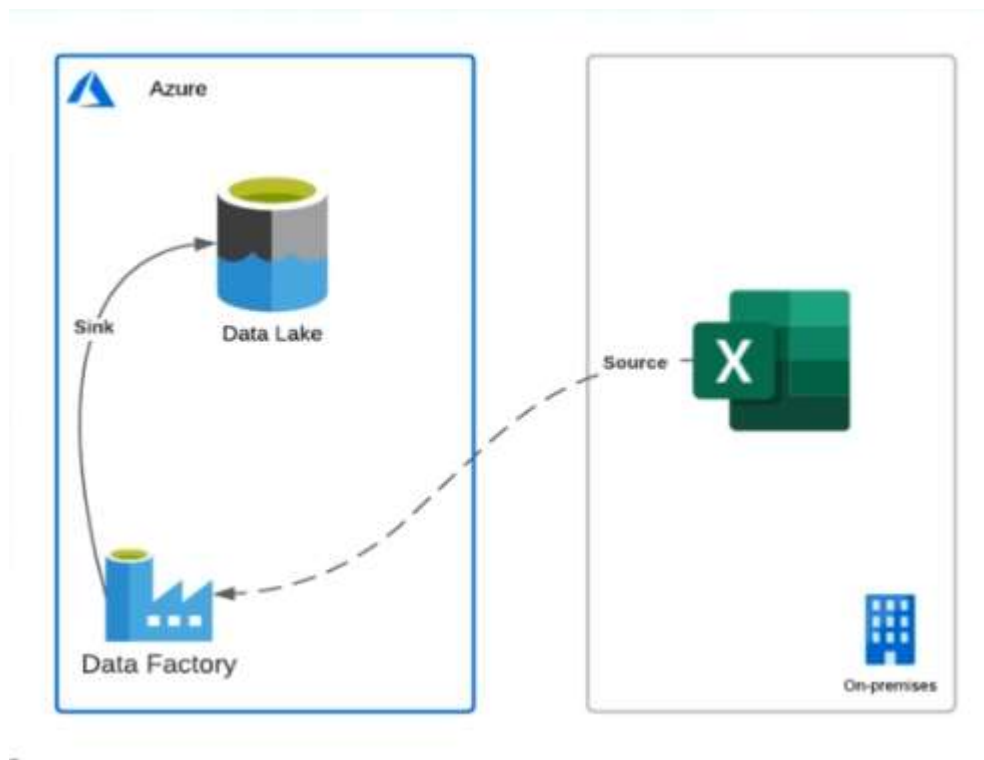
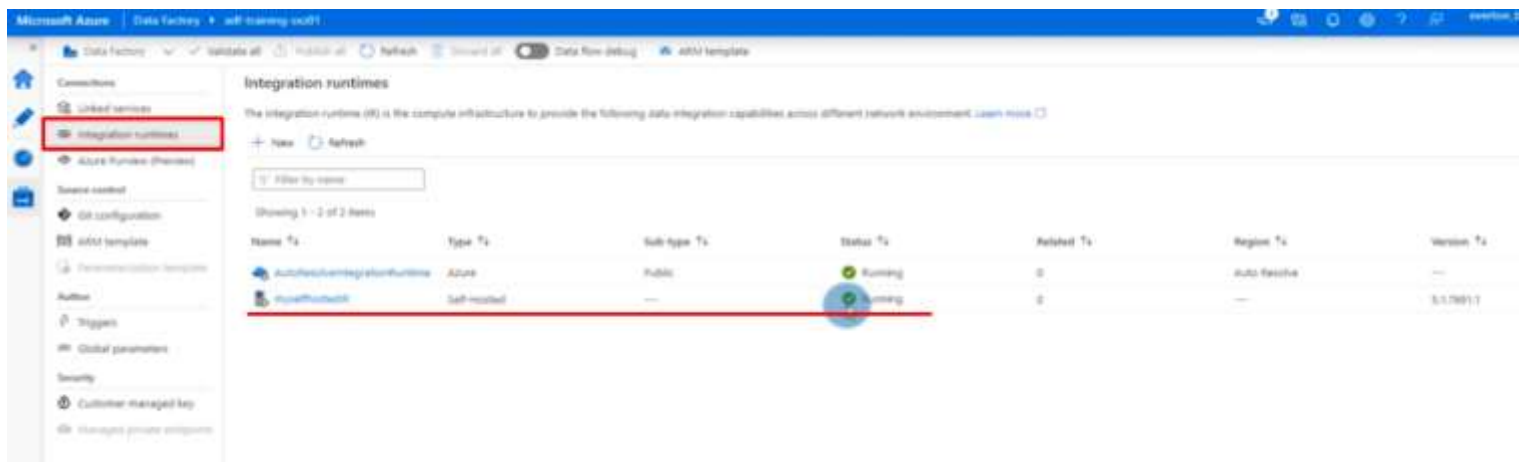


