

Azure Data Factory transformation types

Describe Azure Data Factory transformation types

Mapping Data Flows provides a number of different transformation types that enable you to modify data. They are broken down into the following categories:

Category Name	Description
Schema modifier transformations	These types of transformations will make a modification to a sink destination by creating new columns based on the action of the transformation. An example of this is the Derived Column transformation that will create a new column based on the operations performed on existing column.
Row modifier transformations	These types of transformations impact how the rows are presented in the destination. An example of this is a Sort transformation that orders the data.
Multiple inputs/outputs transformations	These types of transformations will generate new data pipelines or merge pipelines into one. An example of this is the Union transformation that combines multiple data streams.

Below is a list of transformations that is available in the Mapping Data Flows

Name	Category	Description
Aggregate	Schema modifier	Define different types of aggregations such as SUM, MIN, MAX, and COUNT grouped by existing or computed columns.
Alter row	Row modifier	Set insert, delete, update, and upsert policies on rows. You can add one-to-many conditions as expressions. These conditions should be specified in order of priority, as each row will be marked with the policy corresponding to the first-matching expression. Each of those conditions can result in a row (or rows) being inserted, updated, deleted, or upserted. Alter Row can produce both DDL & DML actions against your database.
Conditional split	Multiple inputs/outputs	Route rows of data to different streams based on matching conditions.
Derived column	Schema modifier	generate new columns or modify existing fields using the data flow expression language.
Exists	Multiple inputs/outputs	Check whether your data exists in another source or stream.
Filter	Row modifier	Filter a row based upon a condition.
Flatten	Schema modifier	Take array values inside hierarchical structures such as JSON and unroll them into individual rows.
Join	Multiple inputs/outputs	Combine data from two sources or streams.
Lookup	Multiple inputs/outputs	Enables you to reference data from another source.
New branch	Multiple inputs/outputs	Apply multiple sets of operations and transformations against the same data stream.
Pivot	Schema modifier	An aggregation where one or more grouping columns has distinct row values transformed into individual columns.
Select	Schema modifier	Alias columns and stream names, and drop or reorder columns.
Sink	-	A final destination for your data.
Sort	Row modifier	A data source for the data flow.
Source	-	A data source for the data flow.
Surrogate key	Schema modifier	Add an incrementing non-business arbitrary key value.
Union	Multiple inputs/outputs	Combine multiple data streams vertically.
Unpivot	Schema modifier	Pivot columns into row values.
Window	Schema modifier	Define window-based aggregations of columns in your data streams.

Data Flow Expression Builder

Some of the transformations that you can define have an Data Flow Expression Builder that will enable you to customize the functionality of a transformation using columns, fields, variables, parameters, functions from your data flow in these boxes. To build the expression, use the Expression Builder, which is launched by clicking in the expression text box inside the transformation. You'll also sometimes see "Computed Column" options when selecting columns for transformation. When you click that, you'll also see the Expression Builder launched.

Visual expression builder

Expression reference documentation

FUNCTIONS

Filter...

All Functions Input schema Parameters

abc movie

abc title

abc genres

year <= 1920

+ - * / || && ! ^ == === <=> != > < >= <= []

Data preview

Refresh

Output: ✓	year ¹²³
✓	1920
✗	1921
✓	1920
✗	1921

Save and finish

Cancel

Clear contents

U

sing the expression builder in the Mapping Data Flow

The Expression Builder tool defaults to the text editor option. the auto-complete feature reads from the entire Azure Data Factory Data Flow object model with syntax checking and highlighting