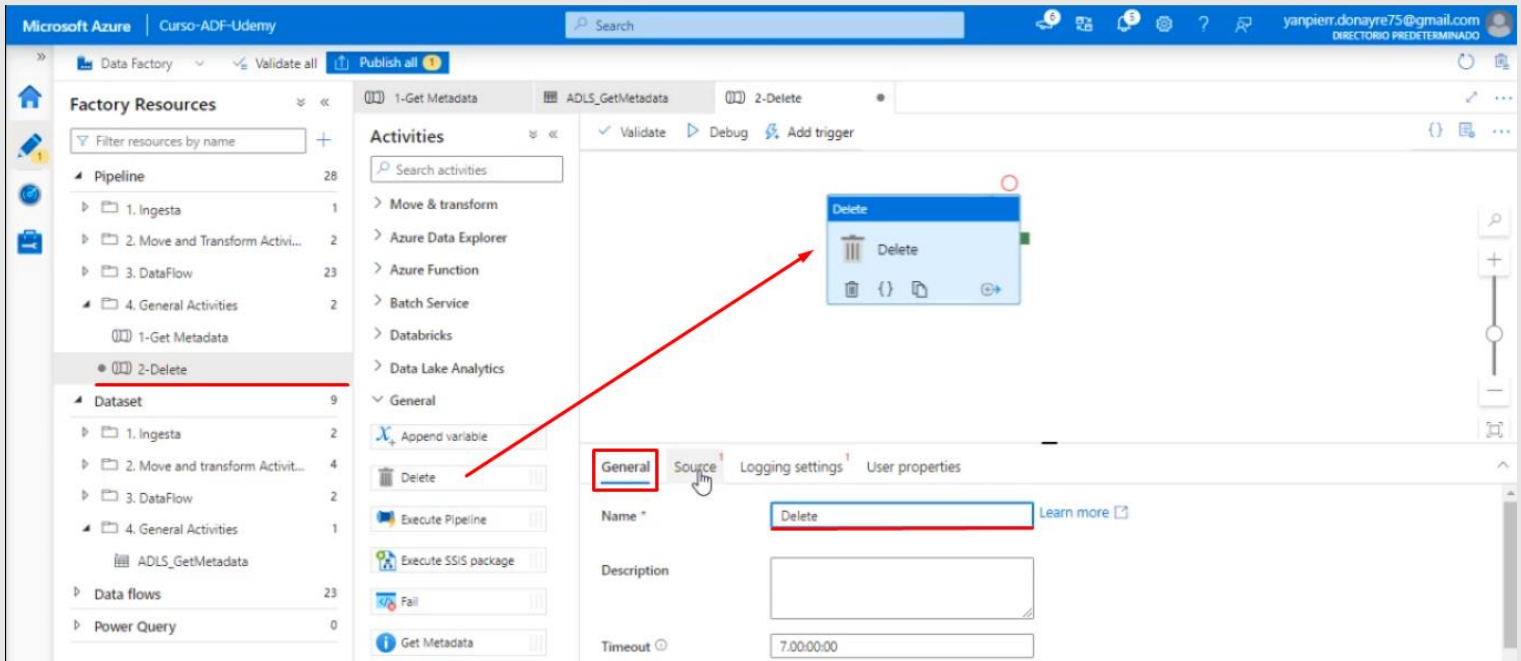
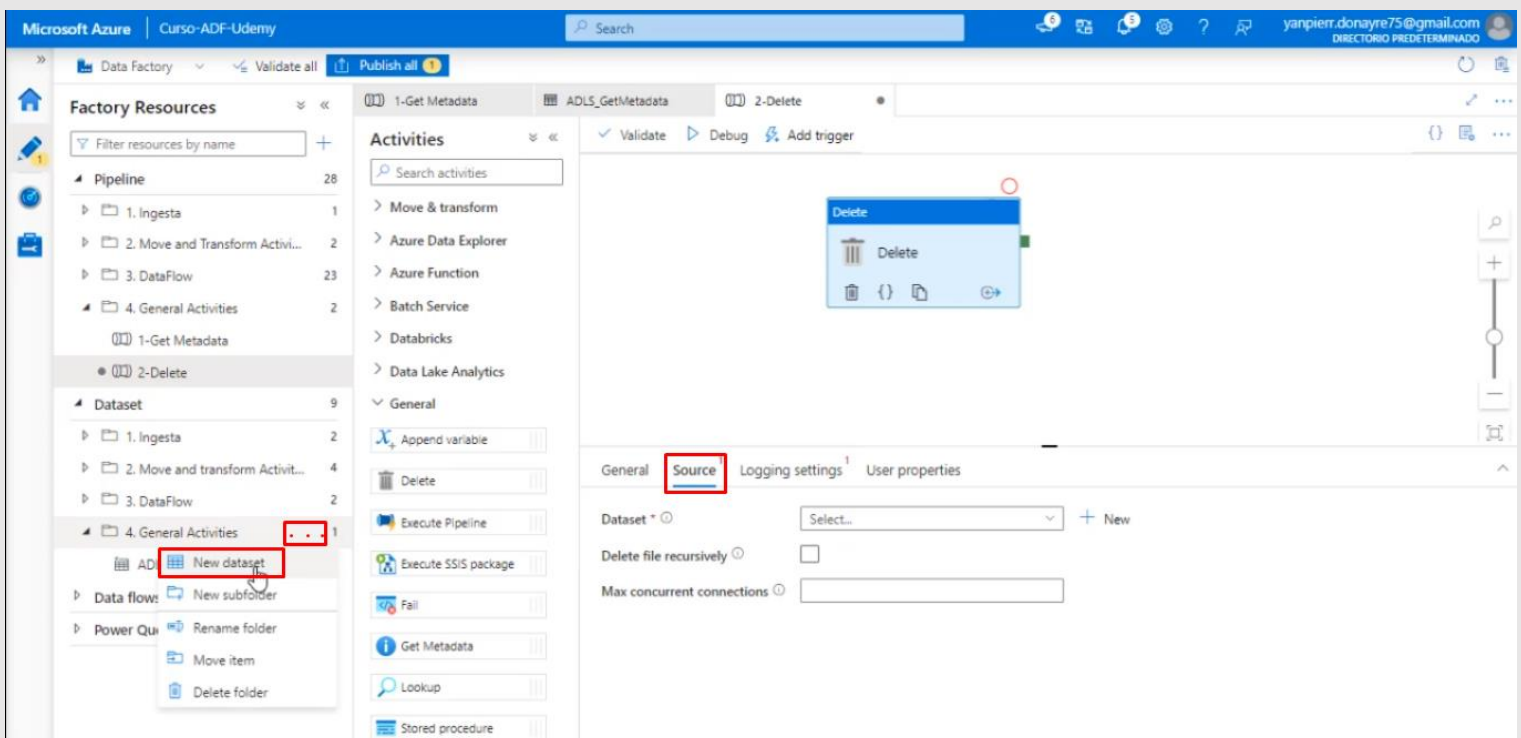


