# SQL SERVER INTEGRATION SERVICES (SSIS)

SSIS is an ETL tool which is used to collecting and cleaning the data from heterogenous / homogenous sources.

SSIS is used for:

1. Importing and exporting the data.
2. Integrating heterogenous data.
3. Clean and standardize the data.
4. Support BI solutions.

Five tabs in SSIS designer:

1. Control Flow.
2. Data Flow.
3. Event Handlers.
4. Package Explorer.
5. Progress / Execution results.

SSIS wizards:

1. SQL Server import export wizards.
2. Package Configuration Wizard (to make packages dynamic).
3. Package Installation Wizard.

# FILE EXTENSION TYPES

Data Transformation Service files are stored in the dtsx (in 2005 and later versions) format.

# TRANSFORMATIONS

Below are some of the useful transformations in SSIS. All the transformations that we perform under data flow are considered as row-wise transformations.

## CHARACTER MAP

Applies string functions to string data. For example, converting column data from lower case to upper case. Only operates on columns with string data type.

## COPY COLUMN

Copy column transformation copies the input column(s) to new columns in the transformed output. For example, if we want to get First Name and last name from a Name column then we use copy column transformation and clean the transformed columns in the subsequent steps.

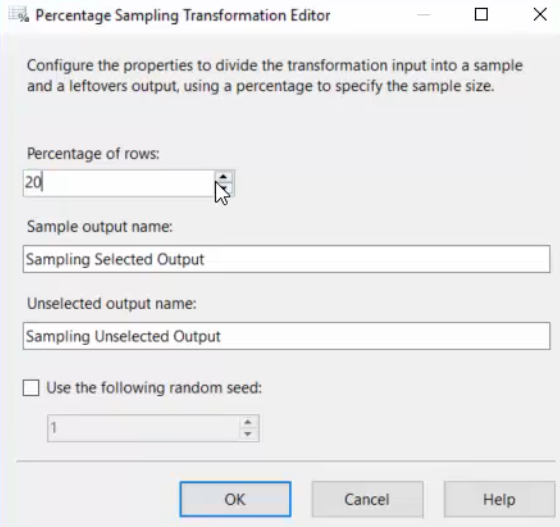
## DERIVED COLUMN

Creates a new column by applying expressions to transformation input columns. It creates either a new column or we can overwrite the computed column data to an existing column. For example, concatenating values from Firstname, Middlename and Lastname columns to Fullname column.

## PERCENTAGE SAMPLING

Randomly samples a percentage of rows from an input data flow. We need to mention the % of rows that we want to sample from the dataset.

The transformation will have sample output and unselected output.



We can either dump the records into different destinations or we can create an Excel destination with two different sheets for the two outputs. This can be done by creating an excel destination and then mapping the out of the row sampling transformation to different sheets.

## ROW SAMPLING

This transformation works similar to percentage sampling, only difference being that instead of specifying % of rows we specify number of rows.

## SORT TRANSFORMATION

Sorts input data in ascending or descending order when it can not be sorted at source. We can have sorting based on multiple columns by specifying sort order.

## UNION ALL

Combines rows from multiple data sources without sorting. Data from the 1st source will be copied to destination and then data from 2nd source will be copied to destination and so on.

Data is not sorted by UNION ALL. If sorting is important then we can use Merge transformation. Number of columns and type of columns should be same in both the data sources given to the transformation as input.

## MERGE TRANSFORMATION

Merge transformation will combine data from multiple sources as UNION ALL but the data is sorted 1st and then copied to destination.

Data should be sorted before feed to MERGE transformation.

## MERGE JOIN

Combines two data flows (inputs) into one using a join (left, full or inner).

## CONDITIONAL SPLIT

This transformation is used to split data based on some condition(s). This transformation routes the data to different destinations depending on the content of the data.