Copiar datos (CSV) desde On-premises hacia ADLS

Este es el diagrama de lo que haremos



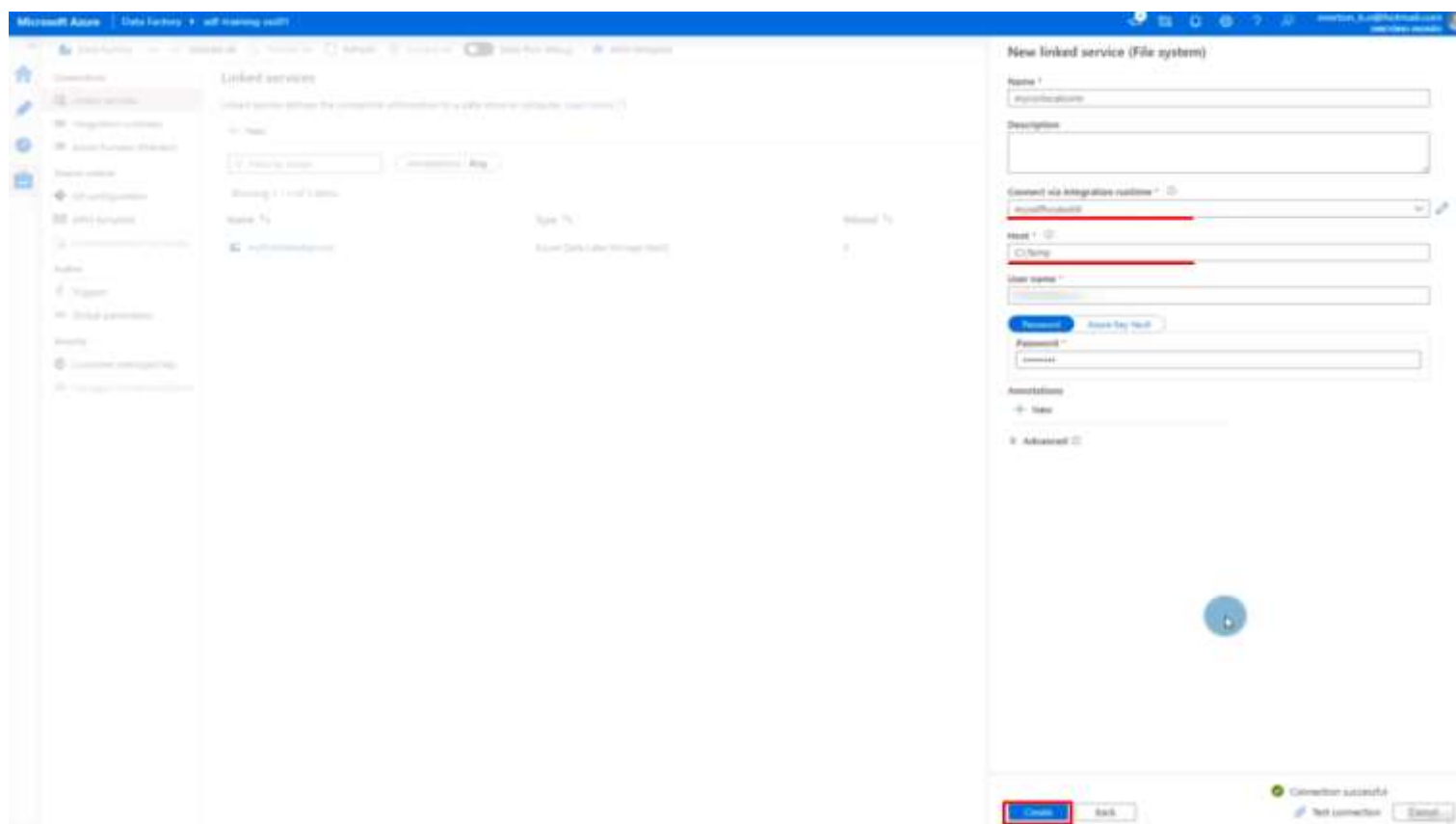
Antes de comenzar debemos tener creado un **Self-Hosted Integration Runtime**. Este nos permitirá generar una puerta de enlace entre la nube y un servidor local.



