the transformation to test.



Copyright © Capgemini 2015. All Rights Reserved 46

Traditional tabulated presentation v/s graphical representation.
Tables don't determine trends quickly as line charts do. Other

Conditional Split Transformation

- Specify the order in which the conditions are evaluated. Order is significant, because a row is sent to the output corresponding to the first condition that evaluates to true.
- Specify the default output for the transformation. The transformation requires that a default output be specified.
- EXAMPLE

Suppose in our scenario we want the employees working for USA or for IND. Here I will put condition as

case 1 CITY==“USA” And

case 2 CITY==“IND”

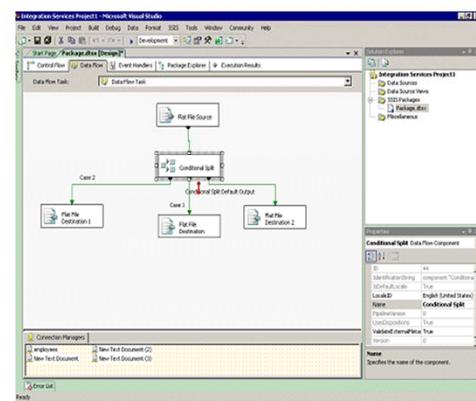
to split the records into two groups. One default group is also created automatically to store unmatched, like CITY other then USA or IND.



Copyright © Capgemini 2015. All Rights Reserved 47

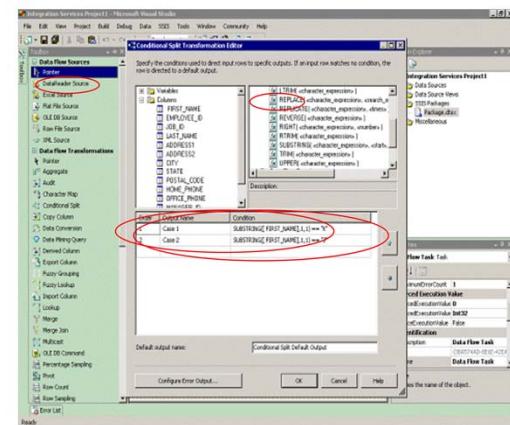
Conditional Split Transformation

- Drag a 'Conditional split transformation' from Data Flow Transformations & drop it in to work area & connect it with the flat file source.
- Double Click or Right Click on Data Flow transformation and select edit to open 'Conditional Split Transformation Editor' window



Conditional Split Transformation

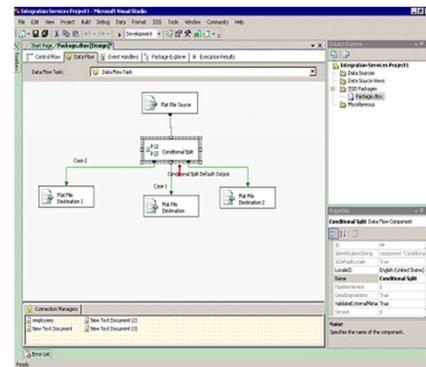
- For example, the following conditions direct any rows in the FirstName column that begin with the letter "K" to one output, rows that begin with the letter "J" to a different output, and all other rows to the default output



Copyright © Capgemini 2015. All Rights Reserved 49

Conditional Split Transformation

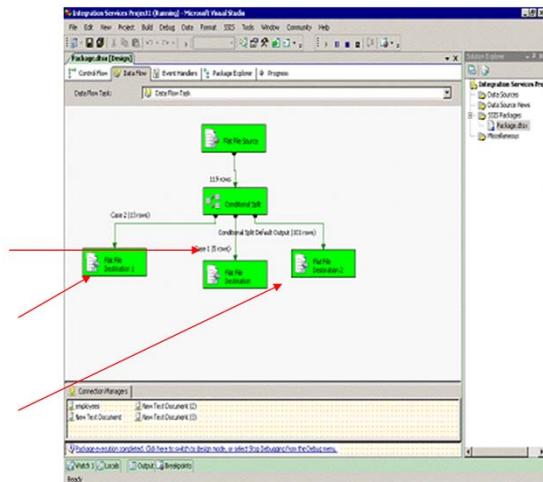
- Depending upon the conditions set the conditional split transformation splits the data into 3 different flat files.
- Three Flat file destinations are connected with the conditional split & corresponding connections for the flat files are done in the connection manager.



