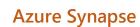
## **Transformations of Mapping Data Flow**

# Azure Data Factory #MappingDataFlows





Activity	Description	SSIS equivalent		SQL Server equivalent
Multiple inputs/outputs				
New branch	Create a new flow branch with the same data	<b>从</b> ⋈	ulticast (+icon)	SELECT INTO SELECT OUTPUT
Join	Join data from two streams based on a condition	<b>*</b> ₩	lerge join	INNER   LEFT   RIGHT JOIN, CROSS   FULL OUTER JOIN
Conditional Split	Route data into different streams based on conditions	<b>₩</b> Co	onditional Split	SELECT INTO WHERE condition1 SELECT INTO WHERE condition2 CASE WHEN
<b>Exists</b>	Check the existence of data in another stream	-	Lookup / Merge Join	SELECT * FROM Table WHERE EXISTS (SELECT )   JOIN   NOT EXISTS
Union	Collect data from multiple streams	<b>T</b> U	nion All	SELECT colla UNION (ALL) SELECT collb
Lookup	Lookup additional data from another stream	Lo	ookup	LEFT   RIGHT JOIN
Schema modifier				
Cast	Cast columns to different types	fx D	erived Column	SELECT CAST (NumStr as INT) as Number, CONVERT (DATE, '14/07/2022', 103)
Derived Column	Compute new columns based on the existing once	fx D	erived Column	SELECT Column1 * 1.09 as NewColumn
<b>Select</b> ■	Choose columns to flow to the next stream		UTPUT in components, apping columns	SELECT Column1, Column4 FROM Table
∑ Aggregate	Calculate aggregation on the stream	<b>(Ξ)</b> Σ Α	ggregate	SELECT Year(DateOfBirth) as Year, MIN(), MAX(), AVG() GROUP BY Year(DateOfBirth)
Surrogate Key	Add a surrogate key column to output stream from a specific value		cript omponent	SELECT ROW_NUMBER()  OVER(ORDER BY n ASC) AS R#, n  FROM sys.databases + Incremental Primary Key (with limited capabilities)
Pivot	Pivots row values into columns, groups columns and aggregates data	Pi	vot	SELECT rowCol, c1, c2 FROM ( SELECT sourceCols FROM Table) PIVOT ( SUM(sumCol) FOR col IN () )
Unpivot	Unpivots columns into row values and ungroups columns	<b>5</b> ⊞ ∪	npivot	SELECT rowCol, col, X FROM ( SELECT rowCol, c1, c2 FROM pvt) UNPIVOT (X FOR col FROM (c1, c2)) AS unpvt
Window	Aggregates data based on a window and joins with original data	11 1	[Sort] + Custom Script	SELECT fun() OVER( PARTITION BY pc ORDER BY oc) newc, pc, oc, otherCols FROM Table
Rank	Generates an ordered ranking based upon sort conditions specified by the user	11 1	[Sort] + Custom Script	SELECT ROW_NUMBER() OVER ( PARTITION BY pc ORDER BY oc) newc, pc, oc, otherCols FROM Table
External Call	Enables custom function by calling out to external REST endpoints row-by-row		cript omponent	Not available out of the box. Some tricks are possible: OLE Automation, CLR function, although not recommended.







Sink

**Azure Data Factory** 

Destination for your data flow

stream

### **Transformations of Mapping Data Flow**

#### #MappingDataFlows



Azure Synapse

#### Activity Description SSIS equivalent SQL Server equivalent **Formatters** SELECT \* FROM OPENJSON(json) WITH ( id INT 'strict \$.id', Takes array values from Script Component or Flatten hierarchical structures such as 3<sup>rd</sup> Party Component Name NVARCHAR(50) '\$.info.name'); JSON (denormalization) (ZappySys SSIS JSON Parser Transform) SELECT \* FROM OPENJSON ... Parses columns in document Script Component or SELECT \* FROM OPENXML ... **Parse** form data (JSON, XML, 3<sup>rd</sup> Party Component SELECT value FROM delimited text) (ZappySys SSIS JSON Parser Transform) STRING SPLIT('clothing, road', ','); Turns complex data types into Script Stringify strings Component **Row modifier** SELECT \* FROM Table Filter rows in the stream Filter **Conditional Split** WHERE [Column] LIKE '%pattern%' based on a condition SELECT \* FROM Table Order data in the stream Sort Sort ORDER BY [Column] ASC based on column(s) MERGE Set action policy on rows **Conditional Split Alter Row** INSERT, UPDATE, DELETE when database is sink + n Destinations IF | WHERE WHERE, REPLACE, SUBSTRING, Enables you to build custom Conditional Split or IF, RAISERROR, THROW rules for data quality and data **DQS Cleansing** validation Source / Destination SELECT \* FROM SourceTable Source for your data flow. **OLE DB Source** Source Obligatory first element of and more ... every Data Flow in ADF

Version: 2.1 (Aug 2022)

INSERT INTO TargetTable





**OLE DB Destination and** 

more...