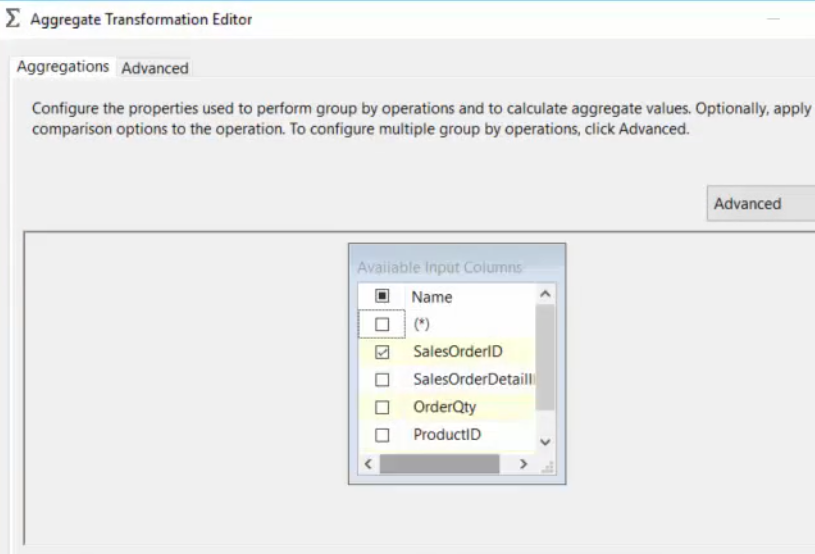
## AGGREGATE TRANSFORMATION

Aggregate data with functions like Sum, Count, Max and Min. For example, group data to find total sales per customer or total sales per product id etc.

Firstly, we have to specify the column based on which we would like to group the data and then we can apply the desired aggregate fucntion.

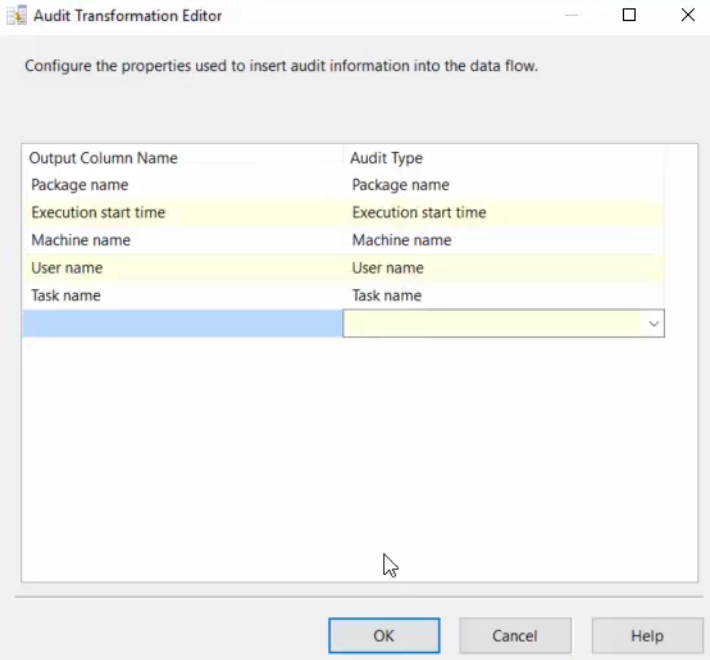
We can have multiple outputs from this transformation if we cerate multiple group by conditions in it. For each group by, we will have a separate output.

This can be by clicking on Advanced button



## AUDIT TRANSFORMATION

Includes the data about the environment in which the package runs.



## EXPORT COLUMN

Reads data from a data flow task and inserts it into a file. With the help of this transformation we can export a specific column of a table to file system.

It takes only DT\_TEXT, DT\_NTEXT AND DT\_IMAGE data types as input.

The table should also have a filepath column so that each rows data is dumped to the filepath.

## IMPORT COLUMN

Imports data from file system and puts it into a column of a table. It takes only DT\_TEXT, DT\_NTEXT AND DT\_IMAGE data types as input.

We should have a text file with the paths for other files to be imported.

For example, we want to import images of employees.

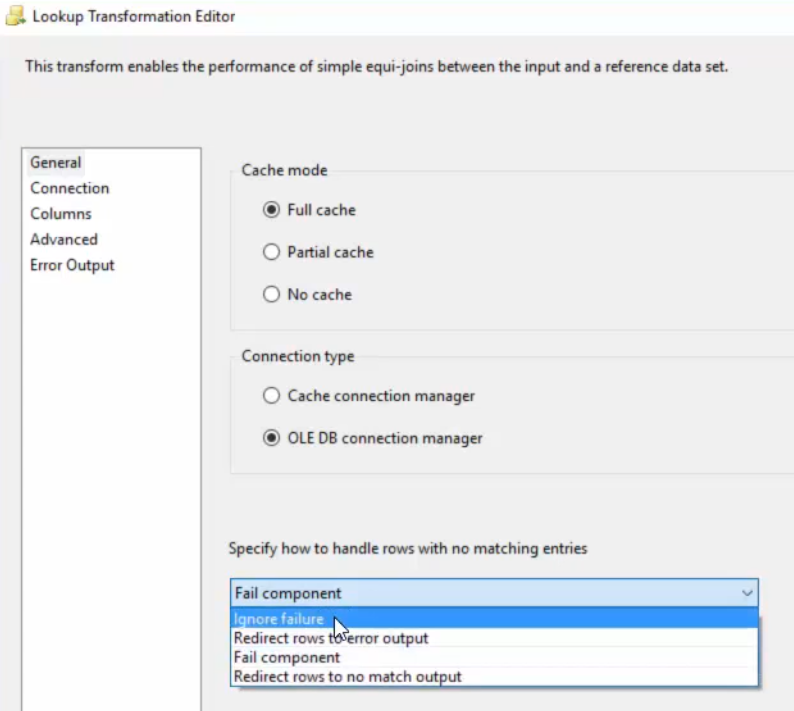
## DATA CONVERSION

Converts data from one type to other. For example, if we want to convert string data type to text data type, we can use this transformation.