Activities Delete

Vamos a comenzar creando un Pipeline y utilizando la actividad de "Delete"



Crearemos un nuevo dataset



Microsoft Azure | Curso-ADF-Udemy

Search

yanpierr.donayre75@gmail.com
DIRECTORIO PREDETERMINADO

Data Factory | Validate all | Publish all

Factory Resources

Filter resources by name

- Pipeline 28
 - 1. Ingesta 1
 - 2. Move and Transform Activ... 2
 - 3. DataFlow 23
 - 4. General Activities 2
 - 1-Get Metadata
 - 2-Delete
- Dataset 9
 - 1. Ingesta 2
 - 2. Move and transform Activ... 4
 - 3. DataFlow 2
 - 4. General Activities 1
 - ADLS_GetMetadata
- Data flows 23
- Power Query 0

Activities

Search activities

- Move & transform
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
 - Append variable
 - Delete
 - Execute Pipeline
 - Execute SSIS package
 - Fail
 - Get Metadata
 - Lookup
 - Stored procedure

1-Get Metadata | ADLS_GetMetadata | 2-Delete

Validate | Debug | Add trigger

Delete

New dataset

In pipeline activities and data flows, reference a dataset to specify the location and structure of your data within a data store. [Learn more](#)

Select a data store

lake

All | Azure | Database | File | Generic protocol | NoSQL | Services and apps

- Azure Data Lake Storage Gen1
- Azure Data Lake Storage Gen2**
- Azure Databricks Delta Lake
- Snowflake

Continue | Cancel

Microsoft Azure | Curso-ADF-Udemy

Search

yanpierr.donayre75@gmail.com
DIRECTORIO PREDETERMINADO

Data Factory | Validate all | Publish all

Factory Resources

Filter resources by name

- Pipeline 28
 - 1. Ingesta 1
 - 2. Move and Transform Activ... 2
 - 3. DataFlow 23
 - 4. General Activities 2
 - 1-Get Metadata
 - 2-Delete
- Dataset 9
 - 1. Ingesta 2
 - 2. Move and transform Activ... 4
 - 3. DataFlow 2
 - 4. General Activities 1
 - ADLS_GetMetadata
- Data flows 23
- Power Query 0

Activities

Search activities

- Move & transform
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
 - Append variable
 - Delete
 - Execute Pipeline
 - Execute SSIS package
 - Fail
 - Get Metadata
 - Lookup
 - Stored procedure

