

# Laboratorio II: Creación de Pipelines en Azure Data Factory

## Introducción

¡Automatiza el flujo de tus datos con Azure Data Factory! Este laboratorio te guía en la creación de un pipeline para ingerir datos desde una **LandingZone** hacia la capa **Bronze**, usando archivos como **Athletes.csv**. Aprenderás a configurar conexiones, mover datos y habilitar el registro para un proceso ETL robusto. Es ideal para dominar la ingesta de datos estructurados. ¡Lleva tus datos al siguiente nivel!

### Paso 1: Ingresamos a Azure Synapse Analytics

1. Primeramente, ingresamos a la plataforma de Microsoft Azure

Buscar recursos, servicios y documentos (G+/)

Copilot

Servicios de Azure

- Crear un recurso
- Azure Synapse Analytics**
- Factorías de datos
- Suscripciones
- Cuentas de almacenamiento
- Grupos de recursos
- Máquinas virtuales
- Azure Lab Services
- Centro de inicio rápido
- Más servicios

Recursos

Reciente Favorito

Nombre	Tipo	Última consulta
<b>syntokyoolympicsgps</b>	Synapse workspace	hace 37 minutos
Azure for Students	Suscripción	hace 15 horas
adftokyoolympicsgps	Factoría de datos (V2)	hace 17 horas
tokio-olympics-gps	Grupo de recursos	hace 17 horas
tokioolympicsgps	Cuenta de almacenamiento	hace 18 horas

Ver todo

2. Ahora ingresamos a la instancia de **Azure Synapse Analytics**.

Microsoft Azure

Buscar recursos, servicios y documentos (G+/)

Copilot

syntokyoolympicsgps

Synapse workspace

Buscar

Información esencial

- Nuevo grupo de SQL dedicado
- Nuevo grupo de Apache Spark
- Actualizar
- Restablecer contraseña de administrador de SQL
- Eliminar

Información general

- Registro de actividad
- Control de acceso (IAM)
- Etiquetas
- Diagnosticar y solucionar problemas
- Visualizador de recursos
- Configuración
- Grupos de análisis
- Seguridad
- Supervisión
- Automation
- Ayuda

Información esencial

Grupo de recursos: [tokio-olympics-gps](#)

Estado: Succeeded

Ubicación: East US 2

Suscripción: [Azure for Students](#)

Id. de suscripción: 941c6b7-cfb5-4898-8095-7eeae75a3e43

Red virtual administrada: No

Id. del objeto de identidad: a2e1f6c-9408-41a7-4938-6f902c7f95c

Dirección URL web del id.: <https://web.azure.synapse.net/te/worksapace/57941c6b7-cfb5-4898-8095-7eeae75a3e43>

Etiquetas: [Azure for Students](#)

Redes: [Mostrar configuración del firewall](#)

Dirección URL de la cuenta: <https://tokioolympicsgps.dfs.core.windows.net>

Sistema de archivos de A.: [dataazkioolympicsgps](#)

Nombre de usuario del a.: rgadminuser

Administrador de Micros.: [rgadmin@synapse.net](mailto:rgadmin@synapse.net)

Punto de conexión de S.: [syntokyoolympicsgps.sql.azure.synapse.net](#)

Punto de conexión de S.: [syntokyoolympicsgps-ondemand.sql.azure.synapse.net](#)

Punto de conexión de de.: <http://syntokyoolympicsgps.dev.azure.synapse.net>

Vista JSON

Abrir Synapse Studio

Empiece a crear su solución de análisis completamente integrada y descubra información nueva.

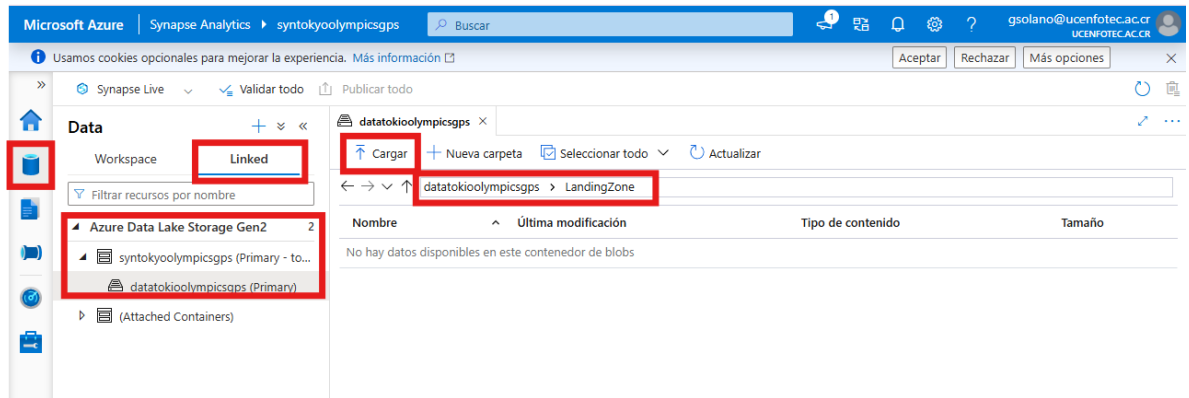
Abrir

Lectura de documentación

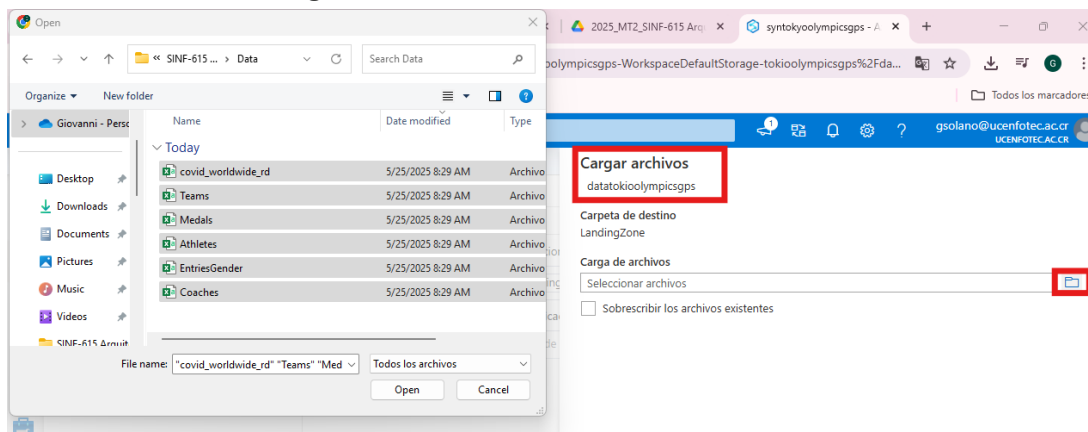
Obtenga información sobre cómo ser productivo rápidamente. Explore los conceptos, los tutoriales y los ejemplos.