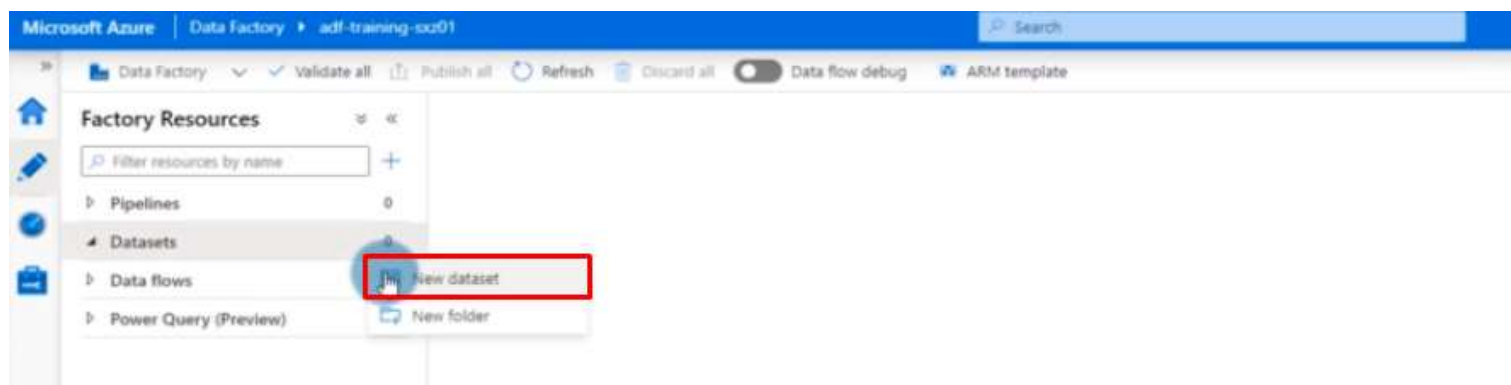
Crearemos un **Linked Service** que haga referencia hacia nuestro **sistema local**