Copyright © Capgemini 2015. All Rights Reserved 50

Conditional Split Transformation

- Start the debugging process and carefully notice result. It is seen that conditional split transformation splits the data depending upon the conditions following conditions.
- Case1: First Name column that begin with the letter "K".(5 rows)
- Case2: First Name column that begin with the letter "J".(13 rows)
- Conditional split default output is the o/p having the remaining data.



Conditional Split Transformation

- Input Source

EMPLOYEE		
First_Name	Last_Name	City_Code
Kashav	ganguly	452001
sachin	tendulkar	400045
vaibhav	dravid	300016
Jai	prakash	400456
Jai	chand	700342
Kishor	kumar	676004

- Case 1 Output

Kashav	ganguly	452001
Kishor	kumar	676004

- Case 2 Output

Jai	prakash	400456
Jai	chand	700342

- Default Output

sachin	tendulkar	400045
vaibhav	dravid	300016



Copyright © Capgemini 2015. All Rights Reserved 52

Copy Column Transformation

Lesson 13:

Copy Column Transformation

- The copy column transformation creates new columns by copying input columns and adding the new columns to the transformation output. Later in the data flow, different transformations can be applied to the column copies.
- For example, you can use the copy column transformation to create a copy of a column and then convert the copied data to uppercase characters by using the character map transformation, or apply aggregations to the new column by using the aggregate transformation.



Copyright © Capgemini 2015. All Rights Reserved 54

Copy Column Transformation

Example::

- The following example is showing a data set comprising of 6 columns. Say now requirement is that we have to create a copy of a column and then convert the copied data to uppercase characters by using the Character Map Transformation. We can achieve this by using Copy Column Transformation.

Account No	First Name	Last Name	Company	Address	City
"10019"	"bobbi"	"Arndt"	"Market Place"	"1000 S Nicolet..."	"Sametown"
"10023"	"bruce"	"Beecher"	"Madson & Hut..."	"1037 W Wisco..."	"Smithville"
"10024"	"bruce"	"Beyer"	"La Salle Clinic"	"108 E Wiscon..."	"Jonestown"
"10025"	"butch"	"Bobbi"	"Town & Count..."	"108 Hilltop Ct"	"Smithville"



Copyright © Capgemini 2015. All Rights Reserved 55

Copy Column Transformation

- Output::

- Output Data set will be as follows comprising of 7 columns which includes newly added company alias column which is nothing but a copy of company column.

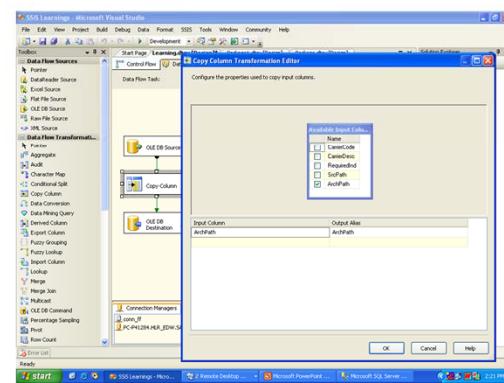
Account No	First Name	Last Name	Company	Address	City	Company alias
"10019"	"bobbi"	"Arndt"	"Market Place"	"1000 S Nicolet..."	"Sametown"	"Market Place"
"10023"	"bruce"	"Beecher"	"Madson & Huth Communication Co"	"1037 W Wisco..."	"Smithville"	"Madson & Huth Communication Co"
"10024"	"bruce"	"Beyer"	"La Salle Clinic"	"108 E Wiscon..."	"Jonestown"	"La Salle Clinic"
"10025"	"butch"	"Bobbi"	"Town & Country Electric Inc."	"108 Hillock Ct"	"Smithville"	"Town & Country Electric Inc."



