



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)	<code>SELECT INTO</code> <code>SELECT OUTPUT</code>
Join	Join data from two streams based on a condition	Merge join	<code>INNER</code> <code>LEFT</code> <code>RIGHT JOIN</code> , <code>CROSS</code> <code>FULL OUTER JOIN</code>
Conditional Split	Route data into different streams based on conditions	Conditional Split	<code>SELECT INTO WHERE</code> condition1 <code>SELECT INTO WHERE</code> condition2 <code>CASE ... WHEN</code>
Union	Collect data from multiple streams	Union All	<code>SELECT</code> colla <code>UNION (ALL)</code> <code>SELECT</code> collb
Lookup	Lookup additional data from another stream	Lookup	<code>LEFT</code> <code>RIGHT JOIN</code>
Derived Column	Compute new columns based on the existing once	Derived Column	<code>SELECT</code> Column1 * 1.09 <code>as</code> NewColumn
Aggregate	Calculate aggregation on the stream	Aggregate	<code>SELECT</code> Year(DateOfBirth) <code>as</code> Year, <code>MIN()</code> , <code>MAX()</code> , <code>AVG()</code> <code>GROUP BY</code> Year(DateOfBirth)
Surrogate Key	Add a surrogate key column to output stream from a specific value	Script Component	<code>SELECT ROW_NUMBER()</code> <code>OVER(ORDER BY n ASC) AS R#, n</code> <code>FROM sys.databases</code> <code>+ Incremental Primary Key (with limited capabilities)</code>
Pivot <small>(NEW)</small>	Pivots row values into columns, groups columns and aggregates data	Pivot	<code>SELECT</code> rowCol, c1, c2 <code>FROM</code> (<code>SELECT</code> sourceCols <code>FROM</code> Table) <code>PIVOT</code> (<code>SUM</code> (sumCol) <code>FOR</code> col <code>IN</code> (...))
Unpivot <small>(NEW)</small>	Unpivots columns into row values and ungroups columns	Unpivot	<code>SELECT</code> rowCol, col, X <code>FROM</code> (<code>SELECT</code> rowCol, c1, c2 <code>FROM</code> pvt) <code>UNPIVOT</code> (X <code>FOR</code> col <code>FROM</code> (c1, c2)) <code>AS</code> unpvt
Window <small>(NEW)</small>	Aggregates data based on a window and joins with original data	[Sort] + Custom Script	<code>SELECT</code> fun() <code>OVER</code> (<code>PARTITION BY</code> pc <code>ORDER BY</code> oc) newc, pc, oc, otherCols <code>FROM</code> Table
Exists	Check the existence of data in another stream	Lookup / Merge Join	<code>SELECT * FROM</code> Table <code>WHERE EXISTS</code> (<code>SELECT ...</code>) <code>JOIN</code>
Select	Choose columns to flow to the next stream	OUTPUT in components, mapping columns	<code>SELECT</code> Column1, Column4 <code>FROM</code> Table
Filter	Filter rows in the stream based on a condition	Conditional Split	<code>SELECT * FROM</code> Table <code>WHERE</code> [Column] <code>LIKE</code> '%pattern%'
Sort	Order data in the stream based on column(s)	Sort	<code>SELECT * FROM</code> Table <code>ORDER BY</code> [Column] <code>ASC</code>
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 ...	<code>SELECT * FROM</code> SourceTable
Sink	Destination for your data flow	OLE DB Destination and more...	<code>INSERT INTO</code> TargetTable

Version: 1.01 (19.01.2019)