1-Get Metadata | ADLS_GetMetadata | 2-Delete

Validate | Debug | Add trigger

Delete

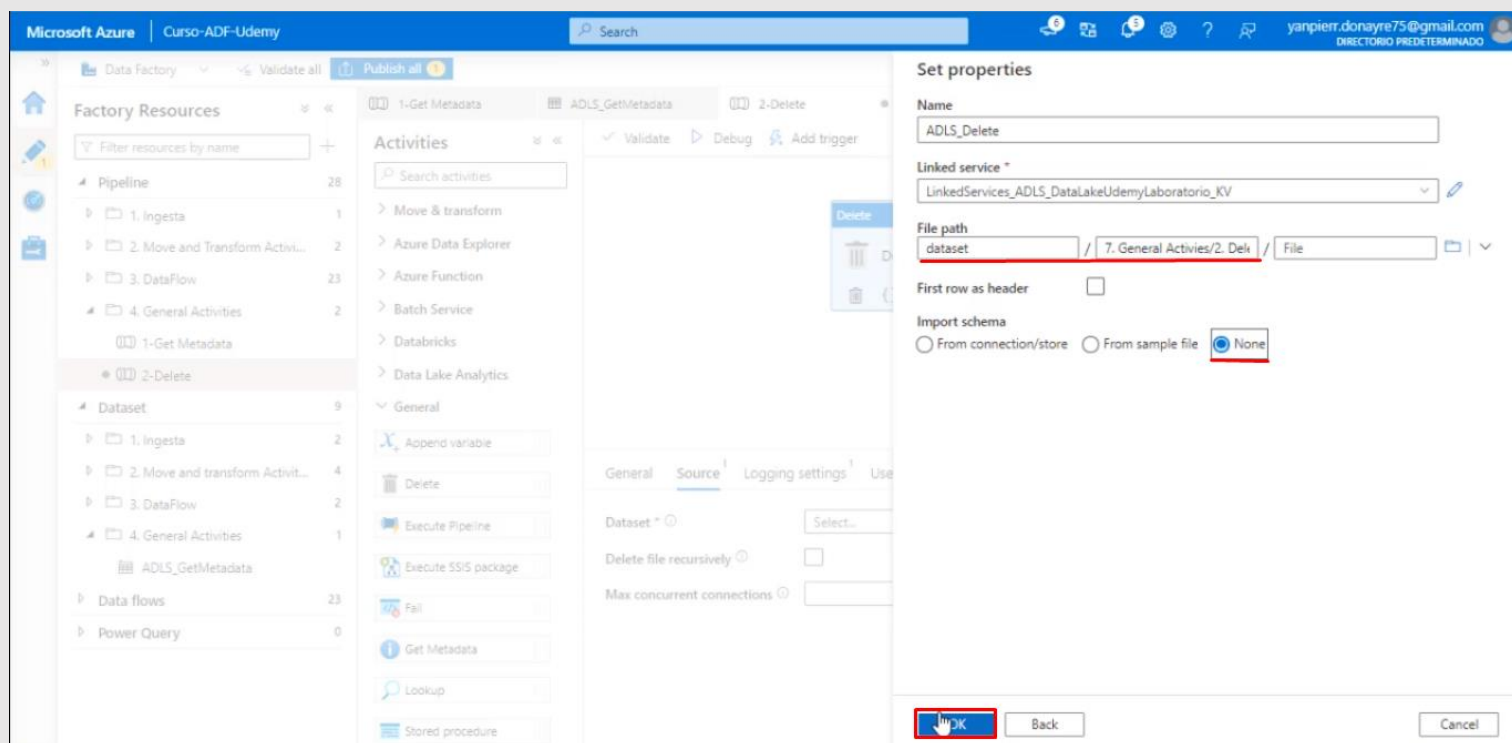
Select format

Choose the format type of your data

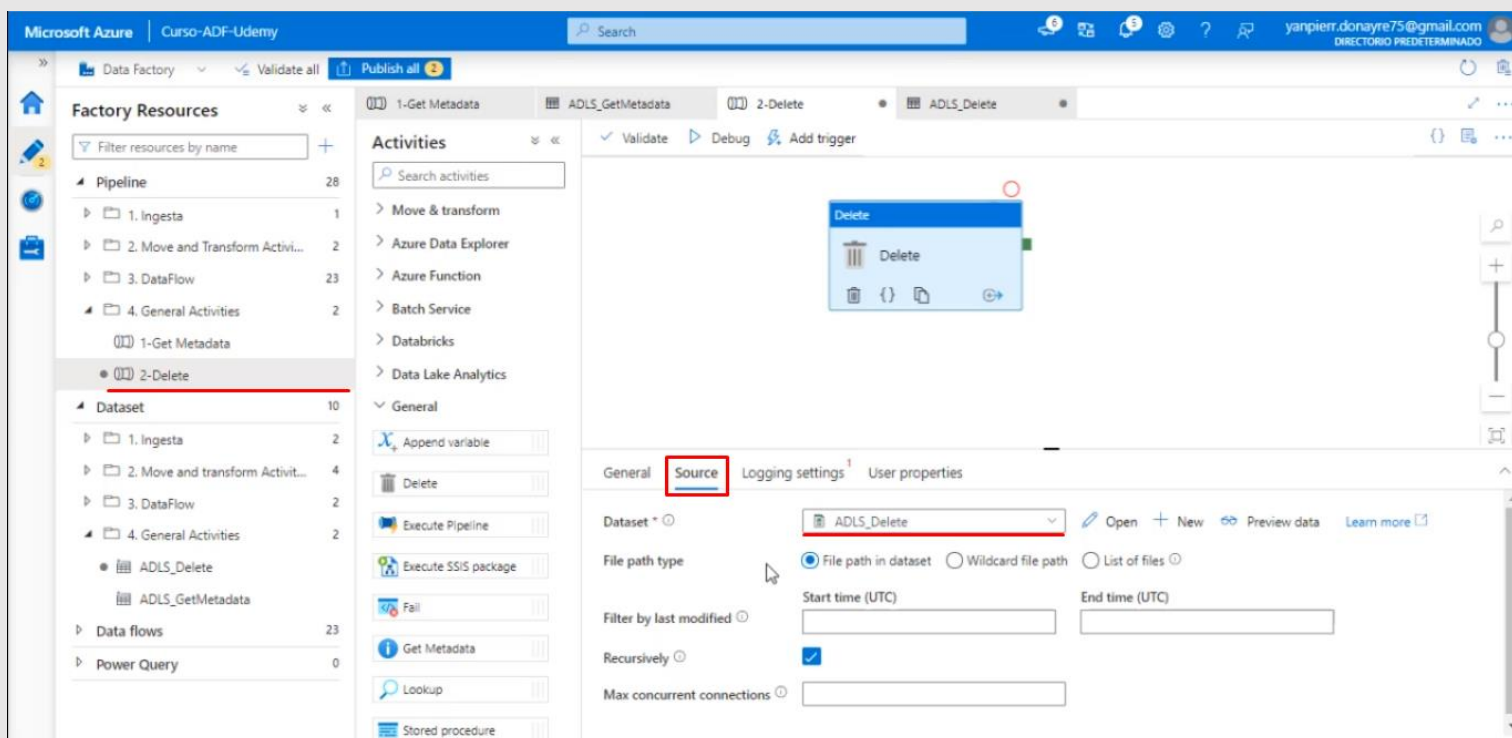
- Avro
- Binary
- DelimitedText**
- Excel
- JSON
- ORC
- Parquet
- XML