Copyright © Capgemini 2015. All Rights Reserved 56

Copy Column Transformation

- Drag Copy Column Transformation from Data Flow Transformations.
- Link it with source.
- Double click on Copy Column Transformation to open copy column Transformation Editor.
- The column to be copied here is ArchPath. The column will be accordingly mapped in the target table.



Copy Column Transformation

Options Available:

- Available Input Columns
- Select columns to copy by using the check boxes. Your selections add input columns to the table below.
- Input Column
- Select columns to copy from the list of available input columns. Your selections are reflected in the check box selections in the Available Input Columns table.
- Output Alias
- Type an alias for each new output column. The default is Copy of, followed by the name of the input column; however, you can choose any unique, descriptive name.



Copyright © Capgemini 2015. All Rights Reserved 58

Data Conversion Transformation

Lesson 14:

Data Conversion Transformation

- “The Data Conversion transformation converts the data in an input column to a different data type and then copies it to a new output column.”
- For example, a package can extract data from multiple sources, and then use this transformation to convert columns to the data type required by the destination data store. You can apply multiple conversions to a single input column.
- You can configure the Data Conversion transformation in the following ways:
- Specify the columns that contain the data to convert and the types of data conversions to perform.
- Specify whether the transformation output columns use the quicker, but locale-insensitive, fast parsing routines that Microsoft SQL Server 2008 Integration Services (SSIS) provides or the standard locale-sensitive parsing routines.



Copyright © Capgemini 2015. All Rights Reserved 60

Data Conversion Transformation

- Example: The Data before Data Conversion is shown below. The column named Expenses contains decimal values.
- We need to have integer values in Expenses Column. In this case we choose Data Conversion Transformation for converting decimal values into integer values.

Employee ID	Expenses	Description	Account No
44894	001.11	INTL	111
44895	0011.9	ALBANY	119
44896	0018.8	ALBUQUERQUE	120
44897	5.6626	HONOLULU	126
44898	005.27	HILO	127
44899	07.128	KONA	128
44900	89.129	MAUI	129
44901	080.31	IMC	131
44902	7684.8	ANCHORAGE	140
44903	7645.5	APPLETON	150
44904	077.53	CHERRY HILL	153



Copyright © Capgemini 2015. All Rights Reserved 61

Data Conversion Transformation

- The Data after conversion will be as shown. All the decimal values in the Expenses column are now converted into Integer values.

Employee ID	Rounded Off Expenses	Description	Account No
44894	1	INTL	111
44895	12	ALBANY	119
44896	19	ALBUQUERQUE	120
44897	6	HONOLULU	126
44898	5	HILO	127
44899	7	KONA	128
44900	89	MAUI	129
44901	80	IMC	131
44902	7685	ANCHORAGE	140
44903	7646	APPLETON	150
44904	78	CHERRY HILL	153



Copyright © Capgemini 2015. All Rights Reserved 62

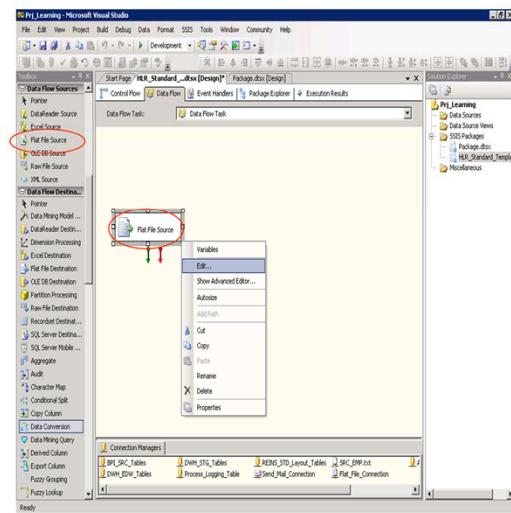
Data Conversion Transformation

- Drag a 'Flat File Source' from Data Flow Source.

