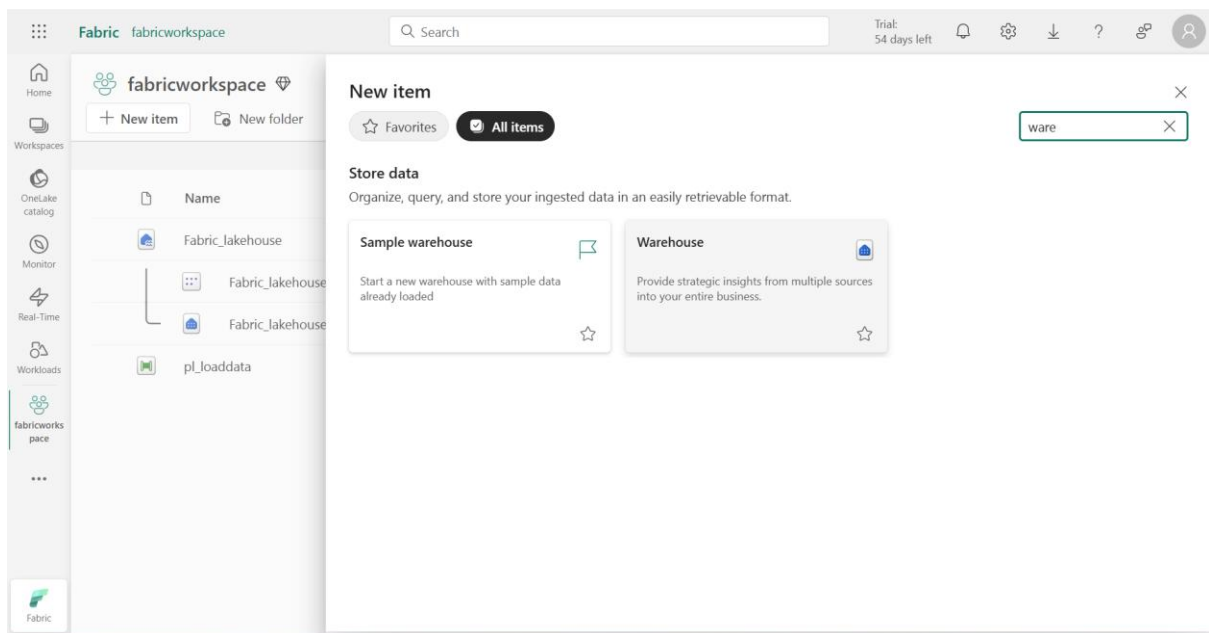# Data warehouse

## Load data dynamically from warehouse to lake house
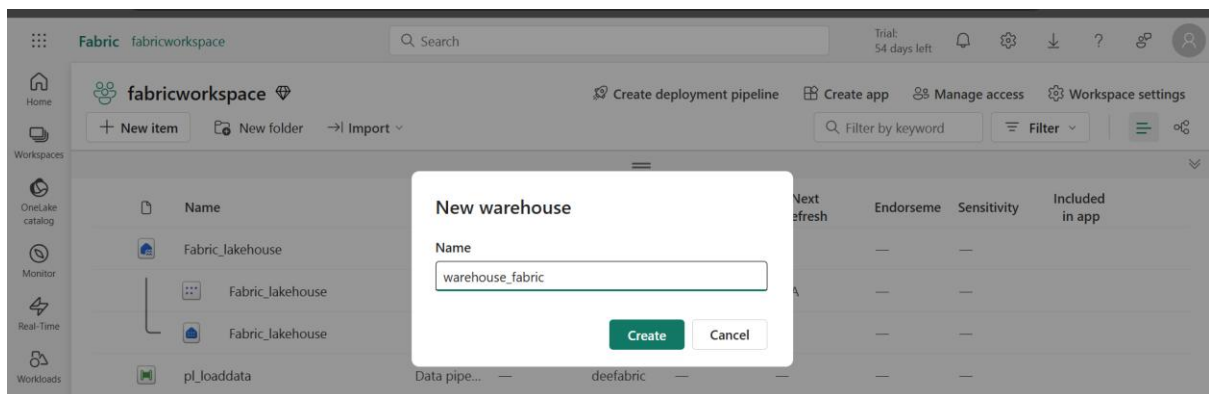
### Data Warehouse

A **data warehouse** is a centralized repository that stores large volumes of structured data from multiple sources. It is designed to support **reporting, analytics, business intelligence (BI)**, and **decision-making** by enabling fast query performance and historical data analysis. Data in a warehouse is typically **cleaned, transformed, and organized** into a consistent format, optimized for analytical workloads rather than transactional processing.
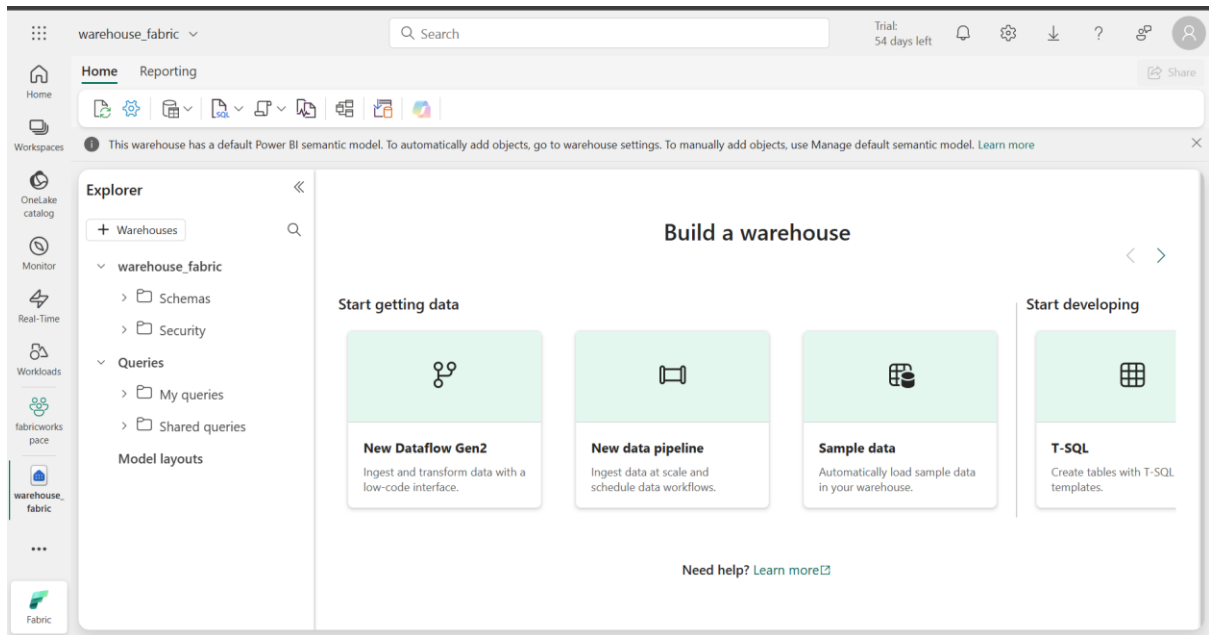
### Create Data warehouse in fabric

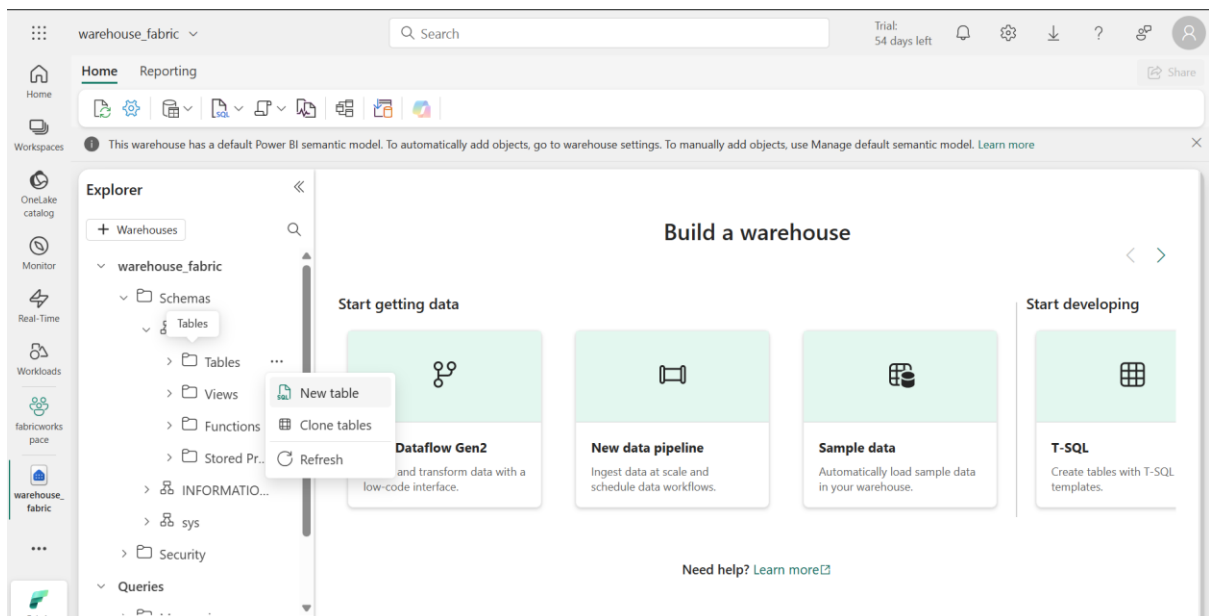Go to workspace -> new item -> search for warehouse and create



Select warehouse

Warehouse created.



Create a table in warehouse

Go to schema->dbo->tables-> 3 dots -> new table



We can create table using below code

CREATE TABLE [warehouse_fabric]. [dbo]. [Customer]
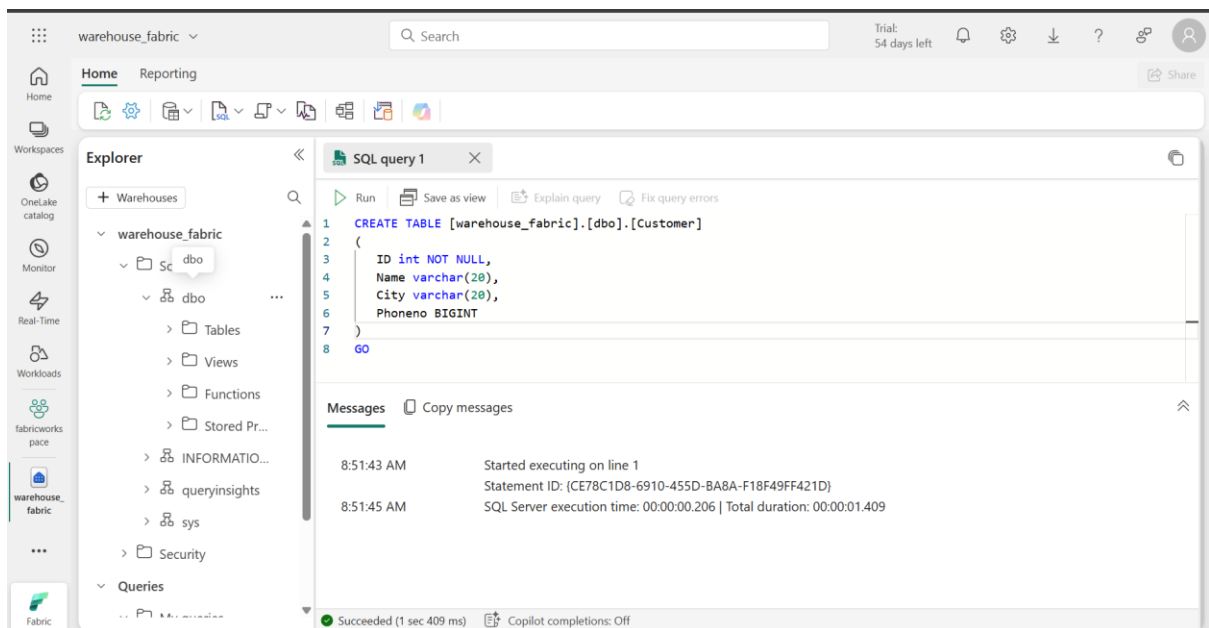(ID int NOT NULL, Name varchar (20), City varchar (20), Phoneno BIGINT )
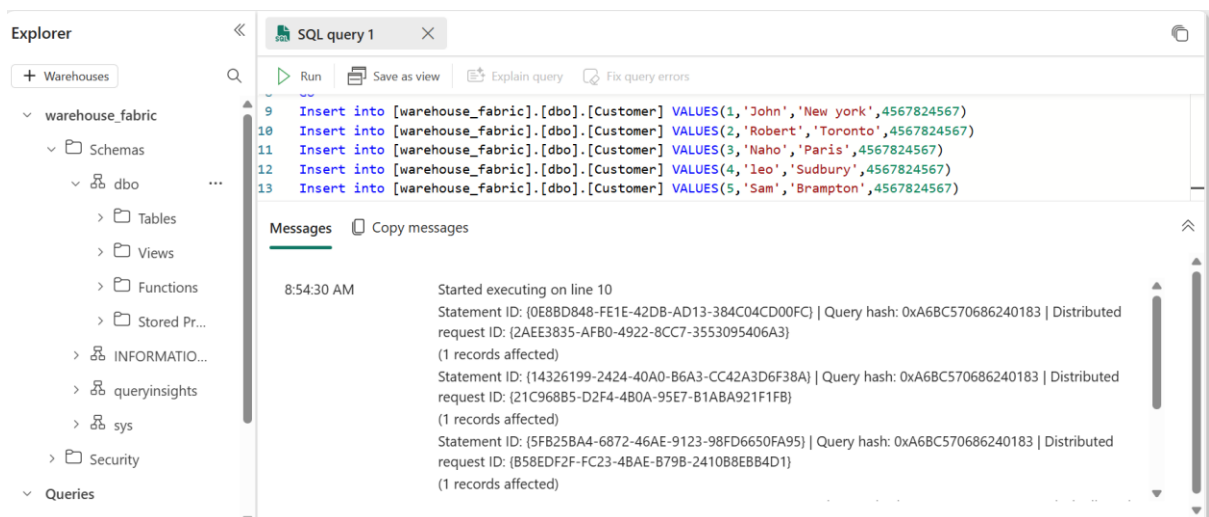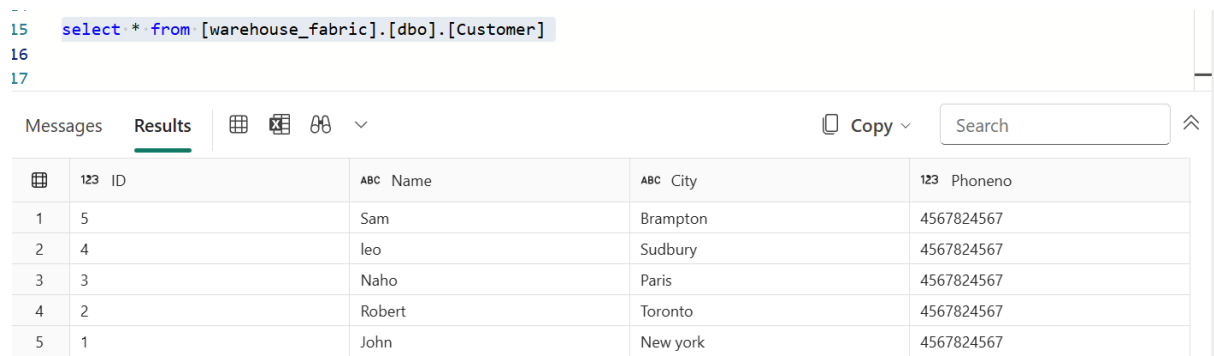GO

Table created successfully

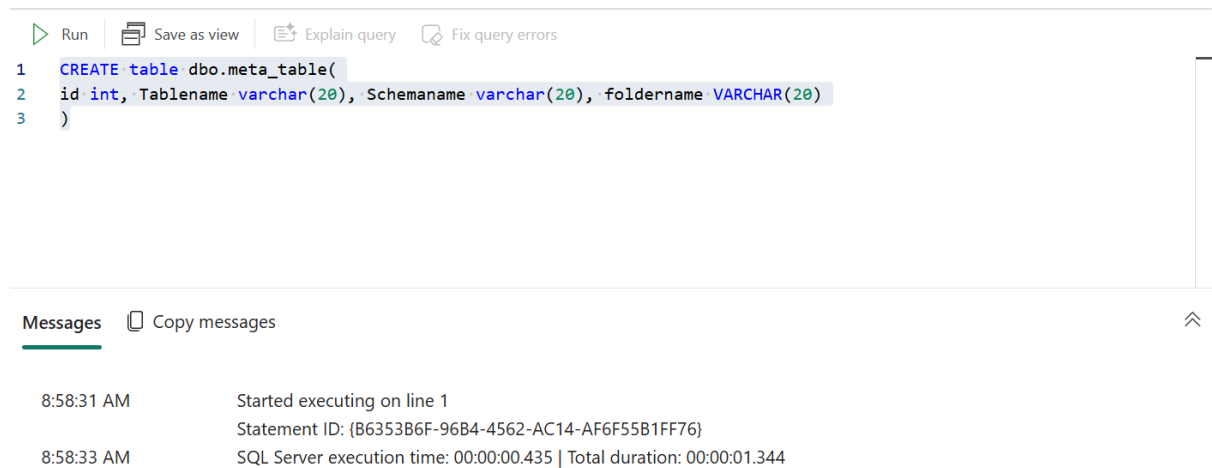Insert data into table



To check data in table use below query.

Load this data into lake house using pipeline dynamically.
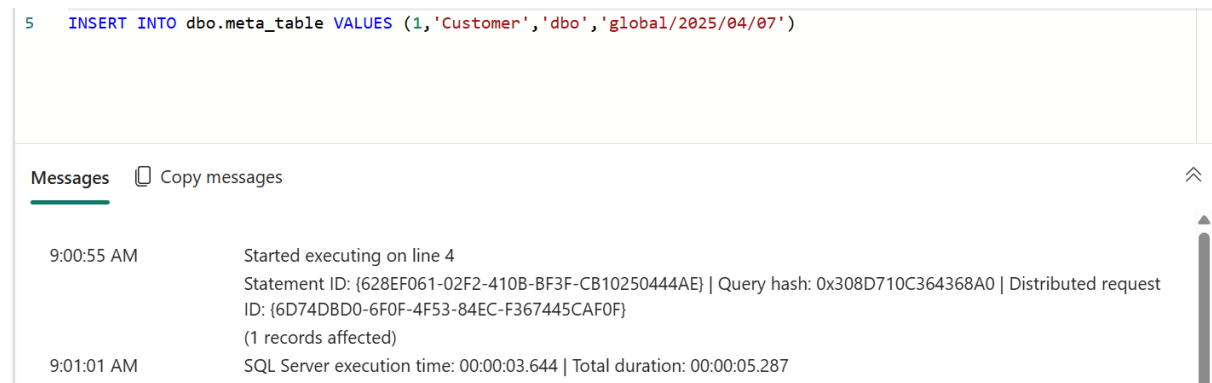
Create a meta_data table in warehouse

This table will store details about all other tables which are in warehouse, by using this table we can get all those table data dynamically.

CREATE table dbo.meta_table(
id int, Tablename varchar(20), Schemaname varchar(20), foldername VARCHAR(20))

```
  ▷ Run      ⊟ Save as view    ⊑⁴ Explain query    ⊘ Fix query errors
1   CREATE table dbo.meta_table(
2   id int, Tablename varchar(20), Schemaname varchar(20), foldername VARCHAR(20)
3   )
```

**Messages**   ☐ Copy messages

| | |
|---|---|
| 8:58:31 AM | Started executing on line 1 |
| | Statement ID: {B6353B6F-96B4-4562-AC14-AF6F55B1FF76} |
| 8:58:33 AM | SQL Server execution time: 00:00:00.435 \| Total duration: 00:00:01.344 |
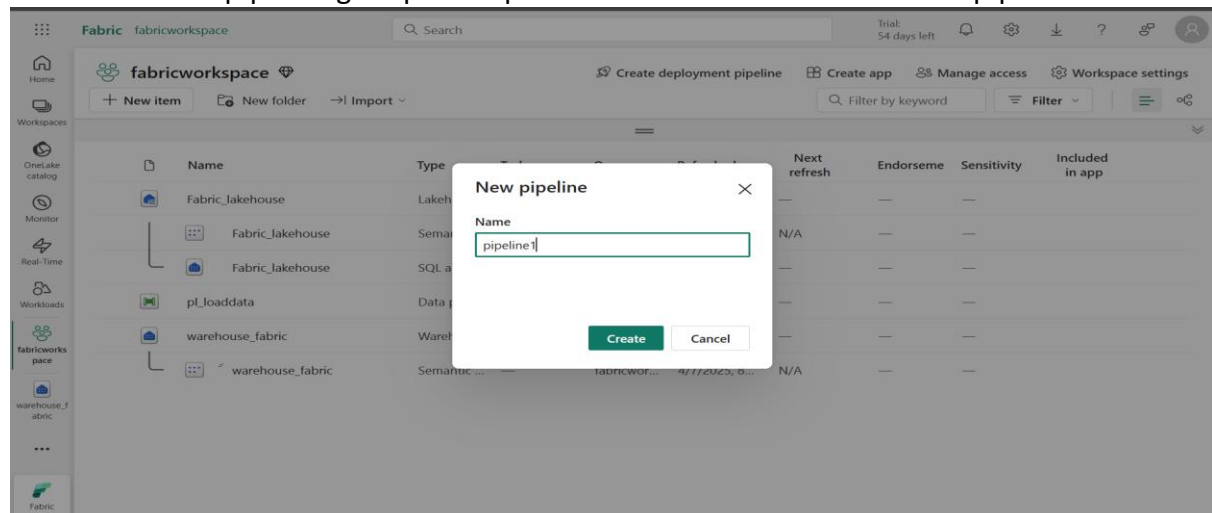
Insert values into this table

```
5     INSERT INTO dbo.meta_table VALUES (1,'Customer','dbo','global/2025/04/07')
```

**Messages**   ☐ Copy messages

| | |
|---|---|
| 9:00:55 AM | Started executing on line 4 |
| | Statement ID: {628EF061-02F2-410B-BF3F-CB10250444AE} \| Query hash: 0x308D710C364368A0 \| Distributed request |
| | ID: {6D74DBD0-6F0F-4F53-84EC-F367445CAF0F} |
| | (1 records affected) |
| 9:01:01 AM | SQL Server execution time: 00:00:03.644 \| Total duration: 00:00:05.287 |

As we have only one table in warehouse only one record is inserted into meta_data table

Now to create a pipeline go top workspace -> new item -> search for data pipeline -> create
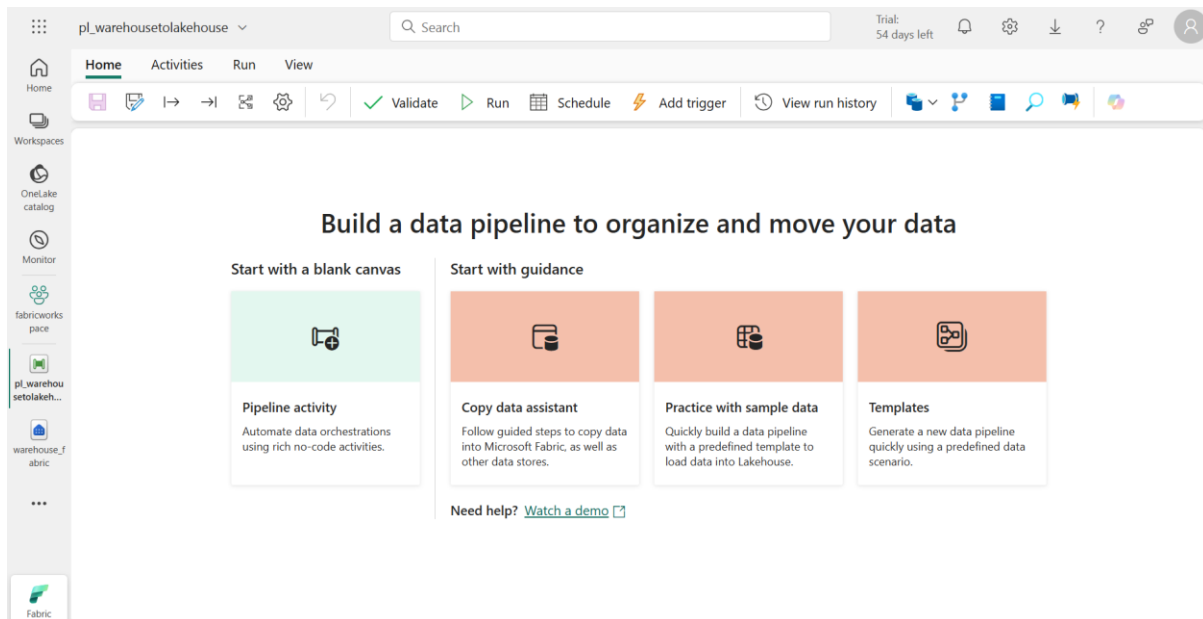
To check what all activities we have in pipeline go to activities tab.
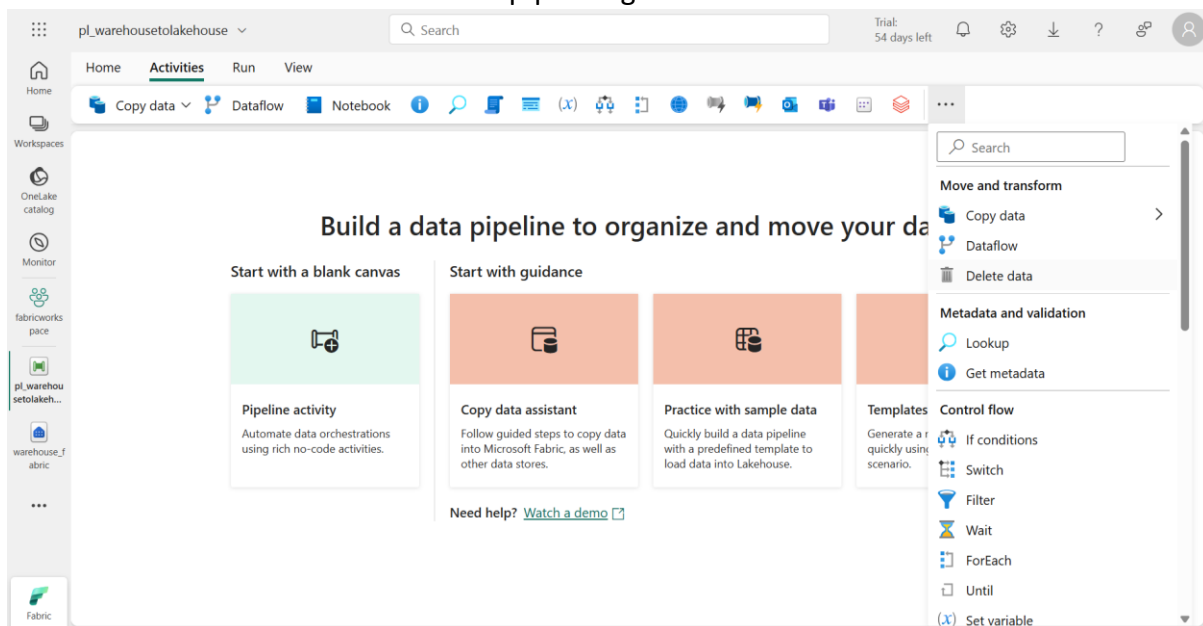


As we are building dynamic pipeline, we need lookup activity, foreach activity and copy activity.

Lookup will take meta_data table as input and will give output as array of all those rows in meta_data table.

Foreach will loop through that array and give input to copy data activity.

Copy data activity will copy data from warehouse to lake house.

Select lookup activity, give name go to settings table and select connection as warehouse and use query as table as we need to get data from meta_data table, and go to table and check enter manually or can select table directly and enter schema and table name and uncheck first row only as we need to loop through the table we need output as array.
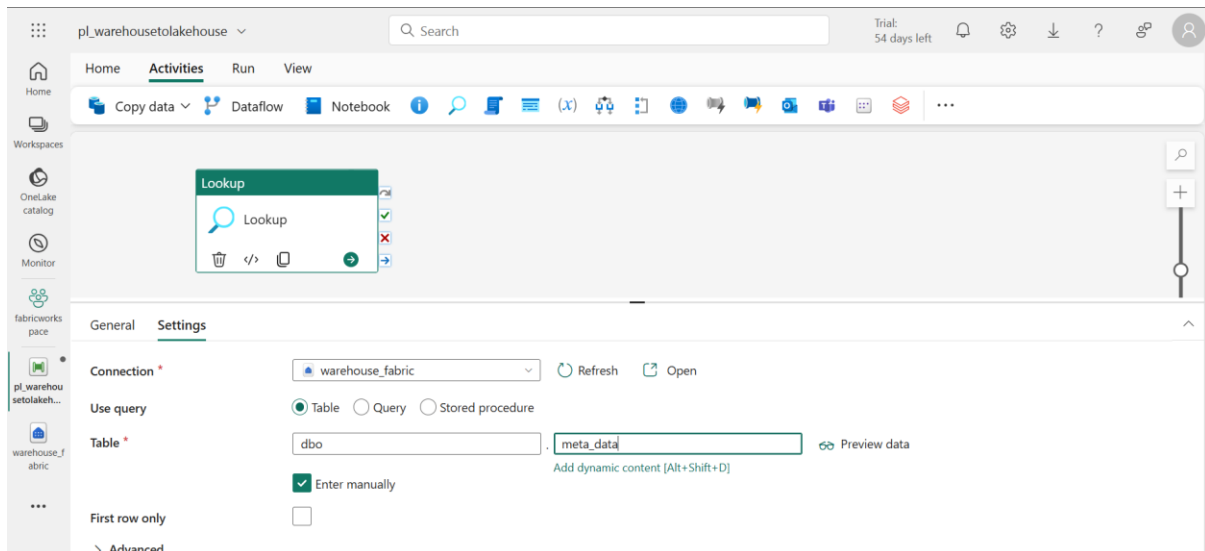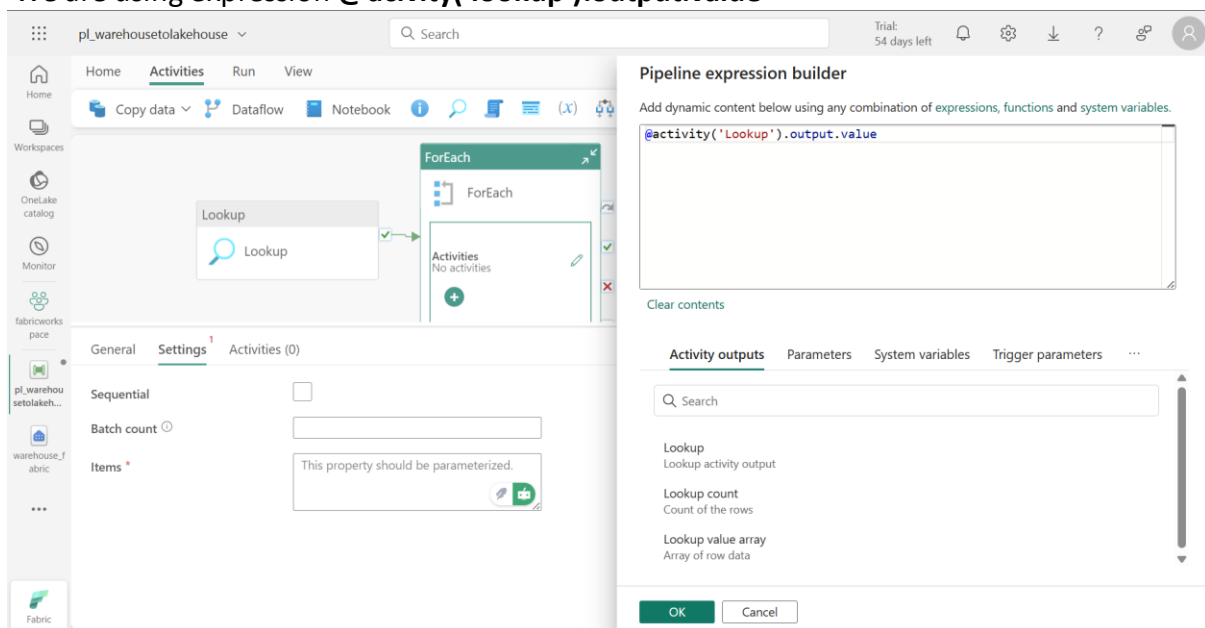
Select foreach activity

Connect lookup to foreach on success.

Go to general and give name then go to settings, items -> add dynamic content and select lookup value array as we need to loop through array.
 We are using expression **@acivity('lookup').output.value**



Now go inside foreach and select copy activity.

Once copy data is selected, go to source select connection as data warehouse , use query as table and go to Table and check enter manually as we have loop we need to give these details dynamically, go to schema name -> add dynamic content -> click on foreach iterator -> expression @item().Schemaname, this schema name is which we given in meta_data table and same we have to do with table name with expression @ item().Tablename

Home    Activities    Run    View

Copy data ∨    Dataflow    Notebook    (i)    🔍    📄    ☰    (x)    🔗

Main canvas > ⬚ ForEach

**Copy data**
Copy data
🗑 </> 📋 ➡

General    **Source**    Destination¹    Mapping    Settings

Connection *    🔷 warehouse_fabric    ∨    ↻ Refresh

Use query    ◉ Table    ○ Query    ○ Stored procedure

Table *    schema name    .    table nam
☑ Enter manually

> Advanced

**Pipeline expression builder**

Add dynamic content below using any combination of expressions, functions and system variables.

@item().Schemaname

Clear contents

**ForEach iterator**    Activity outputs    Parameters    System variables    ···

🔍 Search

ForEach
Current item

OK    Cancel

---

Home    Activities    Run    View

Copy data ∨    Dataflow    Notebook    (i)    🔍    📄    ☰    (x)    🔗

Main canvas > ⬚ ForEach

**Copy data**
Copy data
🗑 </> 📋 ➡

General    **Source**    Destination¹    Mapping    Settings

Connection *    🔷 warehouse_fabric    ∨    ↻ Refresh

Use query    ◉ Table    ○ Query    ○ Stored procedure

Table *    @item().Schemaname    .    table nam
☑ Enter manually

> Advanced

**Pipeline expression builder**

Add dynamic content below using any combination of expressions, functions and system variables.

@item().Tablename

Clear contents

**ForEach iterator**    Activity outputs    Parameters    System variables    ···

🔍 Search

ForEach
Current item

OK    Cancel

---

Home    Activities    Run    View

Copy data ∨    Dataflow    Notebook    (i)    🔍    📄    ☰    (x)    🔗    ⬚    🌐    ⚡    📧    📮    ···

Main canvas > ⬚ ForEach

**Copy data**
Copy data
🗑 </> 📋 ➡

General    **Source**    Destination¹    Mapping    Settings

Connection *    🔷 warehouse_fabric    ∨    ↻ Refresh    ⬈ Open

Use query    ◉ Table    ○ Query    ○ Stored procedure

Table *    @item().Schemaname    .    @item().Tablename    👓 Preview data
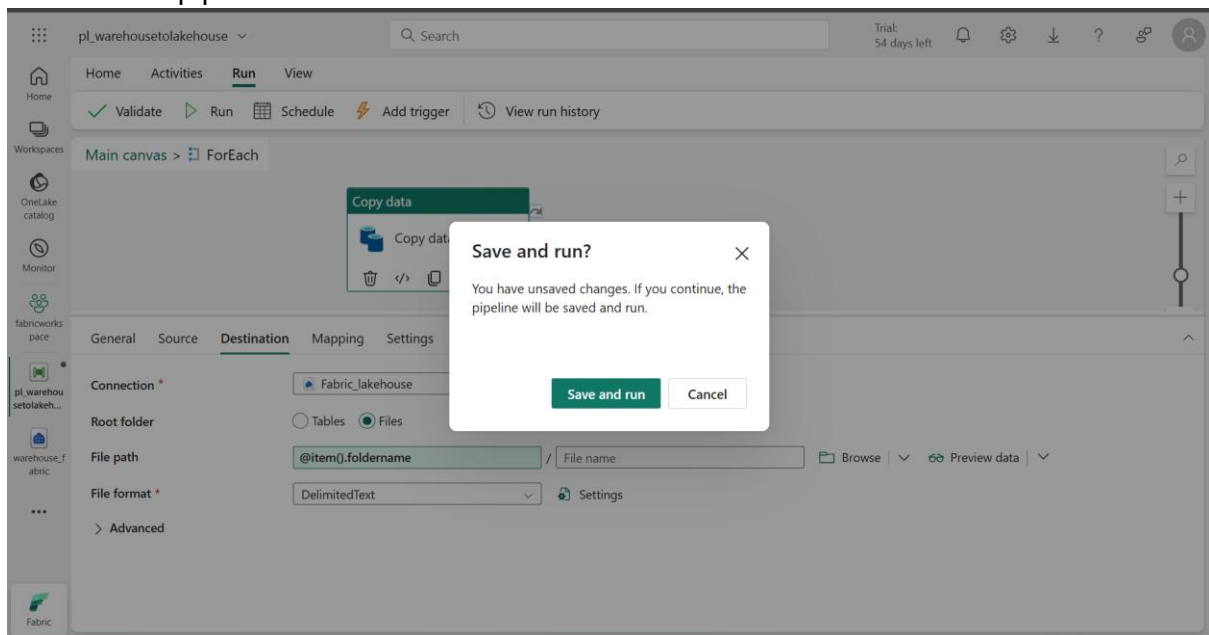☑ Enter manually

> Advanced

Go to destination
Select lakehouse in connection

Select files in root folder as we want to store this in files.
Now to give the file path go to directory -> add dynamic content and give foldername from
meta_data table
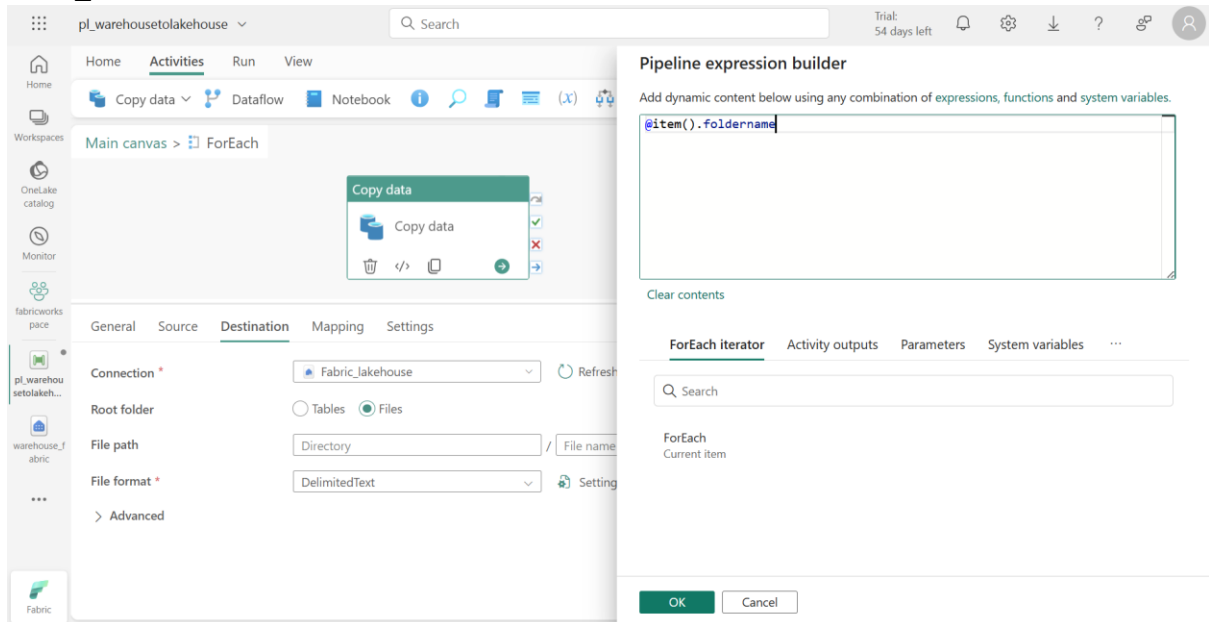


Now run the pipeline



Pipeline ran successfully

To check if data is loaded successfully go to lake house -> check files



Data is stored successfully

**Explore the short cuts feature**

**Shortcut in Microsoft Fabric**

A *Shortcut* in Microsoft Fabric is a reference pointer that allows users to access data stored in other OneLake locations, such as different workspaces, without physically duplicating the data. It enables seamless data sharing across domains, improves collaboration, and optimizes storage by creating a virtual link to existing datasets.

**Key Benefits:**

- Avoids data duplication
- Simplifies data access across workspaces
- Enables unified analytics across domains
- Enhances governance and control by maintaining a single source of truth

To create a short cut to load data from data warehouse to lakehouse

Go to lakehouse -> files /Tables -> 3 dots -> New shortcut
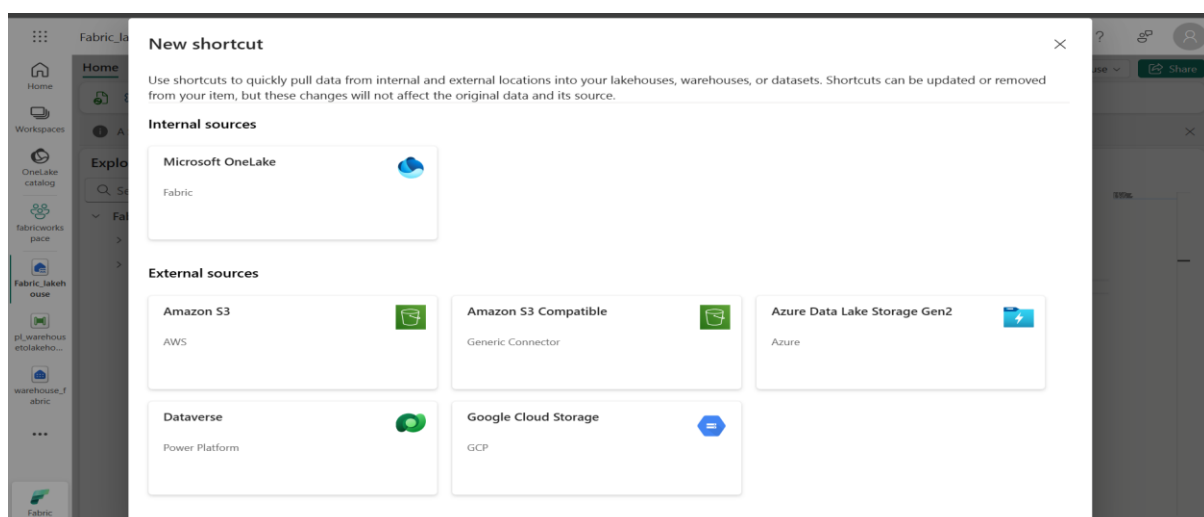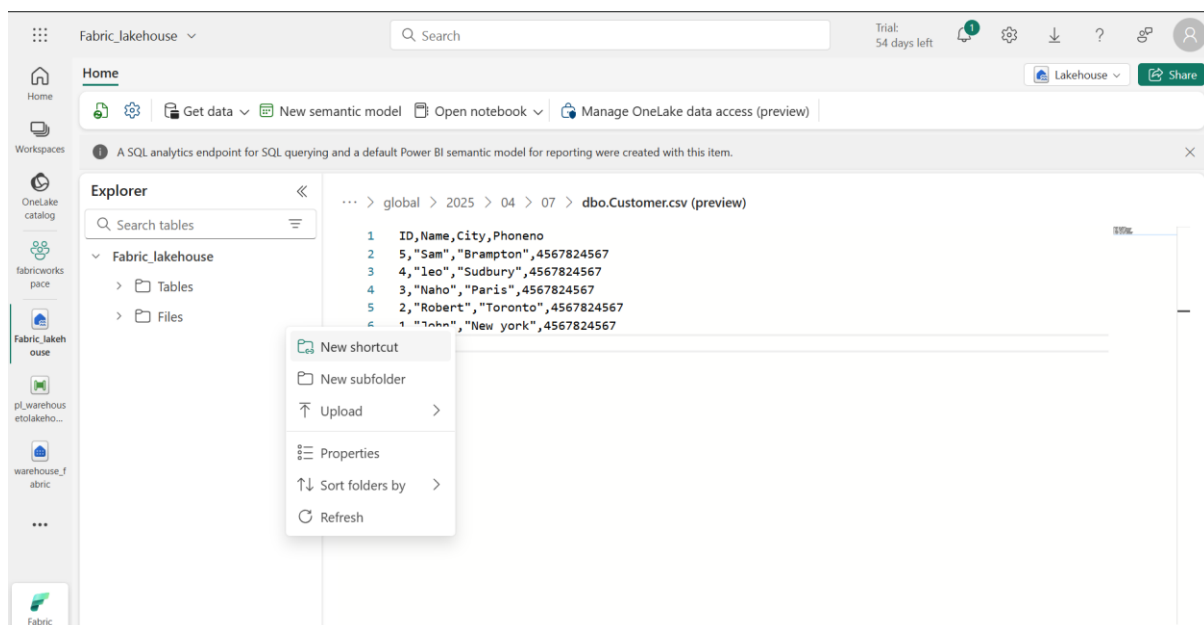
We can create a connection to everything in Onelake and to external sources as well.
Now as we want to create a connection to warehouse select Onelake.
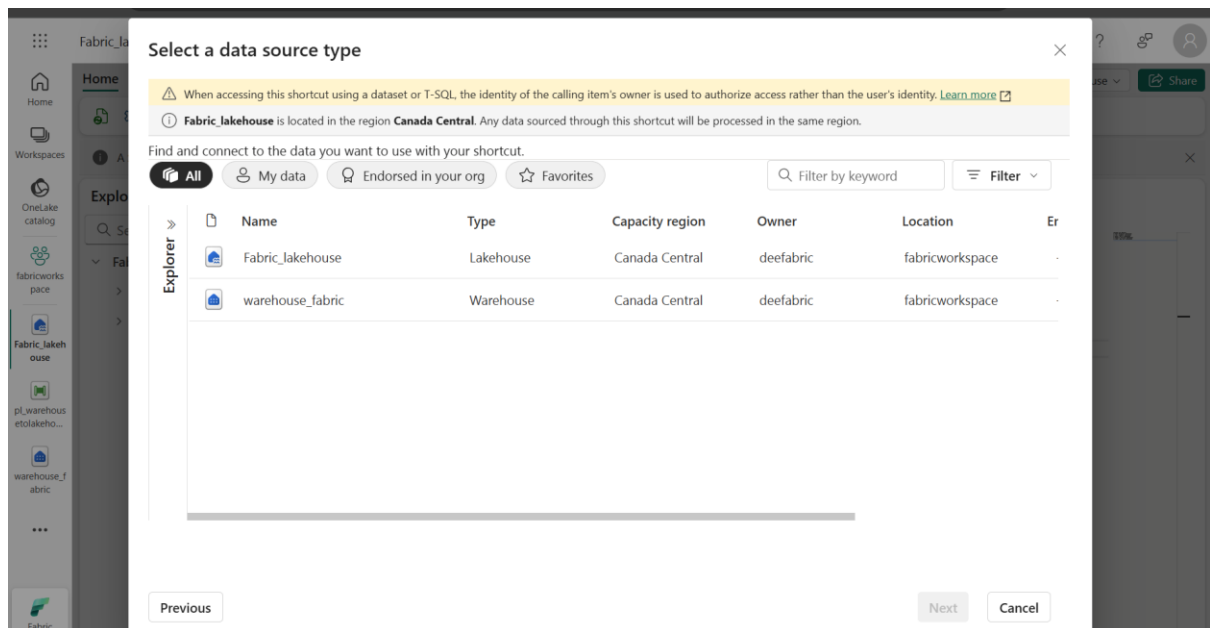