Más información

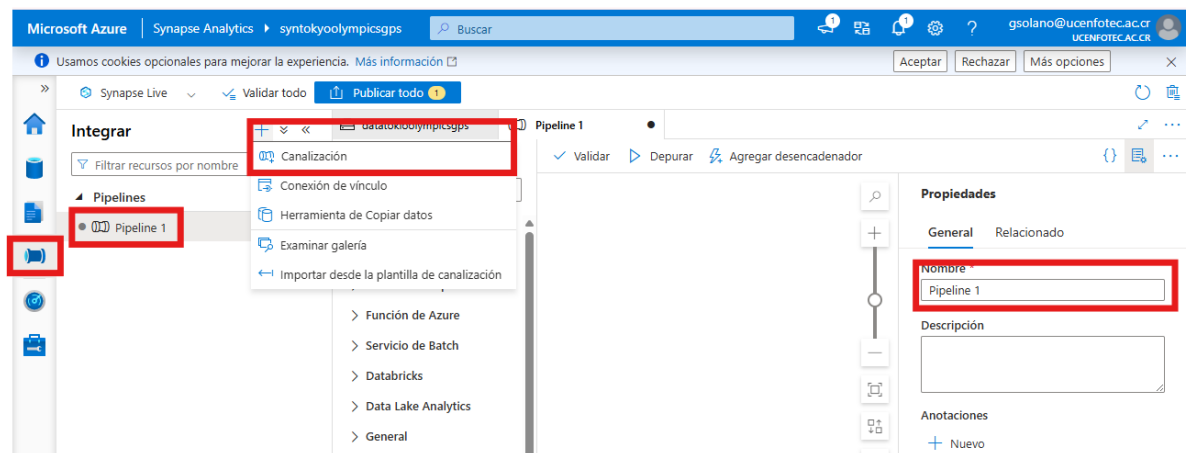
3. Procedemos a cargar los archivos que se encuentran en el aula virtual utilizando el siguiente flujo, clic en el **cilindro de datos**, clic en la pestaña **Linked**, clic en storage container (Data lake), clic en la carpeta **LandingZone** y procedemos a cargar los 6 archivos localizados en la segunda semana como Data.



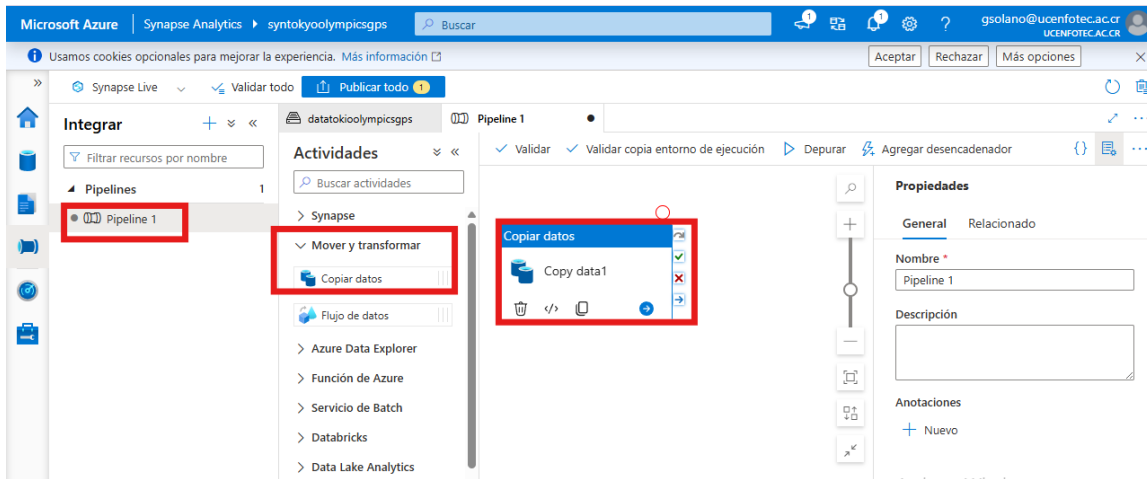
4. Procedemos a cargar los archivos



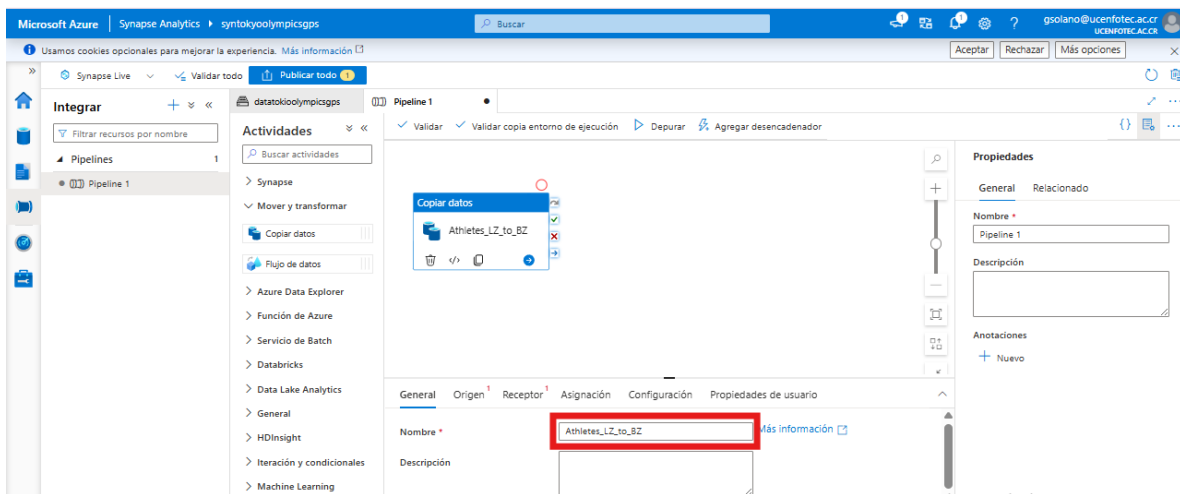
5. Procedemos a crear pipeline.