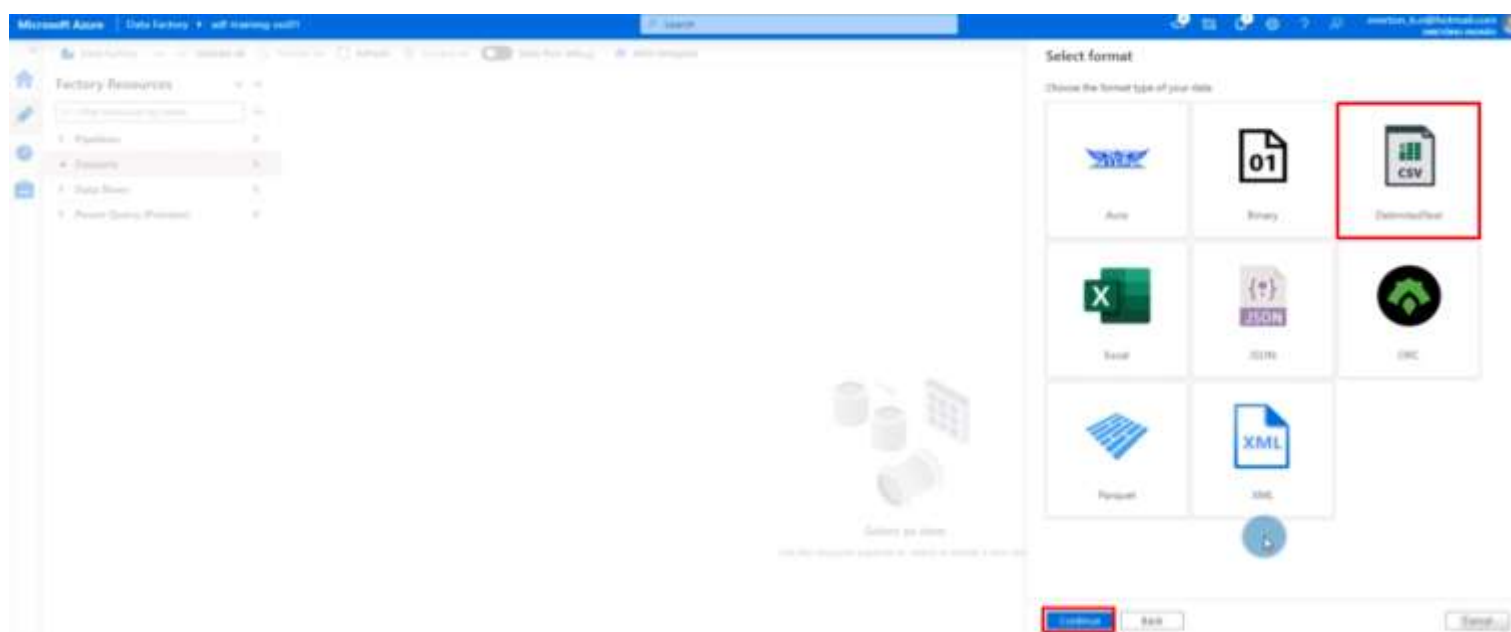
Debemos escoger el **Self-Hosted Integration Runtime** que en este caso lleva el nombre de **myselfhostedIR**. En **Host** indicamos el directorio que almacena el archivo que vamos a utilizar.