Continue | Back | Cancel

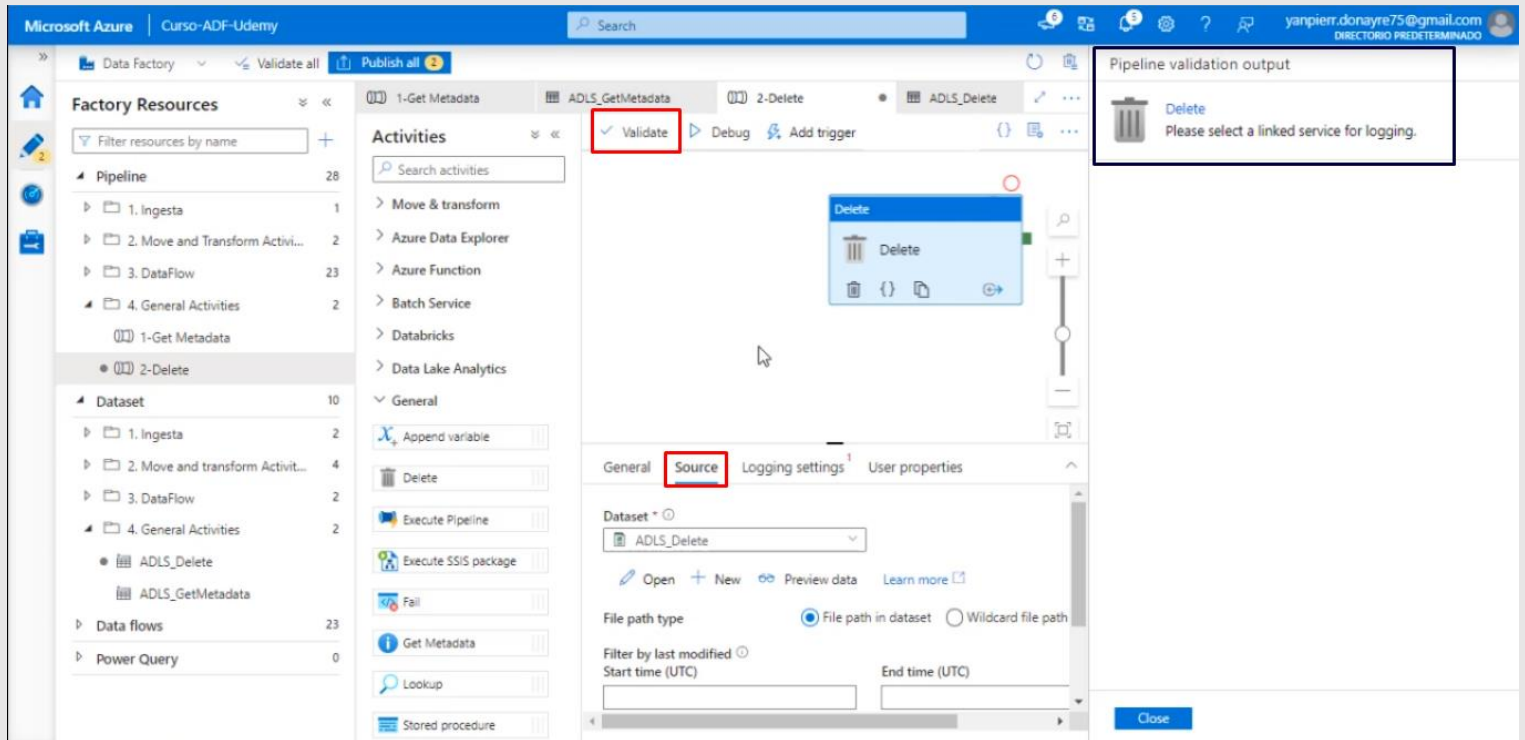
El **File path** apuntará a un directorio en específico, a la carpeta “.../2.Delete” (este directorio contiene el mismo material que el directorio **Delete** que se encuentra en la carpeta **material**). No importaremos el schema, dado que, estamos indicando en el File path una carpeta y no un archivo.



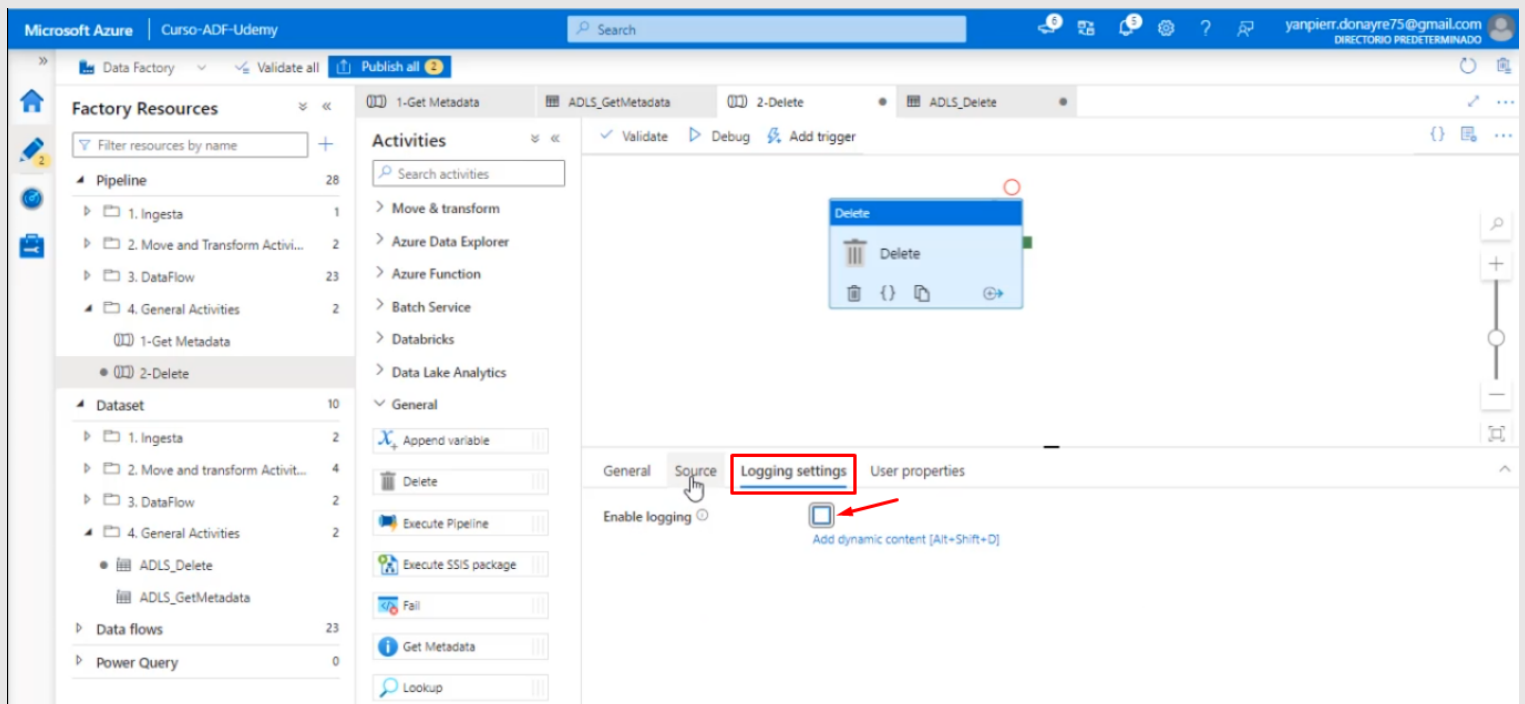
Ahora, si dejamos todo tal cual está y **Valido**