Data in Flat File

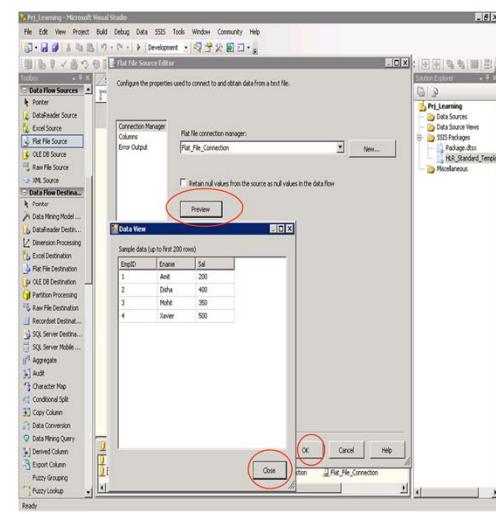
1	Amit	200
2	Disha	400
3	Mohit	350
4	Xavier	500

- Double click on 'Flat File Source' or Right Click on 'Flat File Source' and select 'Edit'
- Make appropriate change in Flat File.(e.g. Choose Flat File Connection, Select a File, Rename Column Name...etc.)



Data Conversion Transformation

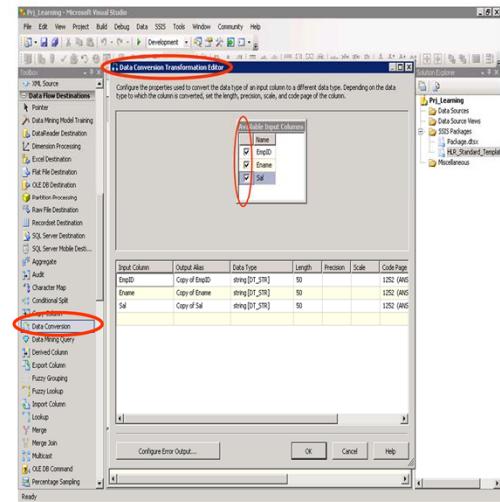
- On Clicking Preview button, Data View window will be open.
- Click 'Close and Then 'OK' button of Flat File Source Editor.



Copyright © Capgemini 2015. All Rights Reserved 64

Data Conversion Transformation

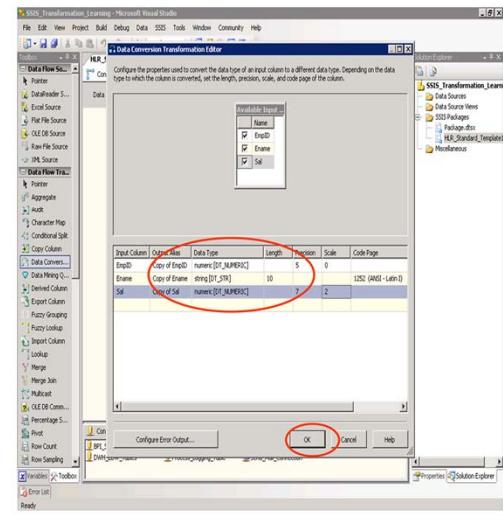
- Drag a ‘Data Conversion’ transformation from Data Flow Destinations Navigation & drop it to work area.
- Double Click or Right Click on Data Flow transformation and select edit to open ‘Data Conversion Transformation Editor’ window.
- Select required columns from ‘Available Input Columns’ .



Data Conversion Transformation

- Change 'Data Type', 'Length',....of required columns.