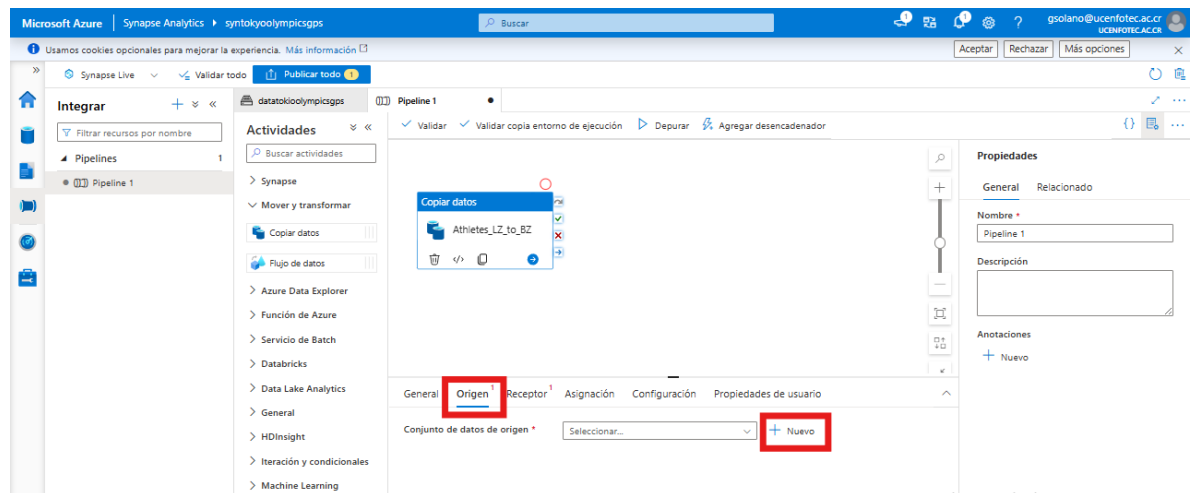
6. Seleccionamos nuestro **pipeline 1** y arrastramos **copiar datos** al lienzo.



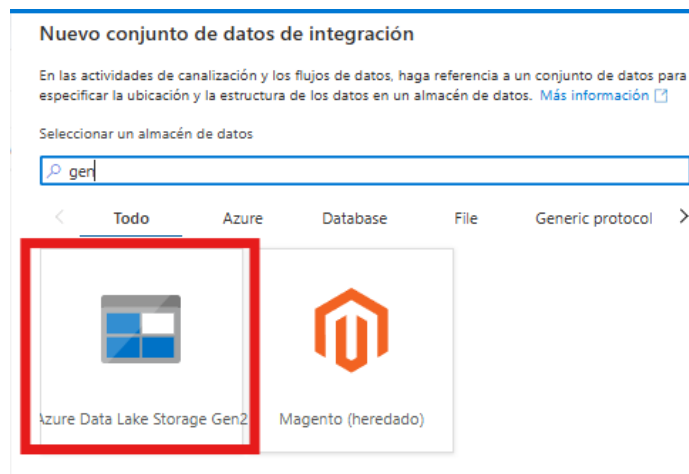
7. Seleccionamos nuestro **copiar datos** y le cambiamos el nombre **Athletes\_LZ\_to\_BZ**



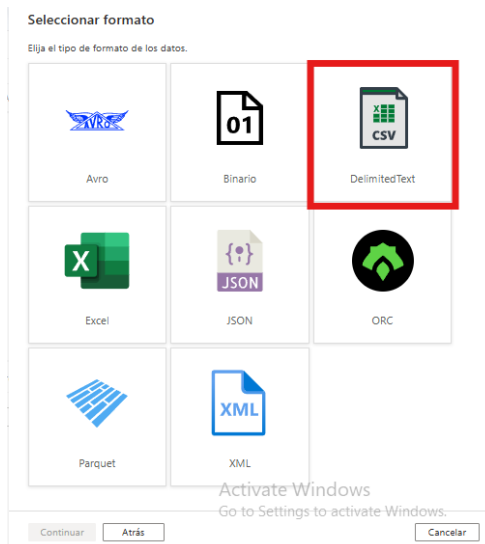
8. Nos dirigimos a la carpeta de **Origen** y le damos a **Nuevo**



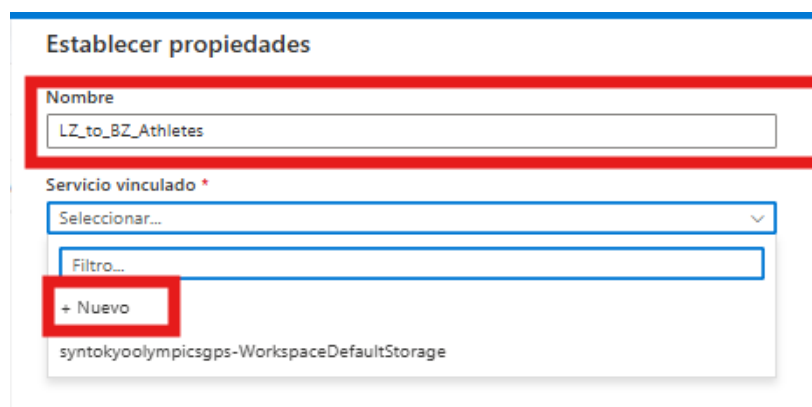
9. Seleccionamos **Azure Data Lake Storage Gen 2** y le damos continuar.