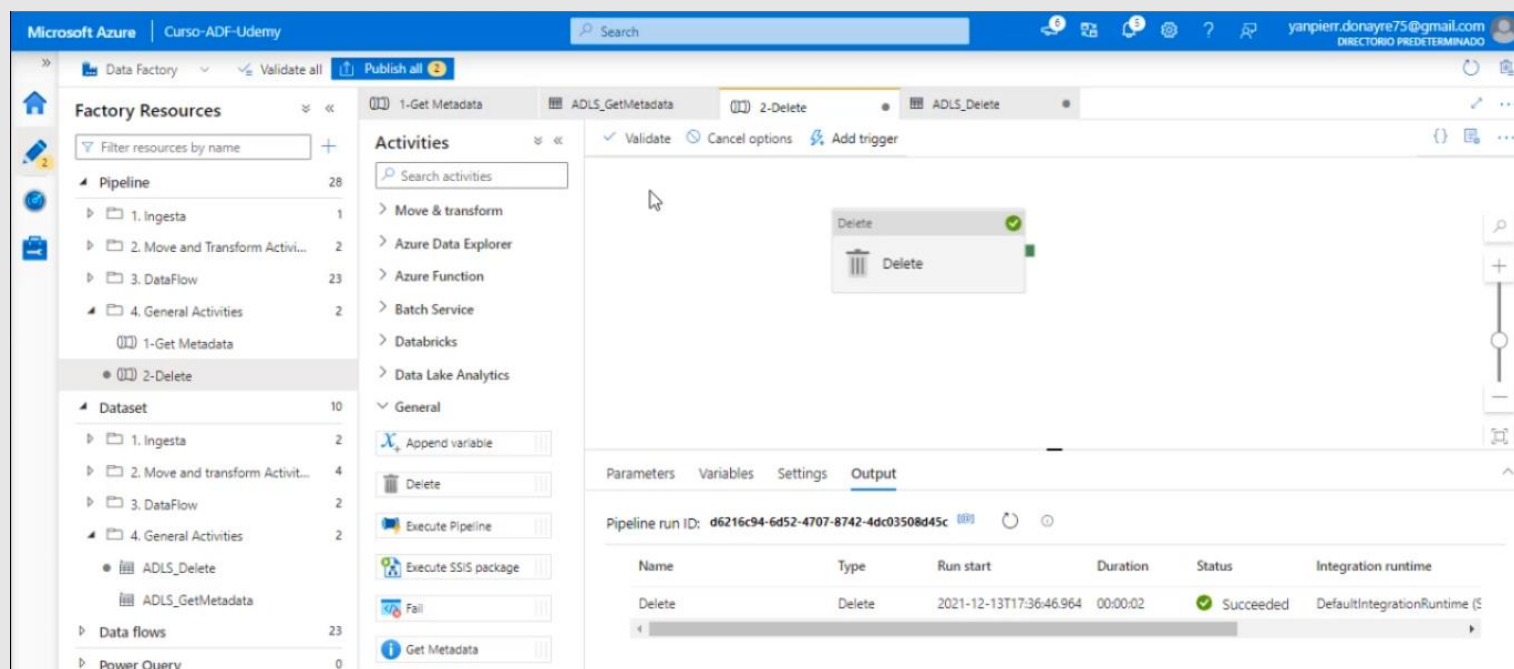
Nos aparece un mensaje pidiéndonos anexar un Linked Service para el Logging



Pero por el momento, deshabilitaremos la opción de obtener un Log



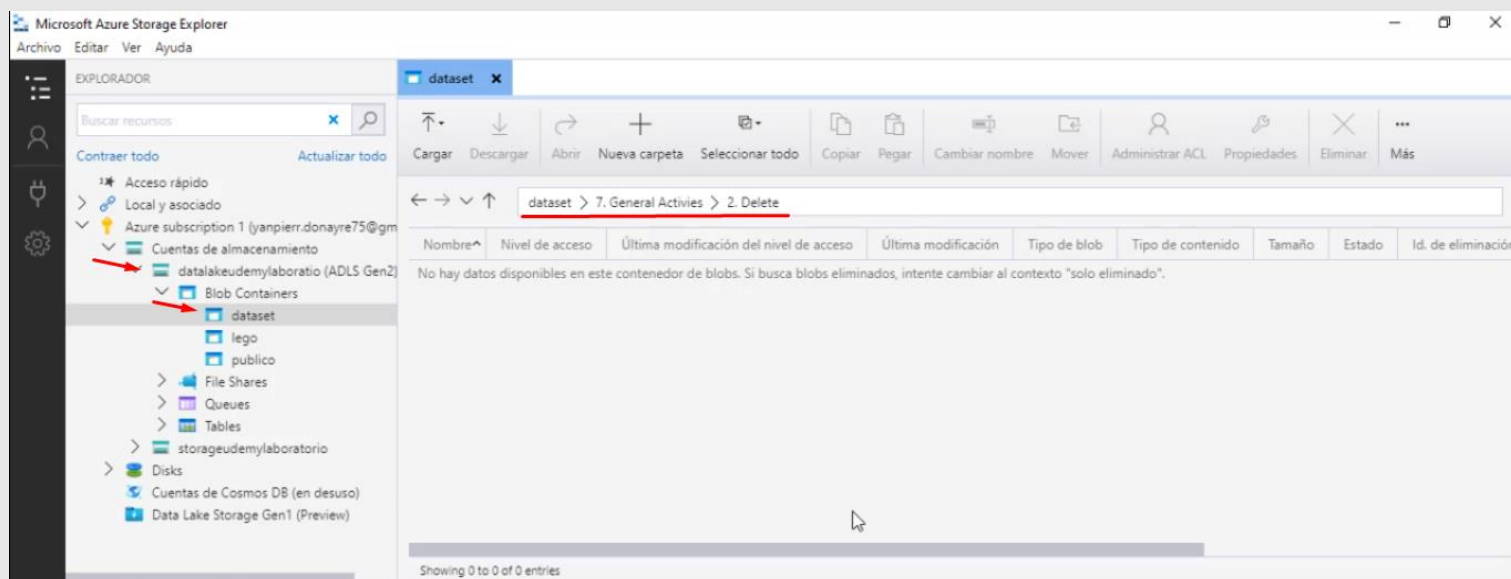
Volvemos a Validar y luego pulsamos Debug