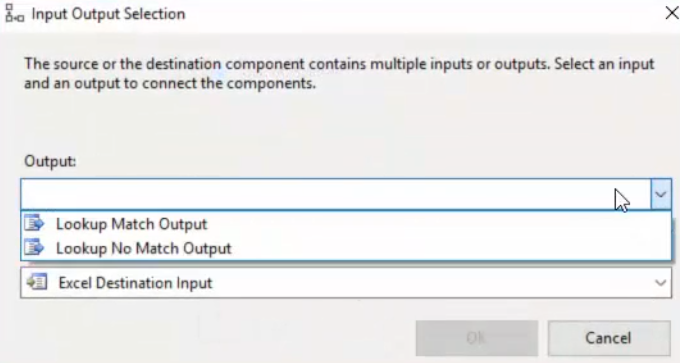
## LOOKUP TRANSFORMATION

This transformation performs equi join between the input dataset and the referenced data set.

Lookup transformation only allows inner join where as merge join allows FULL, LEFT and INNER join. Also, for merge join, the data needs to be sorted where as for lookup transformation it not required.



Output of the transformation will have below options:



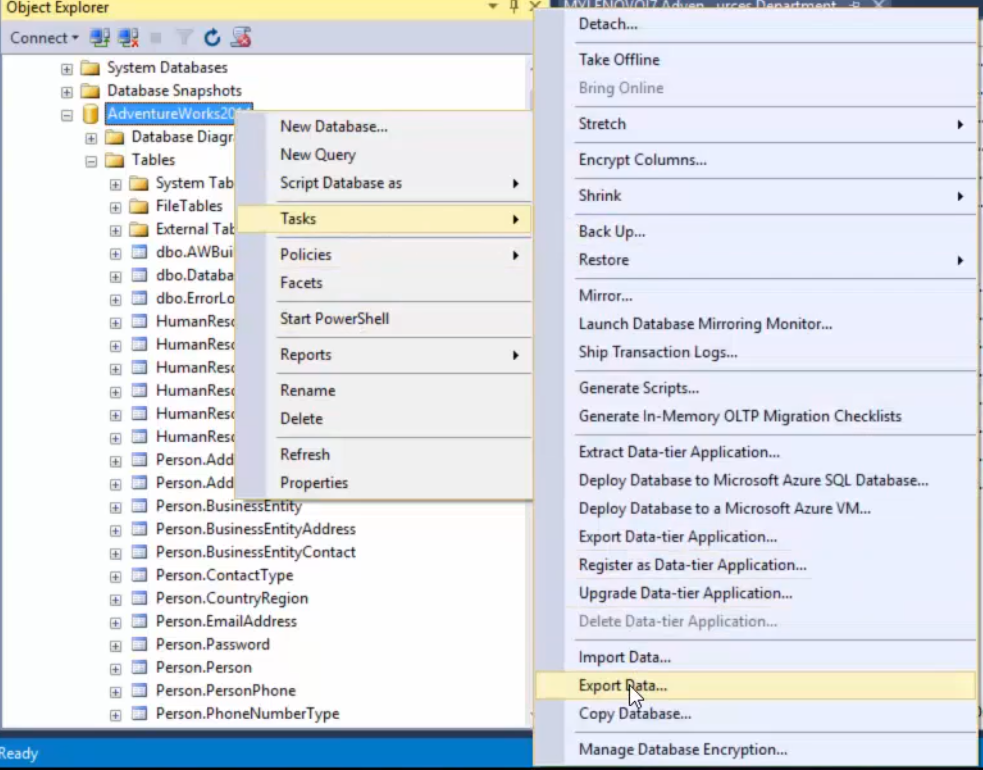
## FUZZY LOGIC TRANSFORMATION

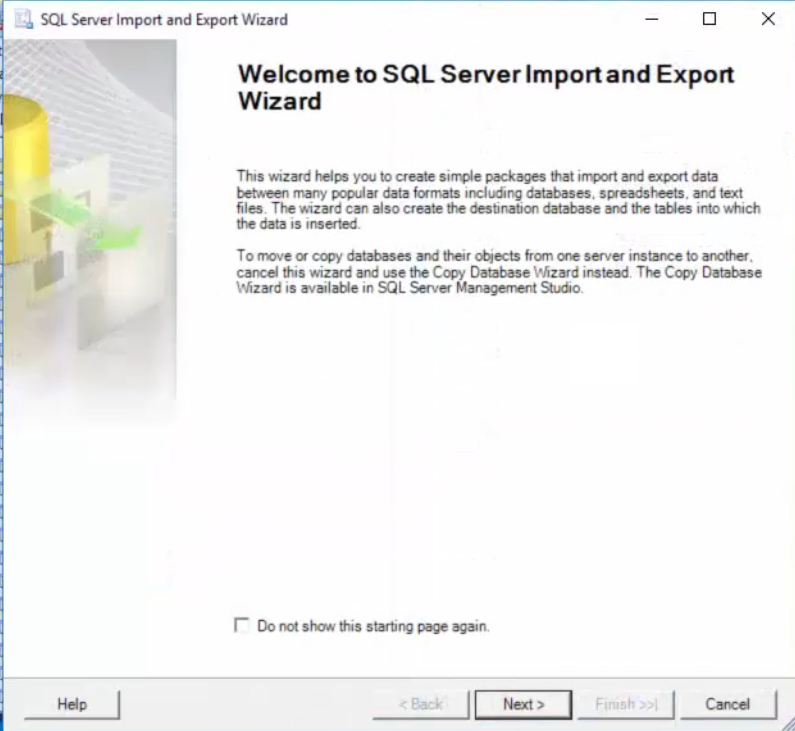