10. Luego seleccionamos el CSV



11. Le asignamos un nombre a la conexión **LZ\_to\_BZ\_Athletes** y creamos un **Nuevo** vínculo



12. Creamos un nuevo servicio con el nombre de **ADLSext**, seleccionamos el **azure for students** en la suscripción y luego **tokioolympicsgps** como la cuenta de almacenamiento, le damos **prueba de conexión** y por último **crear**.

Nuevo servicio vinculado

Azure Data Lake Storage Gen2 Más información

Elija un nombre para el elemento servicio vinculado. Este nombre no se puede actualizar más adelante.

Nombre \*

ADLSext

Descripción

Conectar mediante Integration Runtime \*

AutoResolveIntegrationRuntime

Tipo de autenticación

Clave de cuenta

Método de selección de cuenta

From Azure subscription Enter manually

Suscripción de Azure

Azure for Students (941cafb7-dfb5-4898-80f5-7eeee75a3e43)

Nombre de cuenta de almacenamiento \*

tokioolympicsgps

Prueba de conexión

Conexión correcta

Crear Cancelar

13. Por último, asignamos el nombre a la propiedad **extLZ\_Athletes**, le damos clic a la carpeta y seleccionamos el archivo **Athletes.csv**, nos aseguramos que la función **Primera fila como encabezado** y le damos **aceptar**

Establecer propiedades

Nombre

extLZ\_Athletes

Servicio vinculado \*

ADLSext

Conectar mediante Integration Runtime \*

AutoResolveIntegrationRuntime

Ruta de acceso del archivo

datatokioolympicsgps / LandingZone / Athletes.csv

Primera fila como encabezado

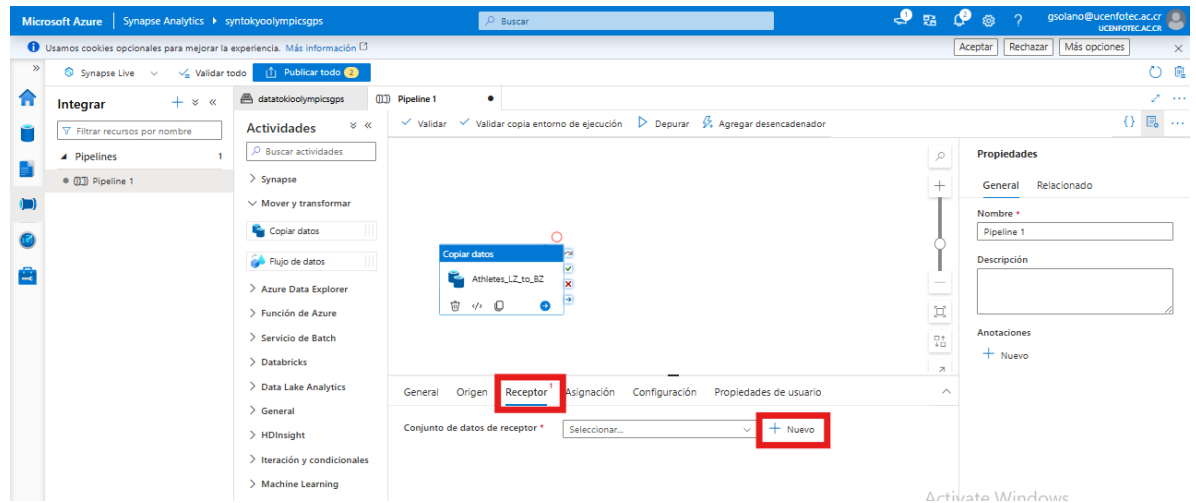
Importar esquema

Desde una conexión o un almacén Desde un archivo de ejemplo Ninguno

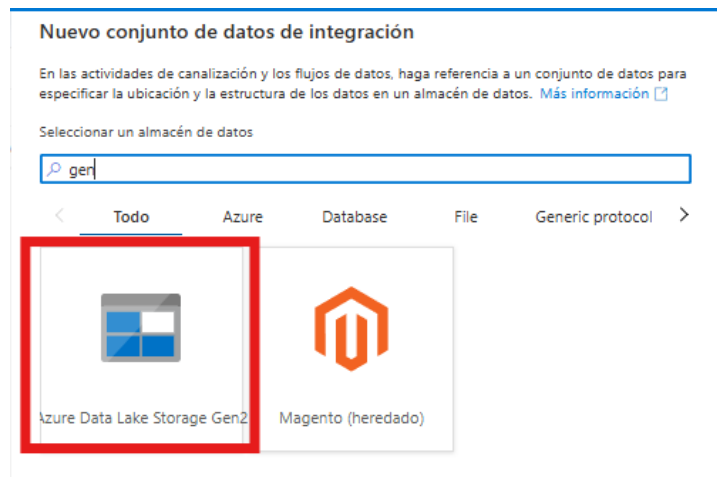
Avanzadas

Aceptar Atrás Cancelar

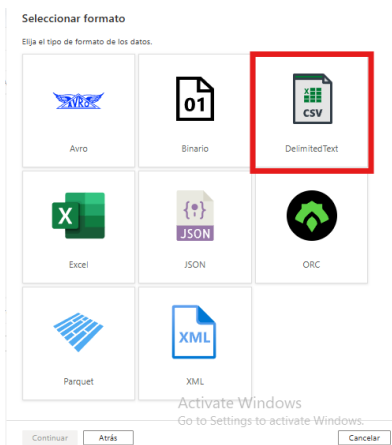
14. Creación del destino (**Receptor**), seleccionamos la pestaña de **receptor** y creamos una **nueva conexión**.



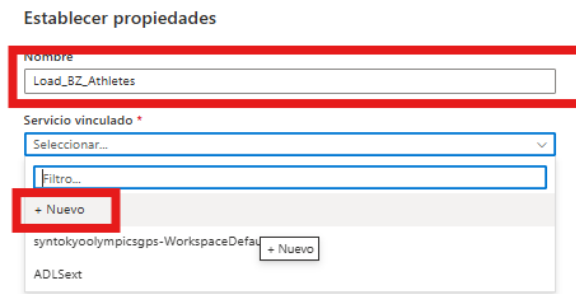
15. Seleccionamos **Azure Data Lake Storage Gen 2** y le damos continuar.