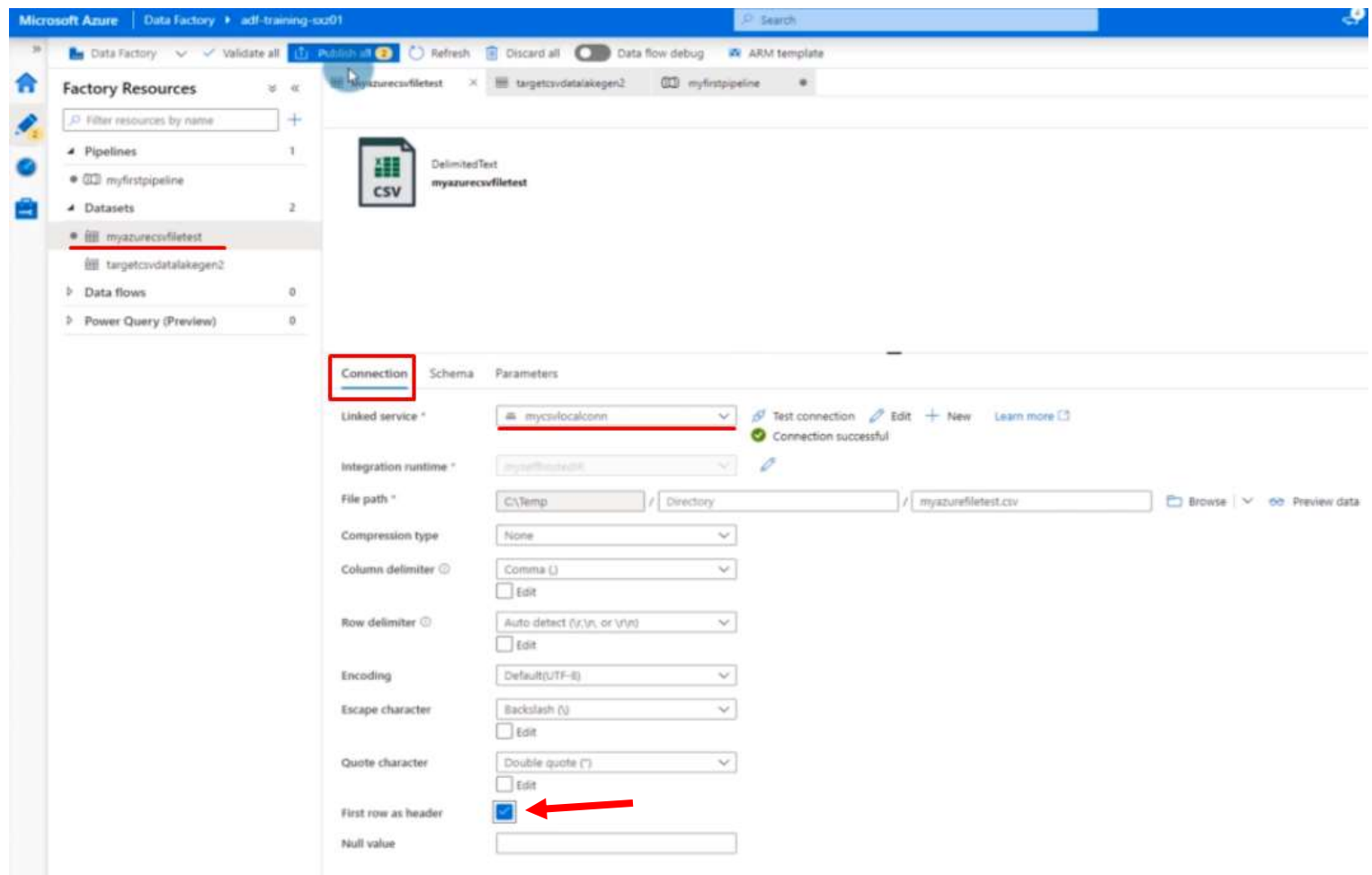
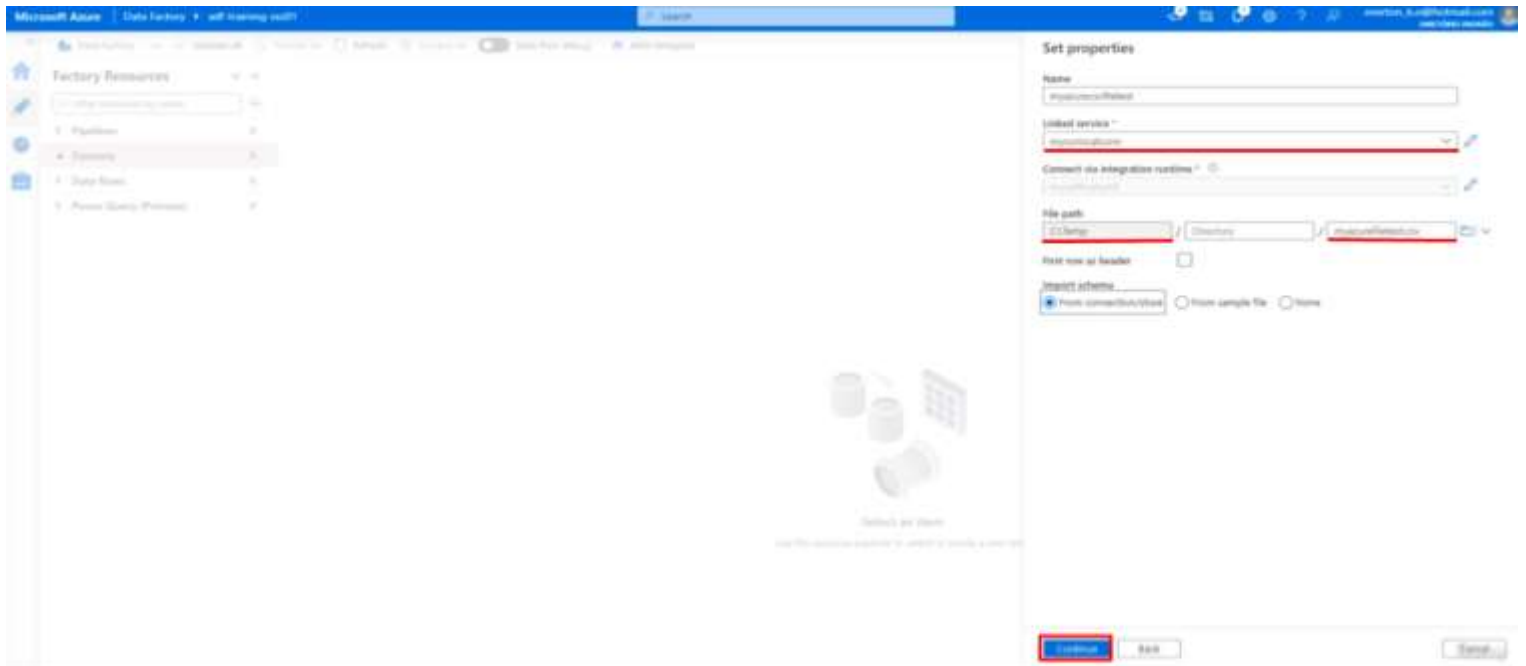