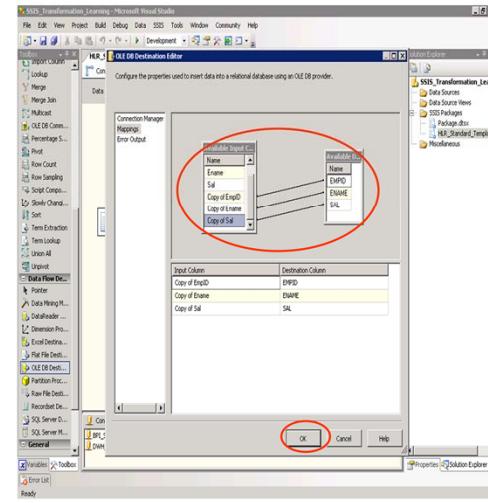
- Click 'OK'



Copyright © Capgemini 2015. All Rights Reserved 66

Data Conversion Transformation

- Drag a ‘OLE DB Destination’ from Data Flow Destination and drop it in work area, and make a connection between ‘Data Conversion Transformation’ and ‘OLE DB Destination’.
- Edit ‘OLE DB Destination’ and make proper connection in ‘Connection Manager’ (e.g. Set Target Database, Target Table.).
- In ‘Mappings’ tab make proper column mapping between ‘Available Input Columns’ and ‘Available Output Columns’.
- Click ‘OK’.



Copyright © Capgemini 2015. All Rights Reserved 67

Derived Column Transformation

Lesson 15:

Derived Column Transformation

“The Derived Column Transformation creates new columns values by applying expressions to transformation input columns.”

The Derived Column can perform following tasks :

- Concatenate data from different column into derived column.
- Extract characters from string data by using functions like SUBSTRING, and then store result in a derived column.
- Apply mathematical functions to numeric data and store the result in the derived column.
- Create expressions that compare input columns and variables.
- Extract part of date time value.



Copyright © Capgemini 2015. All Rights Reserved 69

Derived Column Transformation

- EXAMPLE: The data shown below is in fixed width format. Columns 'First Name' & 'Last Name' both are of length 10.
- Here, we can trim those fields in order to avoid blank spaces to be inserted into target.
- Also we can insert some value incase the input coming is NULL.

Account No	First Name	Last Name	Company
10019	Bobbi	Arndt	Market Place
10023	Bruce	Beecher	Madson & Huth Communication Co
10024	Bruce	Beyer	La Salle Clinic
10025	Butch	Bobbi	Town & Country Electric Inc.
10026	Calla	Boshers	Saturn of Appleton
10027	Carol	Brauer	Benniss Corp.
10028	Carol	Braun	AAL Member Credit Union
10029	Cheri	Bulskyk	Office Support
10030	Chuck	Buss	EAA
10031	Chuck	Carpenter	Kurz Electric
10032	Chuck	Carr	Alpha 1
10034	Colleen	Casperson	Valley Trust Corporation
10035	Connie	Catterton	Fox Community Credit Union
10036	Connie	Clay	Valley Lawn Care



Copyright © Capgemini 2015. All Rights Reserved 70

Derived Column Transformation

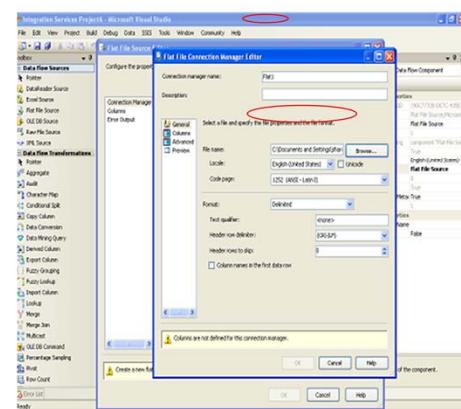
- Modified Date is current system date & Modified User is current system user.
- The following output contains two newly Derived columns called 'Modified Date' & 'Modified User'.
- Fields 'First Name' & 'Last Name' are trimmed where unnecessary spaces are removed.