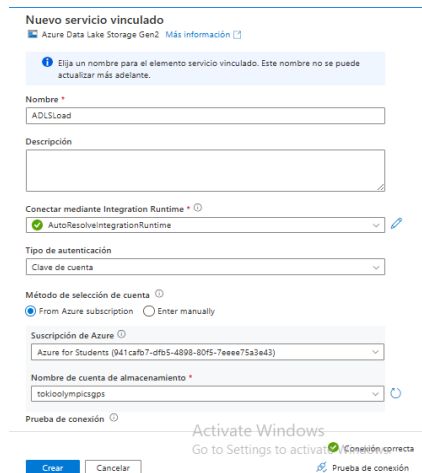
16. Luego seleccionamos el CSV



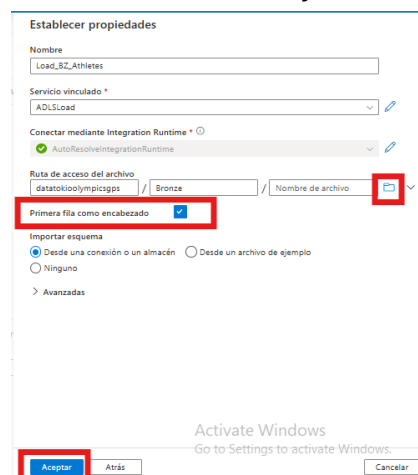
17. Le asignamos el nombre **Load\_BZ\_Athletes** y creamos un **Nuevo**



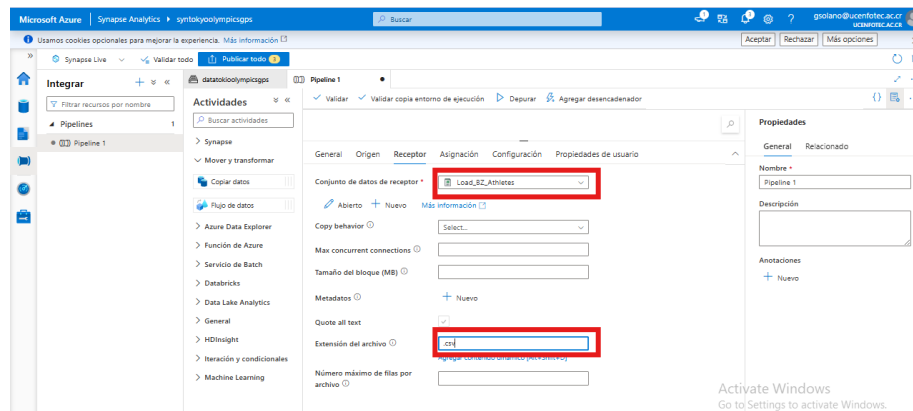
18. le asignamos el nombre **ADLSLoad**, seleccionamos el **azure for students** en la suscripción y luego **tokioolympicsgps** como la cuenta de almacenamiento, le damos **prueba de conexión** y por último **crear**.



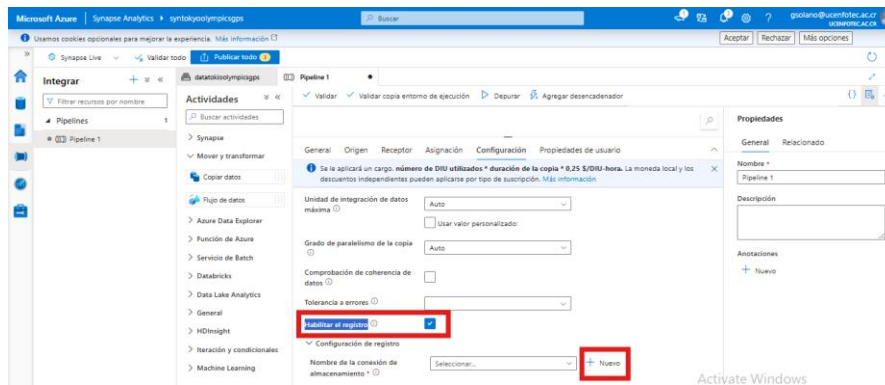
19. asignamos el nombre a la propiedad **Load\_BZ\_Athletes**, le damos clic a la carpeta y seleccionamos la carpeta marcada en rojo, además buscamos **datatokyoolympicsgps/Bronze/Athletes\_LZ.csv**, nos aseguramos de que la función **Primera fila como encabezado** y le damos **aceptar**



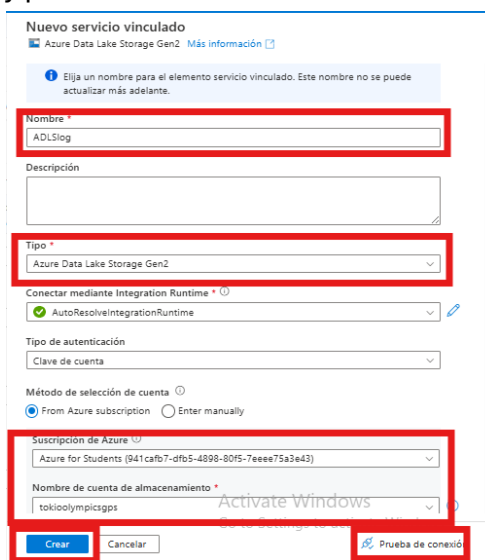
20. luego cambiamos la **Extension del archivo a .csv**



21. además, nos dirigiremos a la pestaña de **configuración** para **Habilitar el registro** y procedemos a crea un **Nuevo data storage**.