Vamos a crear un dataset de origen

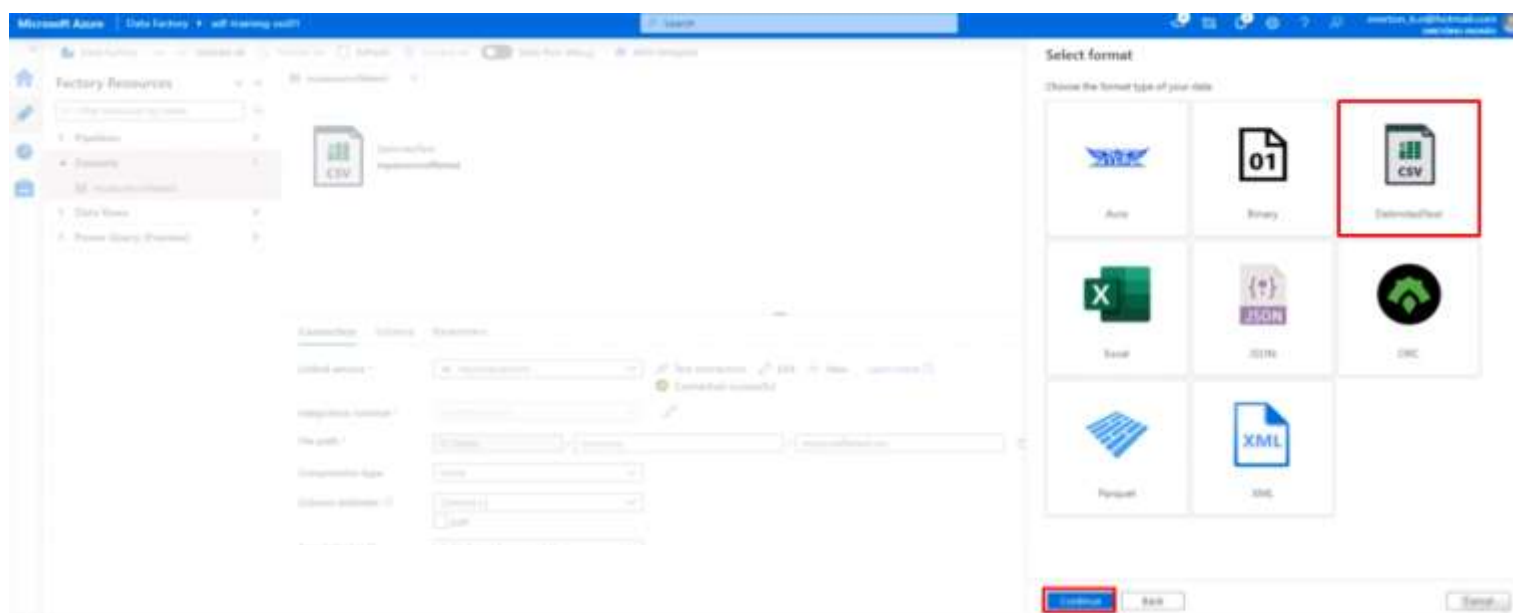


Hará referencia a nuestro **File system (sistema local)**





Ahora debemos crear nuestro dataset de destino



Dejamos vacía la ruta de destino, dado que en nuestro ADLS no tenemos creado ningún contenedor



