Azure Data Factory – Data Flow Components

#ADFDF

Activity	Description	SSIS equivalent	SQL Server equivalent
New branch	Create a new flow branch with the same data	Multicast (+icon)	`SELECT INTO` `SELECT OUTPUT`
Join	Join data from two streams based on a condition	Merge join	`INNER/LEFT/RIGHT JOIN,` `CROSS/FULL OUTER JOIN`

Conditional Split	Route data into different streams based on conditions	Conditional Split	`SELECT INTO WHERE condition1`
			`SELECT INTO WHERE condition2`
			`CASE ... WHEN`

Union	Collect data from multiple streams	Union All	`SELECT col1a UNION (ALL)` `SELECT col1b`
Lookup	Lookup additional data from another stream	Lookup	`LEFT/RIGHT JOIN`
Derived Column	Compute new columns based on the existing once	Derived Column	`SELECT Column1 * 1.09 as NewColumn`

Aggregate	Calculate aggregation on the stream	Aggregate	`SELECT Year(DateOfBirth) as YearOnly,`
			`MIN(), MAX(), AVG()`
			`GROUP BY Year(DateOfBirth)`

Surrogate Key	Add a surrogate key column to output stream from a specific value	Script Component	`SELECT ROW_NUMBER()`
			`   OVER(ORDER BY name ASC) AS Row#,`
			`name`
			`FROM sys.databases`
			
			Incremental Primary Key (with limited capabilities)

Exists	Check the existence of data in another stream	Lookup / Merge Join	`SELECT * FROM Table` `WHERE EXISTS(SELECT ...)`
Select	Choose columns to flow to the next stream	OUTPUT in components, mapping columns	`SELECT Column1, Column4` `FROM Table`
Filter	Filter rows in the stream based on a condition	Conditional Split	`SELECT * FROM Table` `WHERE [Column] LIKE '%pattern%'`
Sort	Order data in the stream based on column(s)	Sort	`SELECT * FROM Table` `ORDER BY [Column] ASC`
Extend	Use any custom logic from an external library	Script Component	SQL CLR
Source	Source for your data flow. Obligatory first element of every Data Flow in ADF	OLE DB Source and more …	`SELECT * FROM SourceTable`
Sink	Destination for your data flow	OLE DB Destination and more…	`INSERT INTO TargetTable`

Version: 1.0 (05.01.2019)

https://SQLPlayer.net/tag/ADFDF

@SQLPlayer