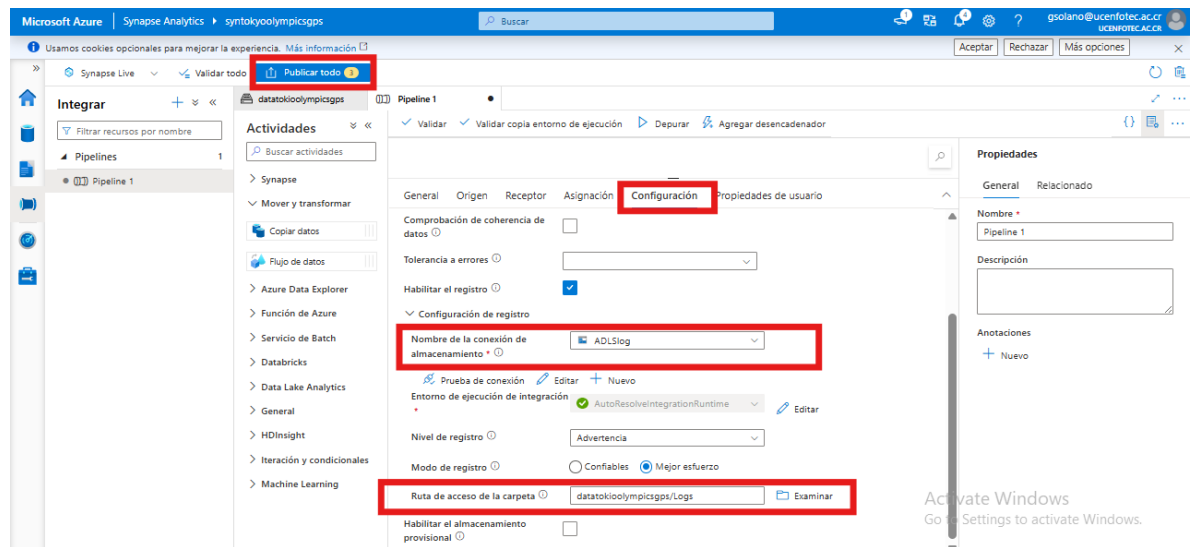


22. le asignamos el nombre de **ADLSlog** seleccionamos el tipo **Azure Data Lake Storage Gen 2**, seleccionamos el **azure for students** en la suscripción y luego **tokioolympicsgps** como la cuenta de almacenamiento, le damos **prueba de conexión** y por último **crear**.

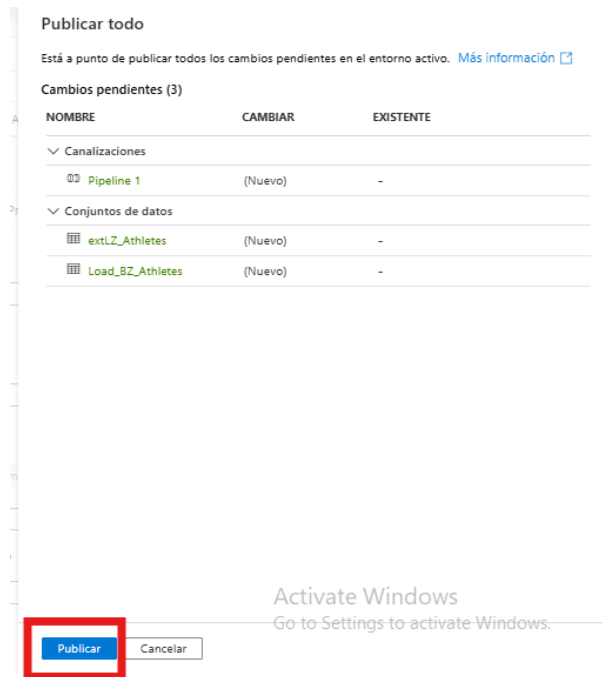




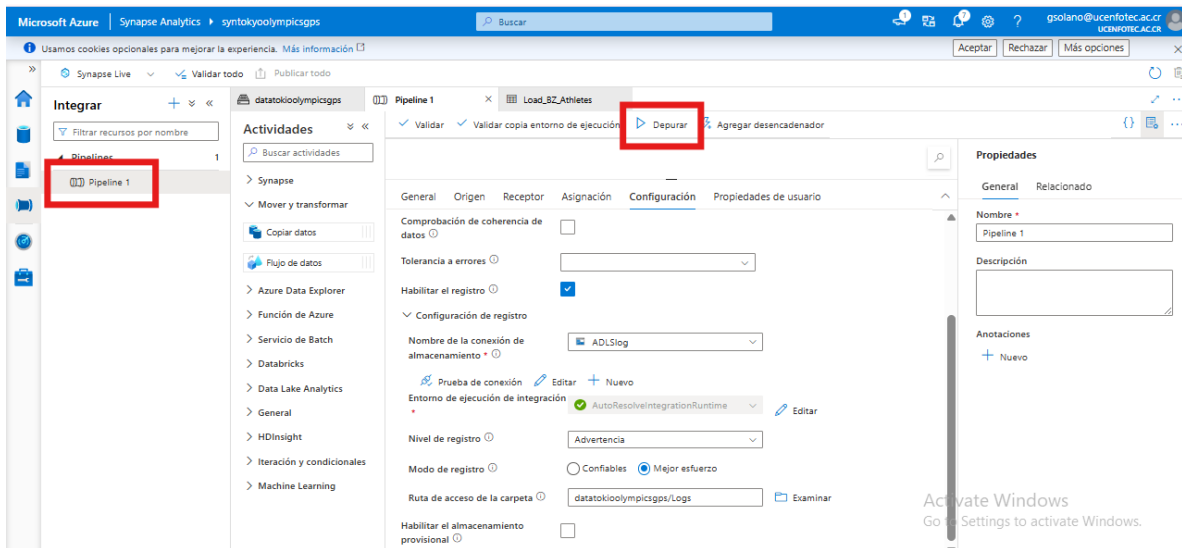
23. Una vez creado la conexión de almacenamiento procedemos a **examinar** la **Ruta de acceso de la carpeta a datatokiolympicsgps/Logs**, por ultimo procedemos a **Publicarlo todo**.