En capítulos anteriores se creó el Linked Service que hace referencia al ADLS llamado **myfirstlinkedservice**

En la ruta de nuestro dataset de destino vamos a ingresar un valor para el contenedor, de esta forma, al no tener ningún contenedor creado en nuestro ADLS, se creará un contenedor llamado **mycsvlocation01** de manera automática. También añadimos un directorio **folder01** que también se creará de manera automática dentro de este contenedor.

The screenshot displays the Microsoft Azure Data Factory portal for a workspace named 'adf-training-ss001'. On the left, the 'Factory Resources' sidebar shows a tree view with 'Pipelines' (1 item: 'myfirstpipeline') and 'Datasets' (2 items: 'myazurecsvfiletest' and 'targetcsvdatalakegen2', which is selected). The main workspace shows a visual representation of the selected dataset as a 'DelimitedText' file icon labeled 'targetcsvdatalakegen2'. Below this, the 'Connection' tab is active, showing the configuration for the dataset. The 'Linked service' is set to 'myfirstlinkedservice'. The 'File path' is configured as 'mycsvlocation01 / folder01 / File', with 'mycsvlocation01' and 'folder01' highlighted by red boxes. Other settings include 'Compression type' (None), 'Column delimiter' (Comma), 'Row delimiter' (Auto detect), 'Encoding' (Default(UTF-8)), 'Escape character' (Backslash), 'Quote character' (Double quote), 'First row as header' (unchecked), and 'Null value' (empty). A 'Test connection' button is visible next to the linked service dropdown.

Property	Value
Linked service	myfirstlinkedservice
File path	mycsvlocation01 / folder01 / File
Compression type	None
Column delimiter	Comma (,)
Row delimiter	Auto detect (\r, \n, or \r\n)
Encoding	Default(UTF-8)
Escape character	Backslash (\)
Quote character	Double quote (")
First row as header	<input type="checkbox"/>
Null value	

Vamos a crear nuestro Pipeline llamado **myfirstpipeline**. Agregamos la actividad **Copy data**. En **Source** llamamos a nuestro dataset de origen **myazurecsvfiletest**.

The screenshot displays the Microsoft Azure Data Factory console. The top navigation bar shows 'Data Factory' and 'adf-training-ssz01'. The left sidebar lists various activities, with 'Copy data' highlighted under the 'Move & transform' category. The main workspace shows a 'Copy data' activity named 'Copy data1' within a pipeline named 'myfirstpipeline'. Below the workspace, the 'Source' tab is selected, showing the configuration for the activity. The 'Source dataset' is set to 'myazurecsvfiletest'. The 'File path type' is set to 'File path in dataset'. The 'Filter by last modified' section has 'Start time (UTC)' and 'End time (UTC)' input fields. The 'Recursively' checkbox is checked, and the 'Enable partition discovery' checkbox is unchecked.

Microsoft Azure | Data Factory | adf-training-ssz01

Search

Data Factory | Validate all | Publish all | Refresh | Discard all | Data flow debug | ARM template