Account No	First Name	Last Name	Company	ModifiedDate	Modified User
10019	Bobbi	Arnold	Market Place	1/8/2007 3:04:03 PM	PATNT\goretush
10023	Bruce	Beecher	Madison & Huth Com...	1/8/2007 3:04:03 PM	PATNT\goretush
10024	Bruce	Beyer	La Salle Clinic	1/8/2007 3:04:03 PM	PATNT\goretush
10025	Butch	Bobbi	Town & Country Elec...	1/8/2007 3:04:03 PM	PATNT\goretush
10026	Calla	Bosher	Saturn of Appleton ...	1/8/2007 3:04:03 PM	PATNT\goretush
10027	Carol	Brauer	Bemiss Corp.	1/8/2007 3:04:03 PM	PATNT\goretush
10028	Carol	Braun	AAL Member Credit U...	1/8/2007 3:04:03 PM	PATNT\goretush
10029	Cheri	Bukysk	Office Support	1/8/2007 3:04:03 PM	PATNT\goretush
10030	Chuck	Buss	EAA	1/8/2007 3:04:03 PM	PATNT\goretush
10031	Chuck	Carpenter	Kurz Electric	1/8/2007 3:04:03 PM	PATNT\goretush
10032	Chuck	Carr	Alpha 1	1/8/2007 3:04:03 PM	PATNT\goretush
10034	Colleen	Casperson	Valley Trust Corporat...	1/8/2007 3:04:03 PM	PATNT\goretush
10035	Connie	Catterton	Fox Community Credi...	1/8/2007 3:04:03 PM	PATNT\goretush



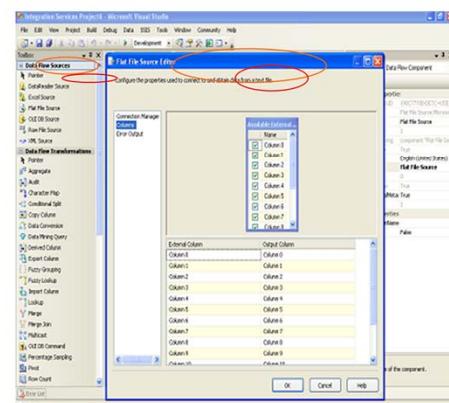
Derived Column Transformation

- Drag a 'Flat File Source' from the Data Flow Task Navigator into Data Flow Task Window.
- Double click or Right click on 'Flat File Source' to open 'Flat File Source Editor'
- Set the 'Connection Manager' by giving appropriate information, e.g. Connection Manager name, File Name.



Derived Column Transformation

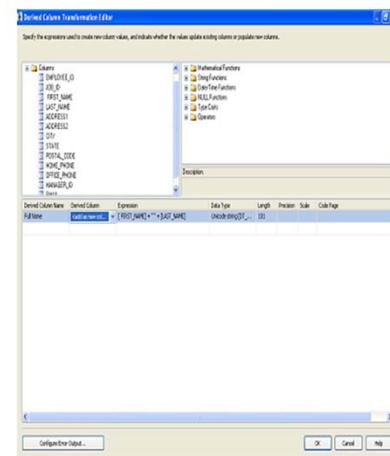
- Select Columns and check required 'Available External Column'.
- Click 'OK'
- Drag 'Derived Column Transformation' from Data Flow Transformation into Data Flow Task Window and connect Flat File Source to Derived Column Transformation.



Copyright © Capgemini 2015. All Rights Reserved 73

Derived Column Transformation

- Double Click or Right Click on the Derived Column Transformation to open the 'Derived Column Transformation Editor'.
- Give Derived Column Name, in Derived Column select 'add as new column', in Expression add the expression which is to be derived, for example we want to concatenate First Name and Last Name of the Employee Table and store it under Separate column 'FULL NAME'.
- Click 'OK'



Derived Column Transformation

- Drag 'Flat File Destination' from the Data Flow Destinations into Data Flow Task Window and connect Derived Column output to the Flat File Destination.
- Set the Connection Manager as explained previously.
- Create the required mapping.
- Click 'OK'

