

Data Flow Flatten



Flatten



Data Flow – Flatten

- Utilice la transformación “Flatten” para tomar valores de matriz dentro de estructuras jerárquicas como JSON y desenrollarlos en filas individuales. Este proceso se conoce como desnormalización.



The screenshot displays the Data Flow tool interface. At the top, there are tabs for 'Save', 'Validate', and 'Debug Settings'. Below these, a workflow is shown with four components: 'source1' (Columns: 11 total), 'Flatten' (Columns: 11 total), 'DerivedColumn1' (Creating/updating the columns: orderID, customerID, employeeID, orderDate, requiredDate, shippedDate), and 'sink1' (Export data to sqlserver). A hand cursor is pointing at the 'Flatten' component. Below the workflow, there is a dashed box labeled 'Add Source'. At the bottom, the 'Data preview' tab is active, showing a table with 5 columns: 'orderID', 'customerID', 'employeeID', 'orderDate', and 'requiredDate'. The table contains 5 rows of data.

orderID	customerID	employeeID	orderDate	requiredDate
10271	SPLIR	6	1996-08-01 00:00:00.000	1996-08-29 00:00:00.000
10266	WARTH	3	1996-07-26 00:00:00.000	1996-09-06 00:00:00.000
10279	LEHMS	8	1996-08-13 00:00:00.000	1996-09-10 00:00:00.000
10292	TRADH	1	1996-08-28 00:00:00.000	1996-09-25 00:00:00.000
10295	VINET	2	1996-09-02 00:00:00.000	1996-09-30 00:00:00.000

Ejercicio – Flatten

- ▶ Se requiere cargar archivos en formato JSON(documento Único, documento por línea y Matriz) hacia Azure SQL.
- ▶ Para realizar la carga de datos se tiene que utilizar la transformación de “Flatten”.



Flatten

The screenshot displays the Azure Data Studio interface for a data pipeline. The pipeline consists of four stages: 'source1' (Columns: 11 total), 'Flatten' (Columns: 11 total), 'DerivedColumn1' (Creating/updating the columns: orderID, customerID, employeeID, orderDate, requiredDate, shippedDate), and 'sink1' (Export data to sqlserver). The 'Flatten' stage is highlighted with a hand icon. Below the pipeline, there is an 'Add Source' button. The bottom section shows the 'Data preview' tab, which displays a table with 5 rows of data. The table has columns: orderID, customerID, employeeID, orderDate, and requiredDate. The data is as follows:

orderID	customerID	employeeID	orderDate	requiredDate
10271	SPLIR	6	1996-08-01 00:00:00.000	1996-08-29 00:00:00.000
10266	WARTH	3	1996-07-26 00:00:00.000	1996-09-06 00:00:00.000
10279	LEHMS	8	1996-08-13 00:00:00.000	1996-09-10 00:00:00.000
10292	TRADH	1	1996-08-28 00:00:00.000	1996-09-25 00:00:00.000
10295	VINET	2	1996-09-02 00:00:00.000	1996-09-30 00:00:00.000