# REFERENCES

Below are some screenshots

## IMPORT EXPORT WIZARD

We can either invoke it from SSMS or via the Programs menu.

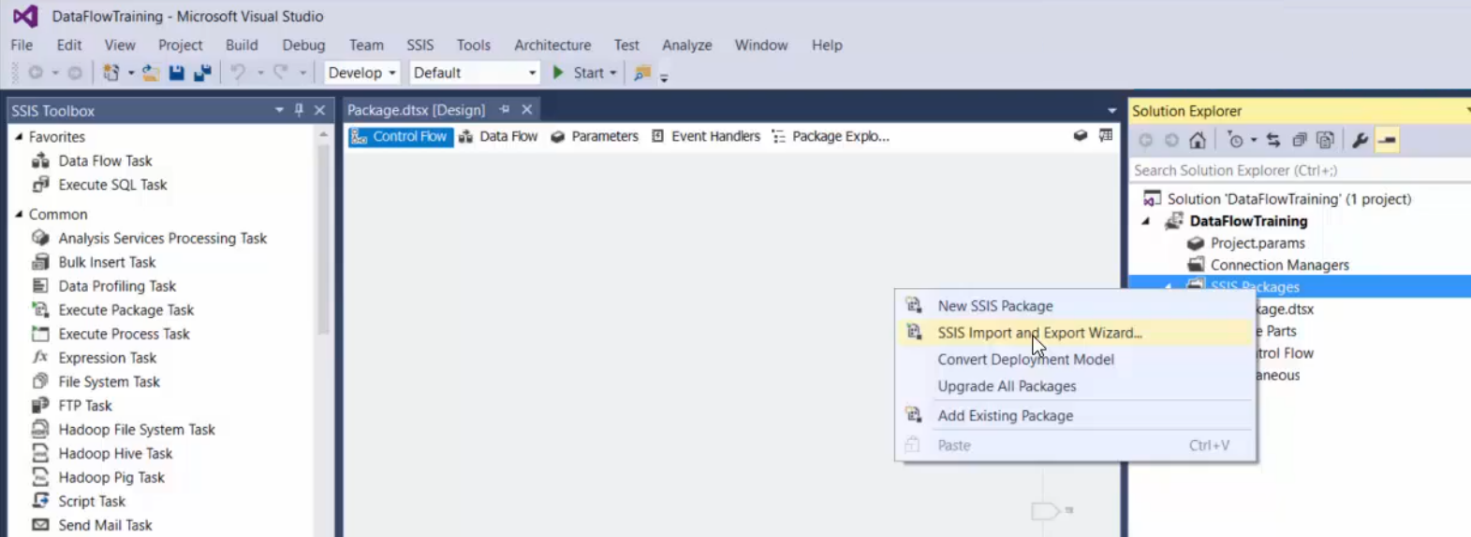




Then

* Choose the data source
* Choose the destination
* Specify table copy or query
* Review data type and Mappings
* Save and run package

We can also invoke it from within SSIS Integration project as below:



## EXCEL DESTINATION ERROR

When we get 64-bit error in our SSIS package, we can change the configuration properties of the package as shown below:

