



Lab 2

Marzo 2024



Power Platform
Bootcamp

Microsoft Fabric

Trabajando con Datos de 0 a 100

Power Platform Madrid



Contenido

Estructura del Documento.....	2
Introducción.....	2
Ingestar Datos con DataFlow Gen 2.....	3
Tarea 1: Cargar Datos desde un Archivo CSV	3
Tarea 2: Verificar los Datos cargados con SQL.....	7

Estructura del Documento

El laboratorio incluye pasos que el usuario debe seguir, junto con capturas de pantalla asociadas que proporcionan ayuda visual. En cada captura de pantalla, las secciones se resaltan con cuadros naranjas para indicar las áreas en las que el usuario debe centrarse.

Introducción

En este laboratorio se trabajará sobre varias características clave de Microsoft Fabric. Se trata de un taller introductorio destinado a presentarte las distintas experiencias de productos y artefactos disponibles en Fabric. Al final de este taller, aprenderá a usar **Lakehouse**, **Datawarehouse**, **Dataflow Gen2**, **Data Pipeline** y la función **DirectLake**.

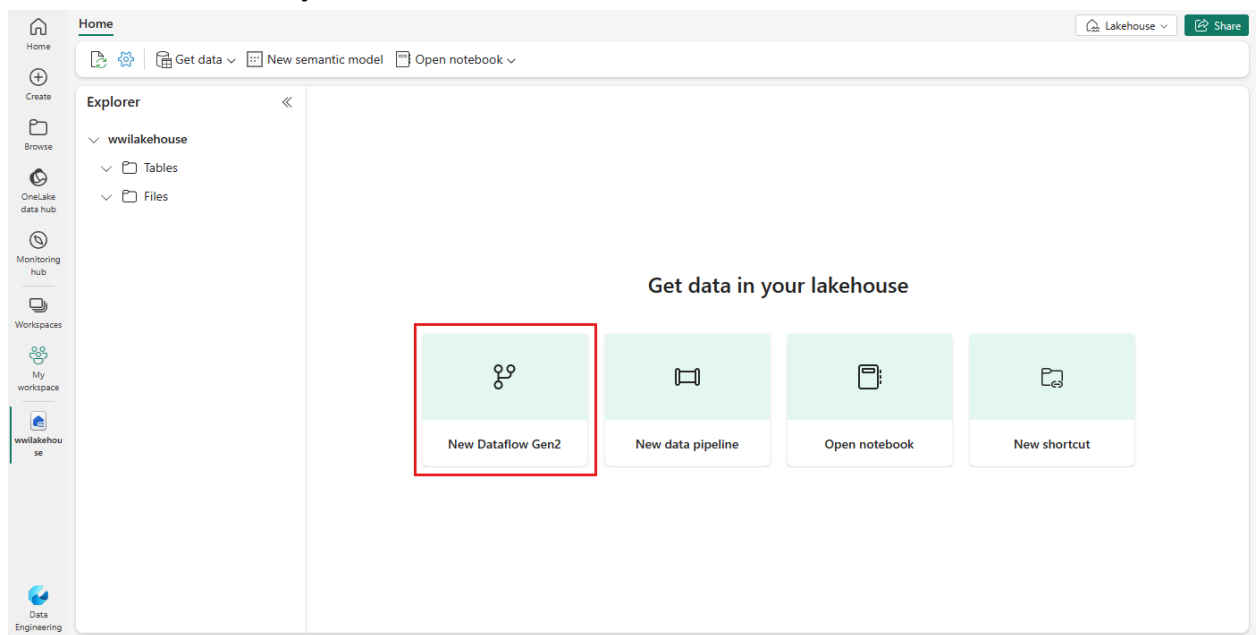
Al final de este laboratorio, se habrá aprendido:

1. Cómo crear un DataFlow Gen 2 en **Fabric**
2. Cómo Refrescar un Dataflow
3. Cómo consultar los datos ingestados

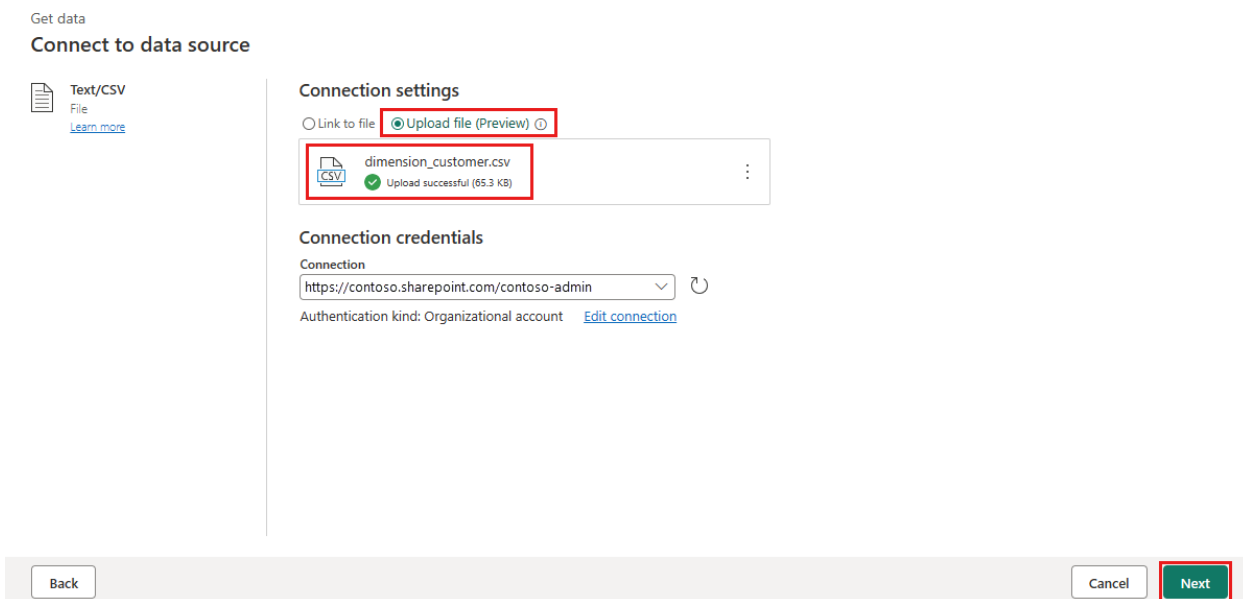
Ingestar Datos con DataFlow Gen 2

Tarea 1: Cargar Datos desde un Archivo CSV

1. Descargue el archivo *dimension_customer.csv* del [repositorio de ejemplos de Fabric](#).
2. En el **explorador de Lago de datos**, verá opciones para cargar datos en el lago.
Seleccione **Nuevo Flujo de datos Gen2**.

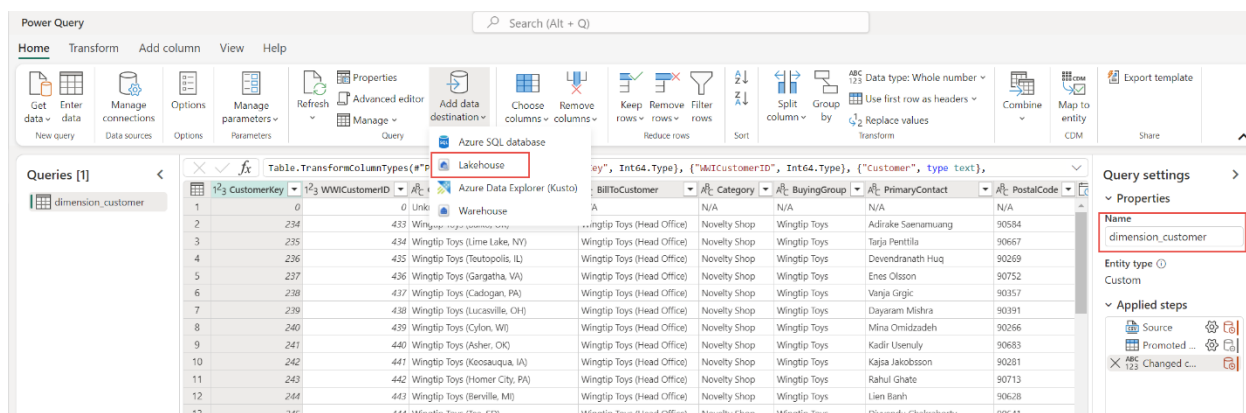


3. En el nuevo panel de flujo de datos, seleccione **Importar desde un archivo Text/CSV**.
4. En el panel **Conectar al origen de datos**, seleccione el botón de radio **Cargar archivo**. Arrastre y coloque el archivo *dimension_customer.csv* que descargó en el paso 1. Una vez cargado el archivo, seleccione **Siguiente**.



5. En la página **Vista previa de los datos del archivo**, obtenga una vista previa de los datos y seleccione **Crear** para continuar y volver al lienzo del flujo de datos.
6. En el panel **Configuración de consulta**, actualice el campo **Nombre** a **dim_customers**.

Nota: Fabric agrega un espacio y un número al final del nombre de la tabla de forma predeterminada. Los nombres de tabla deben estar en minúsculas y no deben contener espacios. Cambie el nombre de forma adecuada y quite los espacios del nombre de la tabla.



7. En este tutorial, ya ha asociado los datos del cliente a una instancia de Lakehouse. En caso de que tenga otros elementos de datos que desee asociar al Lake, puede agregarlos con los pasos siguientes:
 - En los elementos de menú, seleccione **Agregar destino de datos** y seleccione **Lakehouse**. En la pantalla **Conectar al destino de datos**, inicie sesión en su cuenta si es necesario y seleccione **Siguiente**.
 - Seleccione el Lakehouse existente en el área de trabajo.

- Si la tabla **dim_customers** no existe, seleccione la opción **Nueva tabla** y escriba el nombre de la tabla **dim_customers**. Si la tabla ya existe, seleccione la opción **Tabla existente** y elija **dim_customers** en la lista de tablas del explorador de objetos. Seleccione **Siguiente**.

Choose destination target

☒ New table ☐ Existing table

Search

Display options ▾

- My workspace
- wwilakehouse

A new table will be created in wwilakehouse

Table name *

dimension_customer

Back Cancel Next

- En el panel **Elegir configuración de destino**, seleccione **Reemplazar** como **método de Actualización**. Seleccione **Guardar configuración** para volver al lienzo del flujo de datos.
8. Desde el lienzo del flujo de datos, puede transformar fácilmente los datos en función de sus requisitos empresariales. Por motivos de simplicidad, no estamos realizando ningún cambio en este tutorial. Para continuar, seleccione **Publicar** en la parte inferior derecha de la pantalla.

Query settings >

▼ **Properties**

Name

dimension_customer

Entity type ⓘ

Custom

▼ **Applied steps**

Source

Promoted ...

ABC
123 Changed c...

▼ **Data destination**

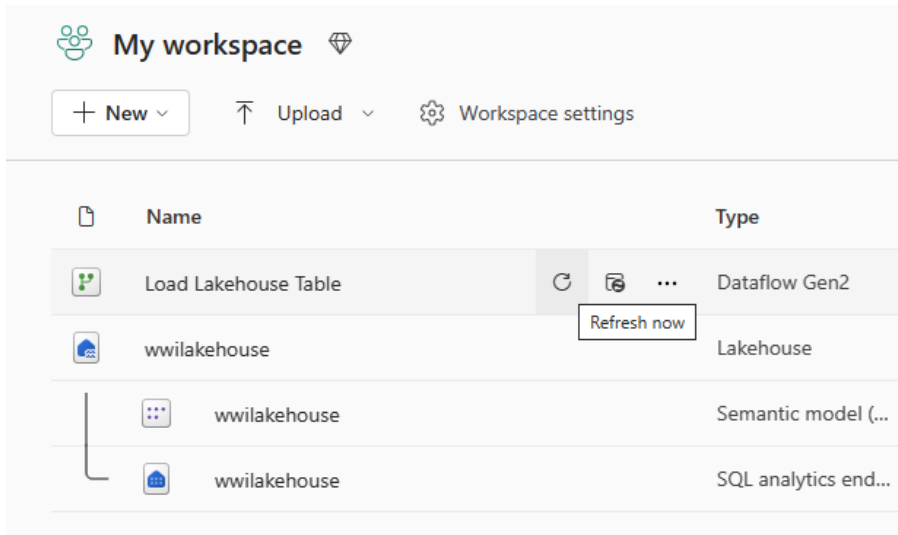
X Lakehouse

Step

Publish

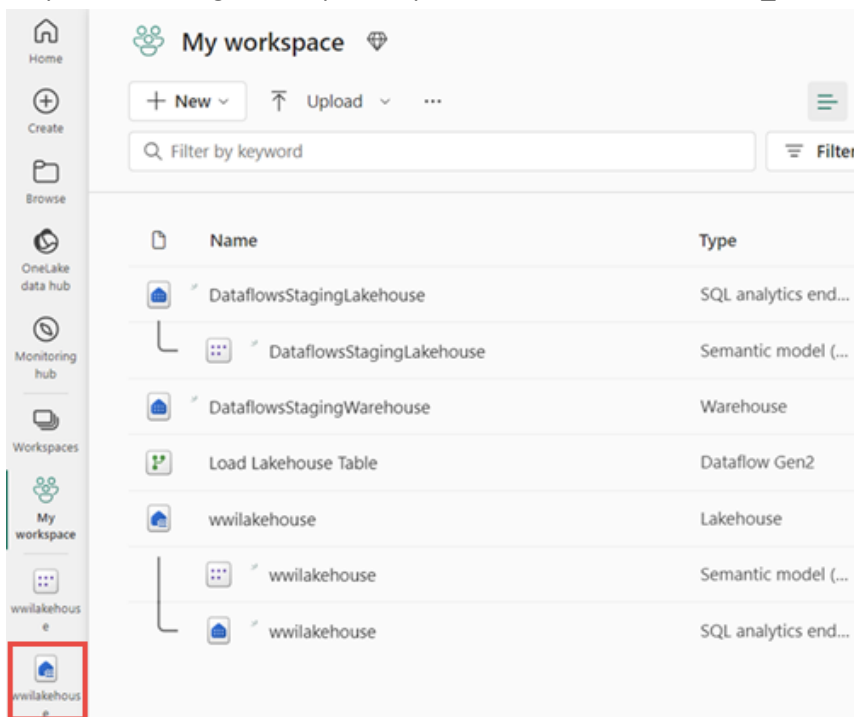
▼

9. Un círculo giratorio junto al nombre del flujo de datos indica que la publicación está en curso en la vista de elementos. Una vez completada la publicación, seleccione ... y seleccione **Propiedades**. Cambie el nombre del flujo de datos a **Cargar tabla de Lakehouse** y seleccione **Guardar**.
10. Seleccione la opción **Actualizar ahora** junto al nombre del flujo de datos para actualizar el flujo de datos. Ejecuta el flujo de datos y mueve los datos del archivo de origen a la tabla de lago de datos. Mientras está en curso, verá un círculo giratorio en la columna **Actualizado** en la vista de elementos.

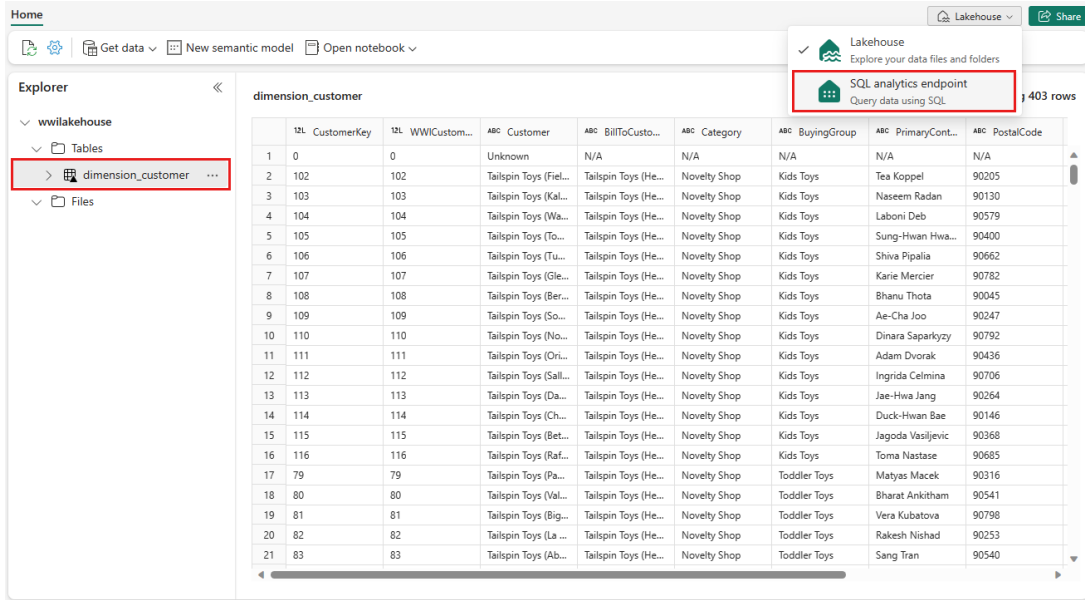


Tarea 2: Verificar los Datos cargados con SQL

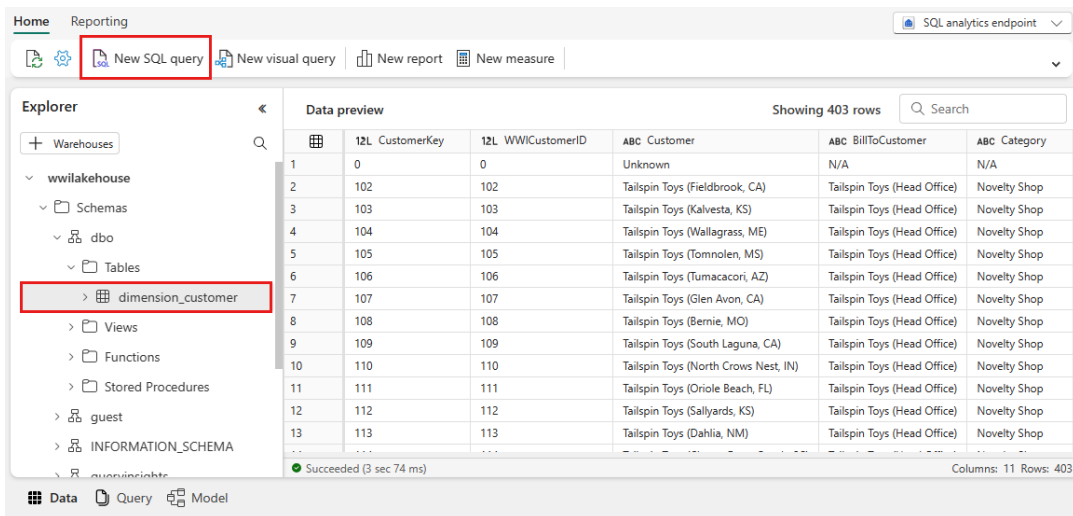
11. Una vez actualizado el flujo de datos, seleccione la nueva instancia de lago de datos en el panel de navegación izquierdo para ver la tabla delta de **dim_customers**.



12. Seleccione la tabla para obtener una vista previa de sus datos. También puede usar el punto de conexión de SQL Analytics del almacén de lago para consultar los datos con instrucciones SQL. Seleccione **Punto de conexión de SQL Analytics** en el menú desplegable de **Almacén de lago** en la parte superior derecha de la pantalla.



13. Seleccione la tabla **dim_customers** para obtener una vista previa de sus datos o seleccione **Nueva consulta SQL** para escribir las instrucciones SQL.



14. La siguiente consulta de ejemplo agrega el recuento de filas en función de la columna *BuyingGroup* de la tabla *dim_customers*. Los archivos de consulta SQL se guardan automáticamente para futuras referencias y puede cambiar el nombre o eliminar estos archivos en función de sus necesidades.

Para ejecutar el script, seleccione el icono **Ejecutar** en la parte superior del archivo de script.

```
SELECT BuyingGroup, Count(*) AS Total
FROM dim_customers
GROUP BY BuyingGroup
```