The screenshot shows the Microsoft Azure Data Factory interface. In the 'Activities' pane, the 'Delete' activity is selected. A confirmation dialog box is displayed, asking to delete the activity. The 'Output' tab shows the pipeline run ID: d6216c94-6d52-4707-8742-4dc03508d45c. The table below shows the activity details:

Name	Type	Run start	Duration	Status	Integration runtime
Delete	Delete	2021-12-13T17:36:46.964	00:00:02	Succeeded	DefaultIntegrationRuntime (S

Si revisamos la carpeta “2.Delete” vamos a ver que se eliminaron todos los archivos y subcarpeta que existían. Si utilizamos la opción **Recursively** se eliminará todo, en el caso de que como dataset tengamos un directorio con archivos, se eliminará todo lo que hay dentro de ella.



The screenshot shows the Microsoft Azure Storage Explorer interface. The 'dataset' folder is selected in the 'Blob Containers' section. The folder is highlighted, and the 'dataset' folder is visible in the list. The interface shows the folder structure and the contents of the 'dataset' folder.

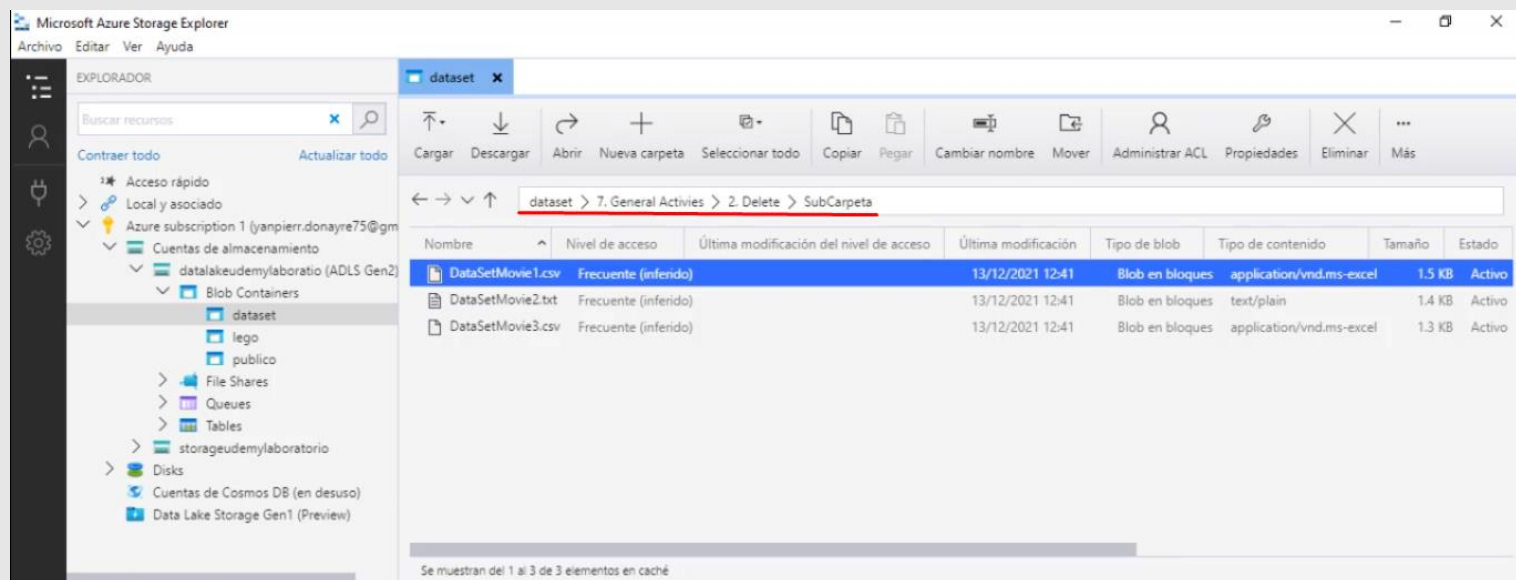
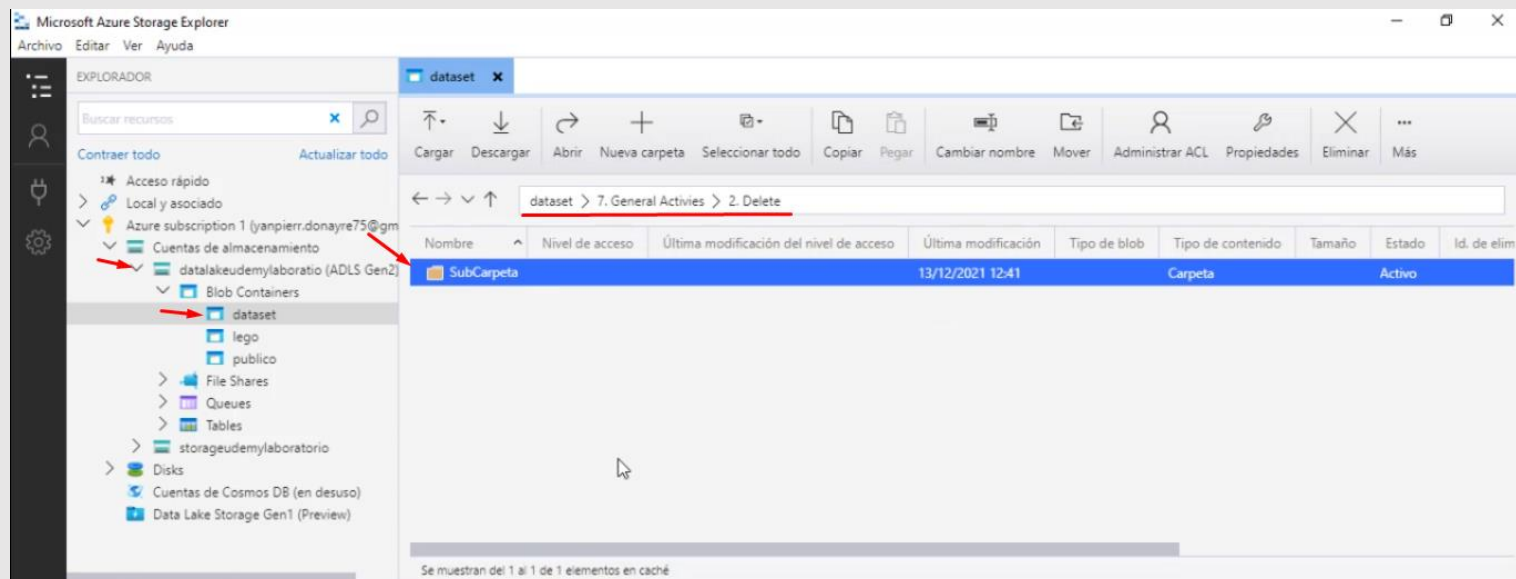
Vamos a volver a cargar la carpeta con sus archivos originales, pero esta vez vamos a **deshabilitar** la opción de **Recursively** y veamos que sucede