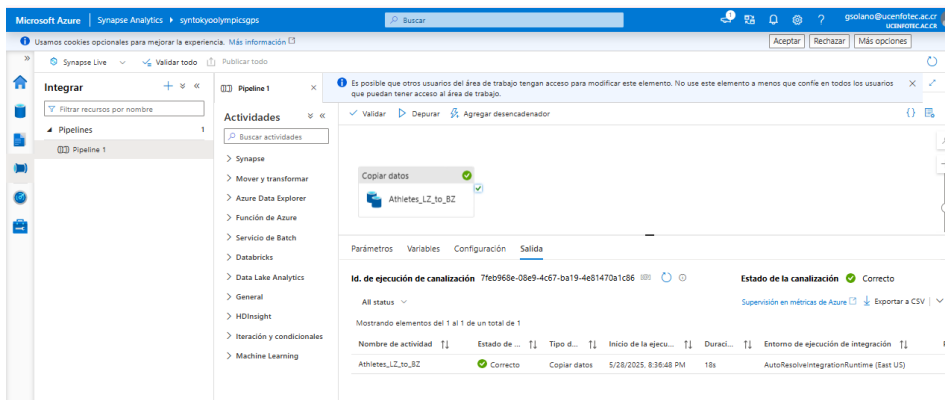
24. Por último, de damos al botón de **publicar**.



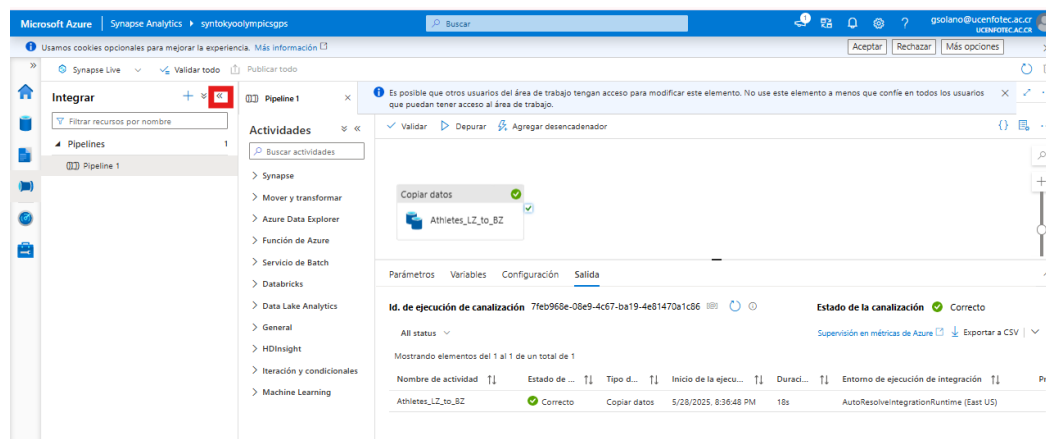
25. Luego, seleccionamos el **pipeline 1** y le damos clic al botón **Depurar**.



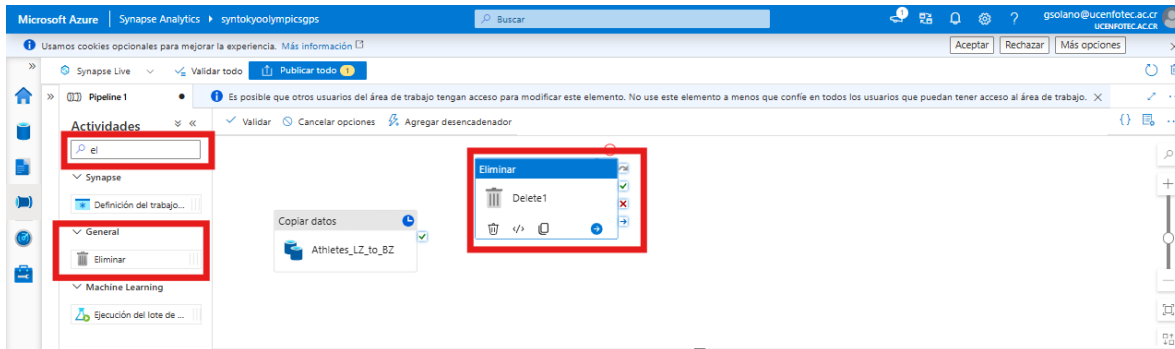
26. Al correr correctamente veremos los siguientes outputs