myazurecsvfiletest | targetcsvdatalakegen2 | myfirstpipeline

Activities

Search activities

Move & transform

Copy data

Data flow

Azure Data Explorer

Azure Function

Batch Service

Databricks

Data Lake Analytics

General

HDInsight

Iteration & conditionals

Machine Learning

Power Query

Save as template | Validate | Validate copy runtime | Debug | Add trigger

Copy data

Copy data1

General | **Source** | Sink | Mapping | Settings | User properties

Source dataset * | myazurecsvfiletest | Open | New | Preview data | Learn more

File path type

☒ File path in dataset | ☐ File filter | ☐ Wildcard file path | ☐ List of files

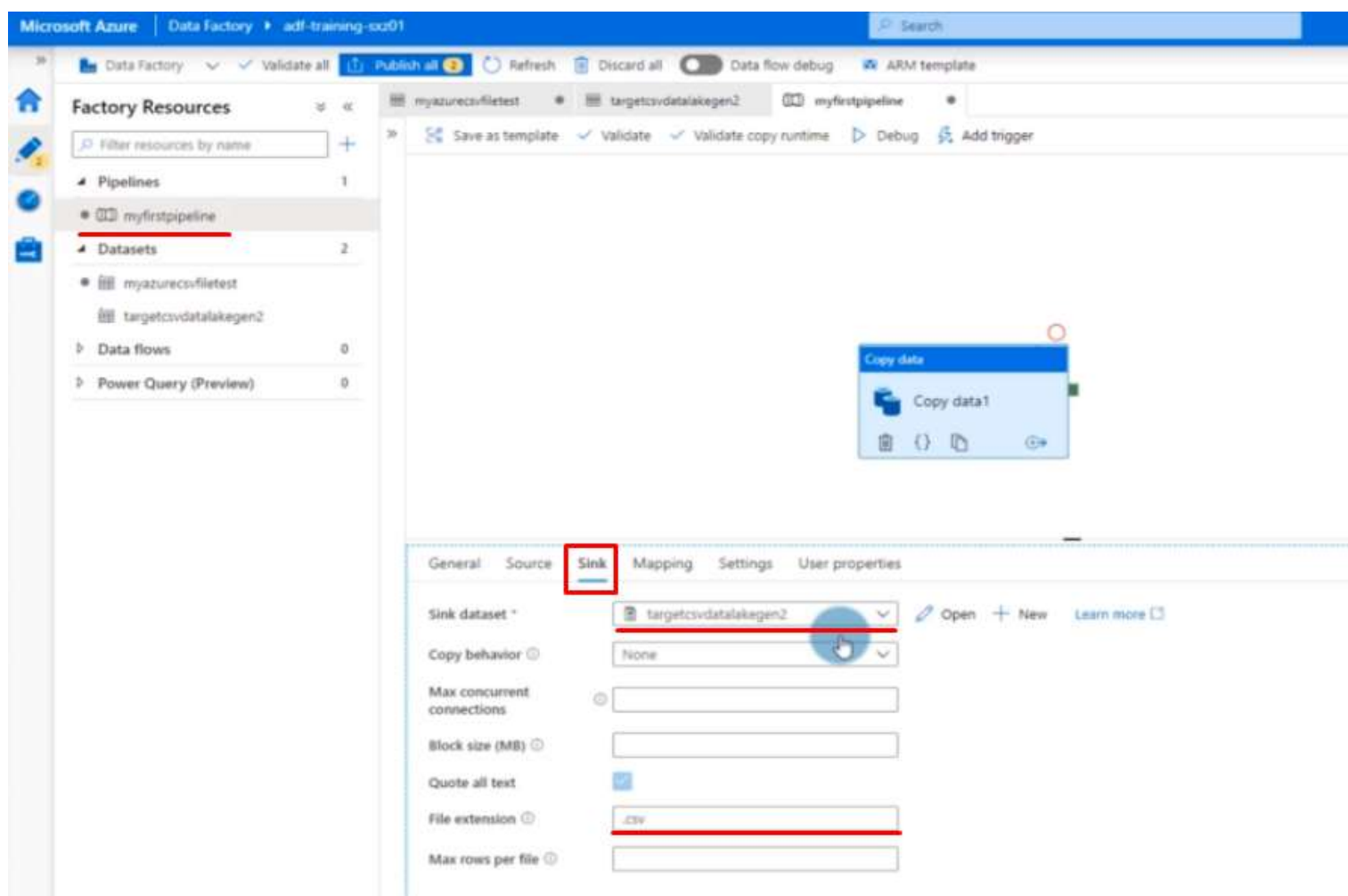
Filter by last modified

Start time (UTC) | End time (UTC)

Recursively ☒

Enable partition discovery ☐

En **Sink** llamamos a nuestro dataset de destino **targetcsvdatalakegen2**.



Después de **Validar** y ejecutar nuestro pipeline utilizando **Debug**, verificamos si copiamos el archivo en nuestro ADLS