The screenshot shows the Microsoft Azure Data Factory portal interface. On the left, the 'Factory Resources' pane lists the pipeline '2-Delete' under the 'General Activities' folder. The main pane displays the 'Activities' list with 'Delete' selected. The 'Source' tab is active, showing the configuration for the 'Delete' activity. The 'Dataset' is set to 'ADLS_Delete'. The 'File path type' is 'File path in dataset'. The 'Filter by last modified' section has 'Start time (UTC)' and 'End time (UTC)' fields. The 'Recursively' checkbox is unchecked, with a red arrow pointing to it. The 'Max concurrent connections' field is empty.

The screenshot shows the Microsoft Azure Data Factory portal interface. On the left, the 'Factory Resources' pane lists the pipeline '2-Delete' under the 'General Activities' folder. The main pane displays the 'Activities' list with 'Delete' selected. The 'Output' tab is active, showing the results of the pipeline run. The 'Pipeline run ID' is '20978ce6-2369-4fda-8e78-b12895b18429'. The table below shows the results of the pipeline run.

Name	Type	Run start	Duration	Status	Integration runtime
Delete	Delete	2021-12-13T17:41:43.477	00:00:02	Succeeded	DefaultIntegrationRuntime (S

Si no activamos la casilla **Recursively** eliminará todo menos las subcarpetas y los archivos que se encuentren dentro de ella.



Vamos a volver a cargar la carpeta con sus archivos originales, pero esta vez vamos a utilizar una ruta de acceso con comodín. En esta ocasión, **únicamente se eliminarán los archivos con extensión .csv** que estén dentro del directorio, pero no dentro de las subcarpetas.

The screenshot shows the Microsoft Azure Data Factory console. On the left, the 'Factory Resources' pane shows a pipeline named '2-Delete' selected. The 'Activities' pane on the right shows the 'Delete' activity. The 'Source' tab is active, showing the configuration for the 'Delete' activity. The 'Dataset' dropdown is set to 'ADLS_Delete'. The 'File path type' is set to 'Wildcard file path'. The 'Wildcard file name' is set to '*.csv'. The 'Filter by last modified' section is empty. The 'Recursively' checkbox is unchecked. The 'Max concurrent connections' is set to 1. A 'Delete' dialog box is visible in the center of the screen.

The screenshot shows the Microsoft Azure Data Factory console. On the left, the 'Factory Resources' pane shows a pipeline named '2-Delete' selected. The 'Activities' pane on the right shows the 'Delete' activity. The 'Output' tab is active, showing the output of the 'Delete' activity. The 'Pipeline run ID' is 'a9a48587-bb7c-429d-b69e-aacfd0d5b482d'. The 'View debug run consumption' link is visible. The 'Output' table shows the following data:

Name	Type	Run start	Duration	Status	Integration runtime
Delete	Delete	2021-12-13T17:45:19.995	00:00:03	Succeeded	DefaultIntegrationRuntime (S

Microsoft Azure Storage Explorer

Archivo Editar Ver Ayuda

EXPLORADOR

Buscar recursos

Contraer todo Actualizar todo

Acceso rápido

Local y asociado

Azure subscription 1 (yanpierr.donayre75@gm)

Cuentas de almacenamiento

datalakeudemylaboratio (ADLS Gen2)

Blob Containers

dataset

lego

publico

File Shares

Queues

Tables

storageudemylaboratorio

Disks

Cuentas de Cosmos DB (en desuso)

Data Lake Storage Gen1 (Preview)

dataset

Cargar Descargar Abrir Nueva carpeta Seleccionar todo Copiar Pegar Cambiar nombre Mover Administrar ACL Propiedades Eliminar Más

dataset > 7. General Activies > 2. Delete

Nombre	Nivel de acceso	Última modificación del nivel de acceso	Última modificación	Tipo de blob	Tipo de contenido	Tamaño	Estado
SubCarpeta			13/12/2021 12:41		Carpeta		Activo
DataSetMovie2.txt	Frecuente (inferido)		13/12/2021 12:42	Blob en bloques	text/plain	1.4 KB	Activo
Lista.txt	Frecuente (inferido)		13/12/2021 12:42	Blob en bloques	text/plain	36 B	Activo

Se muestran del 1 al 3 de 3 elementos en caché

Microsoft Azure Storage Explorer

Archivo Editar Ver Ayuda

EXPLORADOR

Buscar recursos

Contraer todo Actualizar todo

Acceso rápido

Local y asociado

Azure subscription 1 (yanpierr.donayre75@gm)