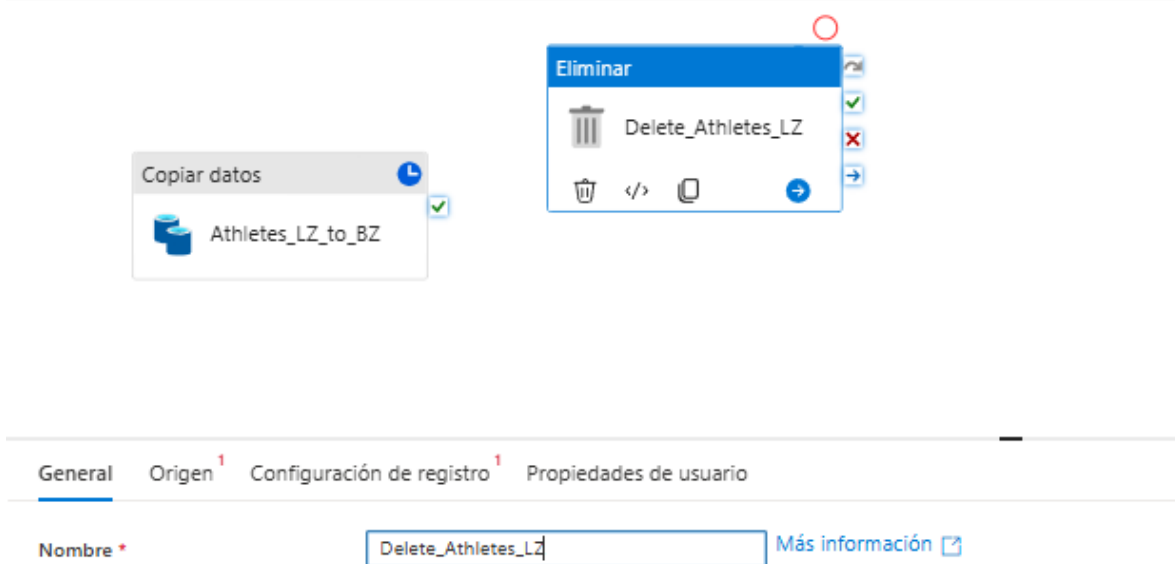
27. Luego le damos clic al botón <<



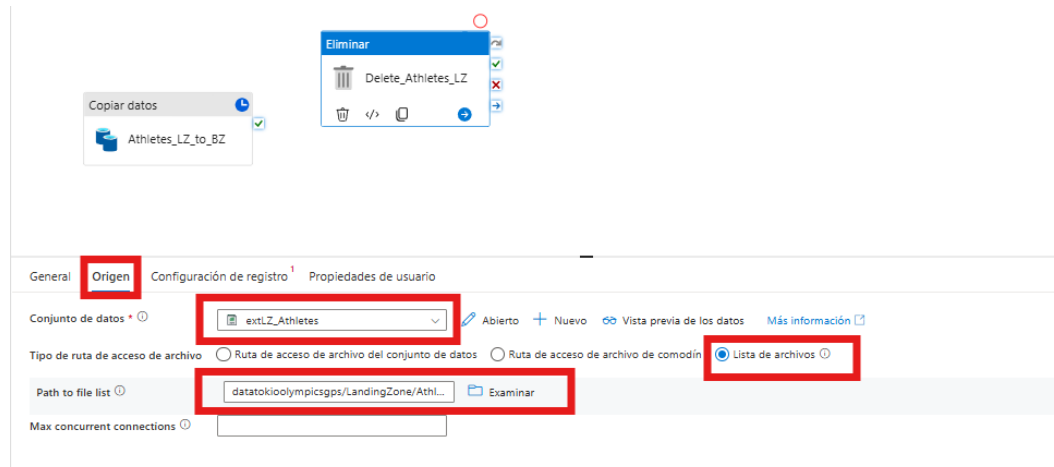
28. Después buscamos **eliminar** y lo agregamos al lienzo



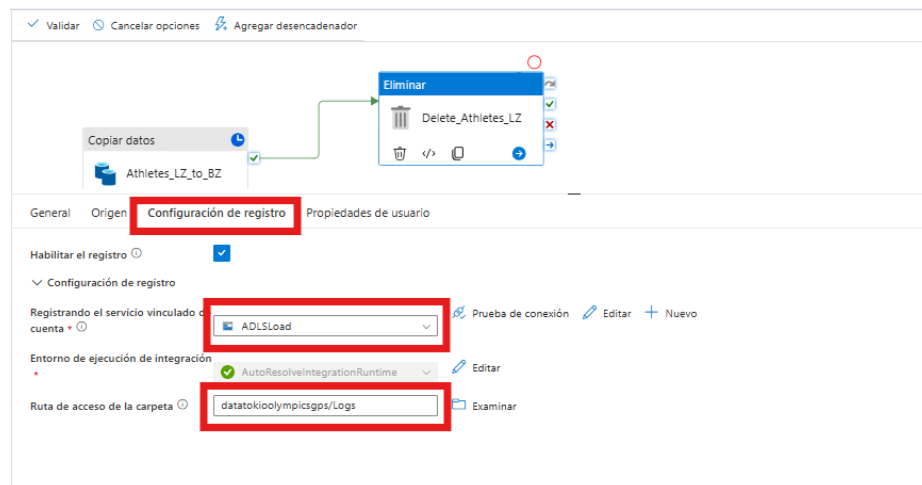
29. Luego editamos el nombre Delete\_Athletes\_LZ



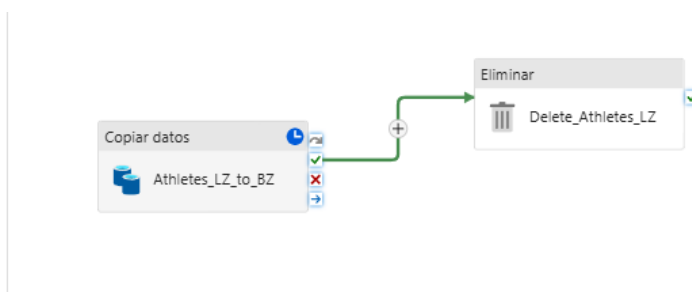
30. Procedemos a movernos a la pestaña de **Origen** seleccionamos el conjunto de datos **extLZ\_Atheletes** luego damos clic a **lista de archivos** y buscamos el path **datatokioolympicsgps/LandingZone/Athletes.csv**



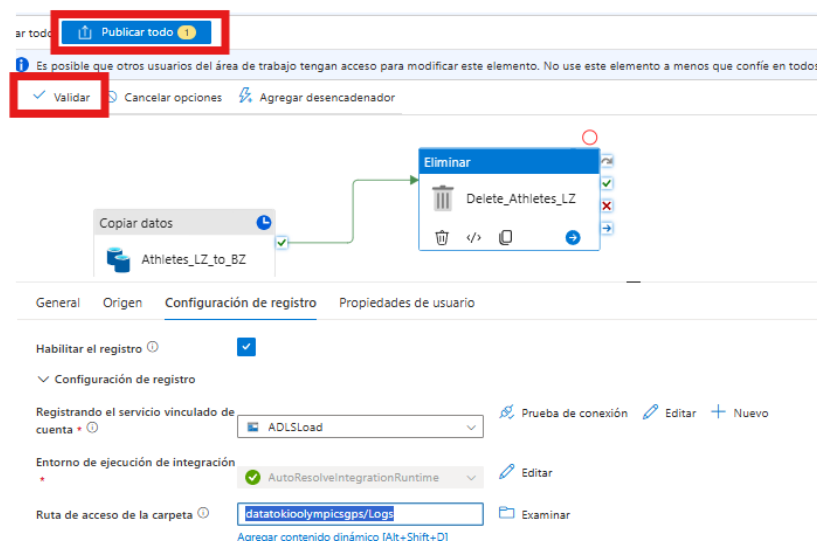
31. Nos movemos a la pestaña de **configuración de registro** y seleccionamos el **ADLSLoad** y la ruta de acceso de la carpeta **datatokiolympicsgps/Logs**



32. Anclamos el **copiar datos a eliminar**, cuando el proceso de copiar datos sea exitoso.



33. Validamos y procedemos a **publicar**



**YOU DID IT!**  
CONGRATULATIONS!