Cuentas de almacenamiento

datalakeudemylaboratio (ADLS Gen2)

Blob Containers

dataset

lego

publico

File Shares

Queues

Tables

storageudemylaboratorio

Disks

Cuentas de Cosmos DB (en desuso)

Data Lake Storage Gen1 (Preview)

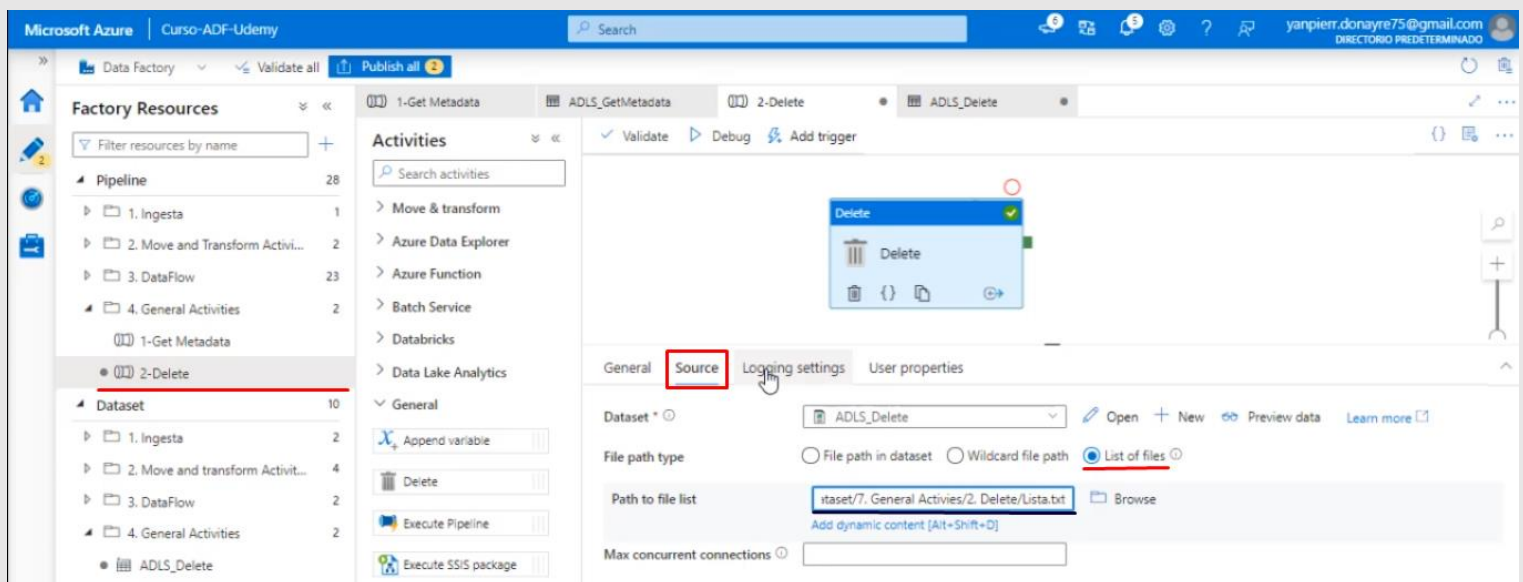
dataset

Cargar Descargar Abrir Nueva carpeta Seleccionar todo Copiar Pegar Cambiar nombre Mover Administrar ACL Propiedades Eliminar Más

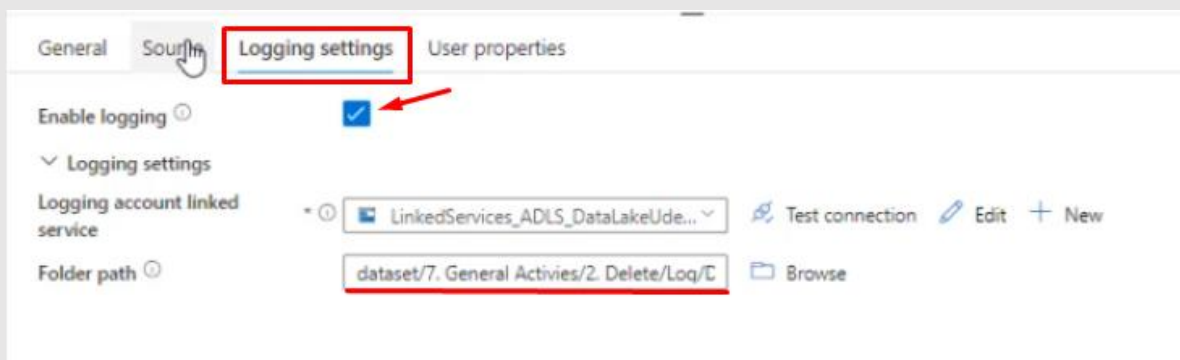
dataset > 7. General Activies > 2. Delete > SubCarpeta

Nombre	Nivel de acceso	Última modificación del nivel de acceso	Última modificación	Tipo de blob	Tipo de contenido	Tamaño	Estado
DataSetMovie1.csv	Frecuente (inferido)		13/12/2021 12:42	Blob en bloques	application/vnd.ms-excel	1.5 KB	Activo
DataSetMovie2.txt	Frecuente (inferido)		13/12/2021 12:42	Blob en bloques	text/plain	1.4 KB	Activo
DataSetMovie3.csv	Frecuente (inferido)		13/12/2021 12:42	Blob en bloques	application/vnd.ms-excel	1.3 KB	Activo

Vamos a volver a cargar la carpeta con sus archivos originales, pero esta vez vamos a **eliminar archivos de acuerdo a los elementos que aparecen en un archivo txt**. Para ello vamos a utilizar el archivo **lista.txt**.



Se utiliza el Linked Service que apunta hacia ADLS e indicaremos la ruta donde se almacenará el Log que se genere, en el directorio **.../2.Delete/Log/Delete_Log**



Podemos ver que se eliminaron los archivos indicados en el archivo **Lista.txt** y se creó el directorio **Log** y dentro de ella subcarpetas donde se encontrará el archivo Log.