Select warehouse
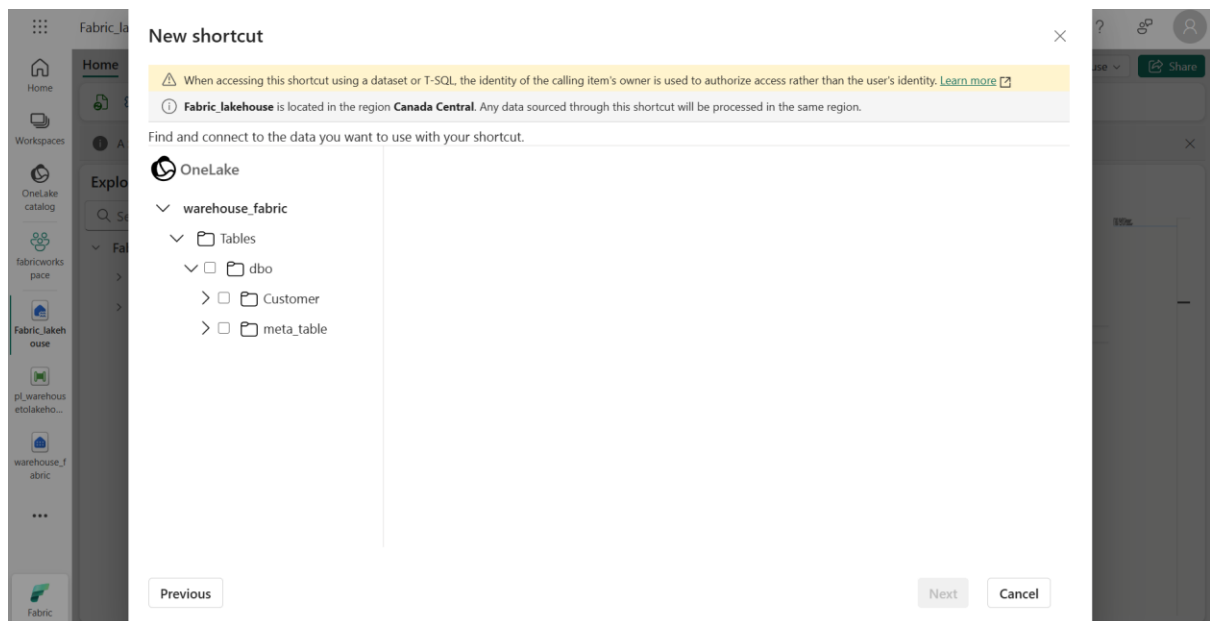


We can select what we want to connect one table or everything.
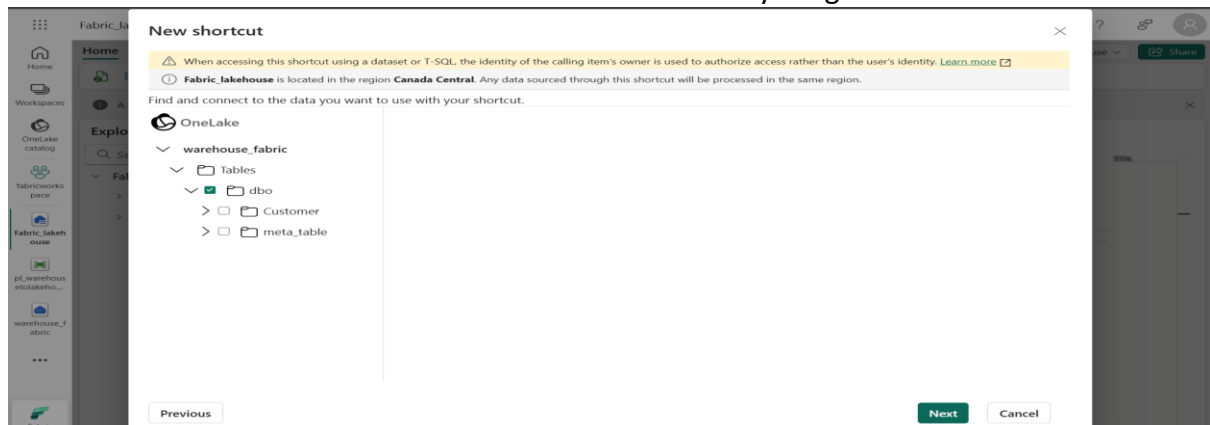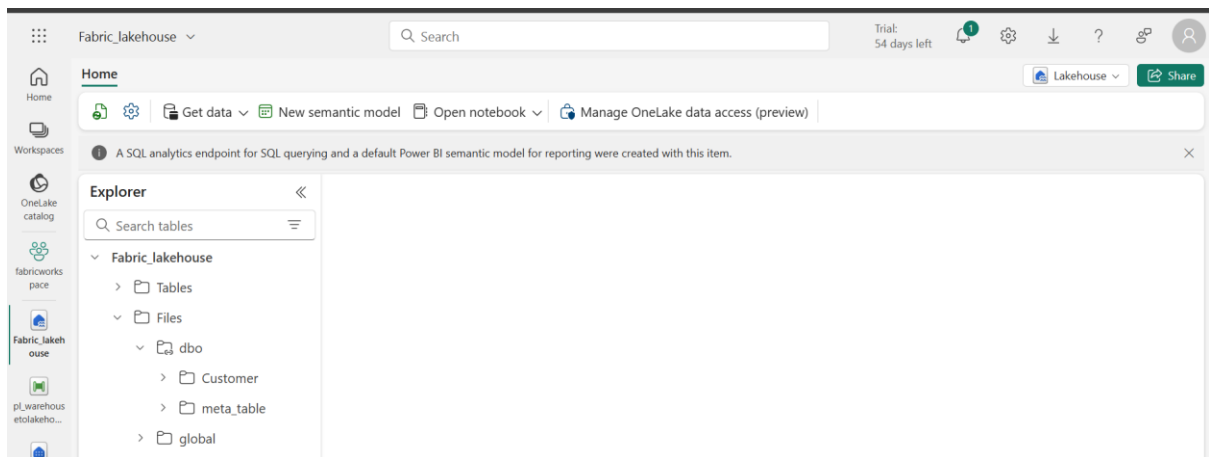
A link is created between lakehouse and Warehouse, we see tables in lakehouse without loading them manually.

As these are tables in files, they are stored in delta format.

We can also create a link to store these tables in Tables and then we can view data in SQL analytics.

We can also create a short cut to external source, let create a shot cur between lakehouse and external Azure data lake gen 2.

As we don't have any existing connection, we need to create a new connection.
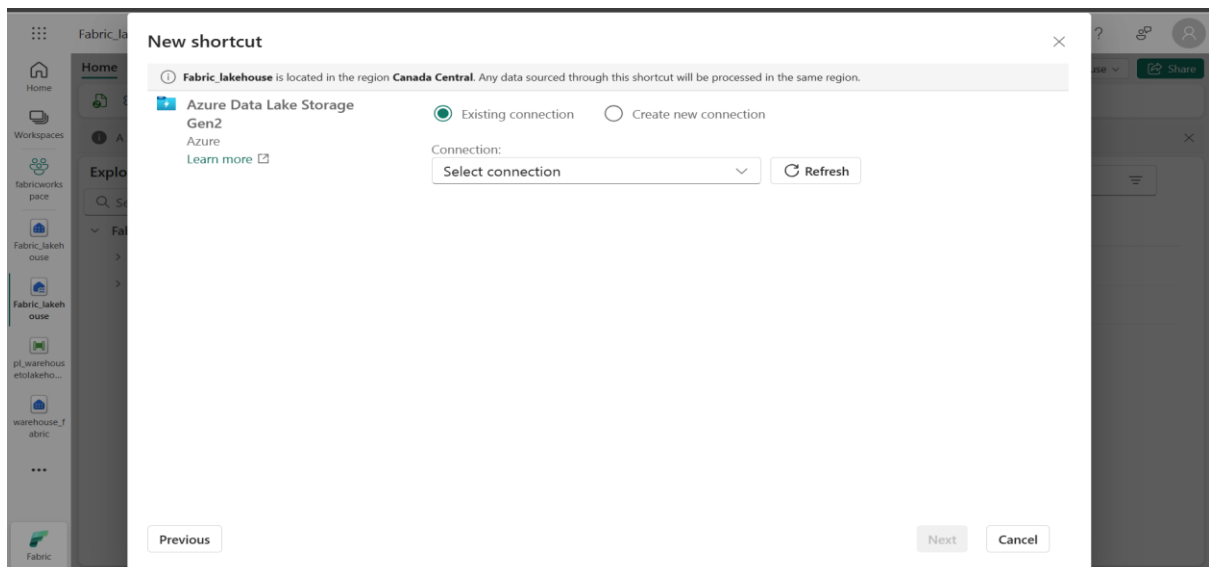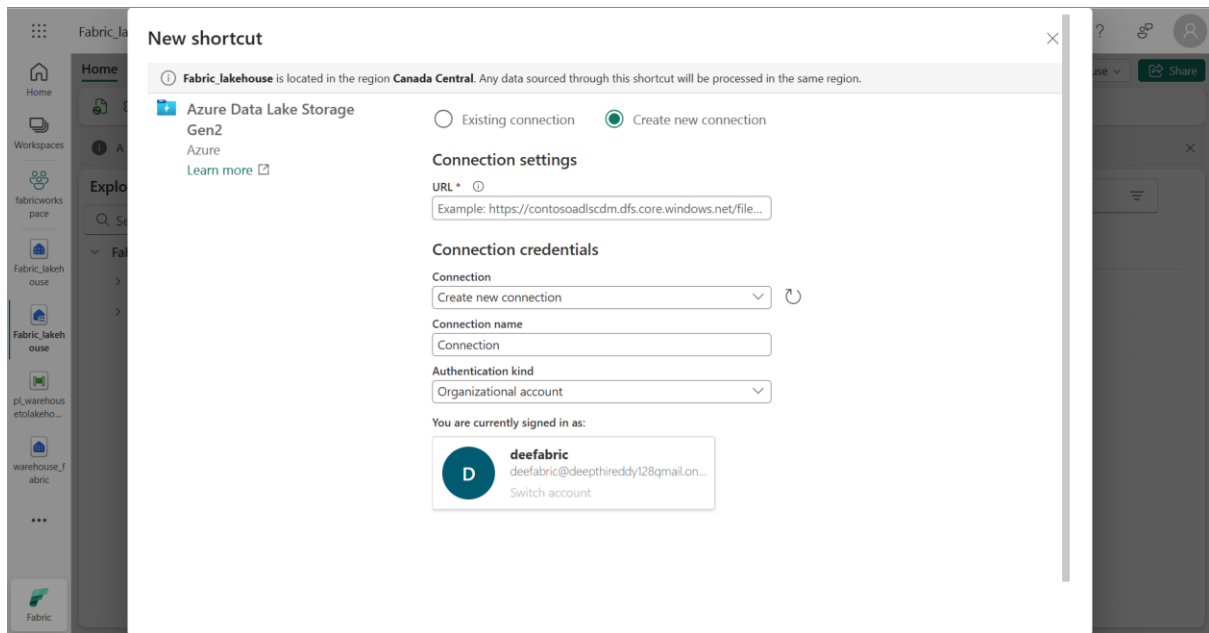
To get a connection string we need to go to Azure data lake gen 2 -> Settings-> Endpoints -> **Data Lake Storage -> Data Lake Storage URL.**
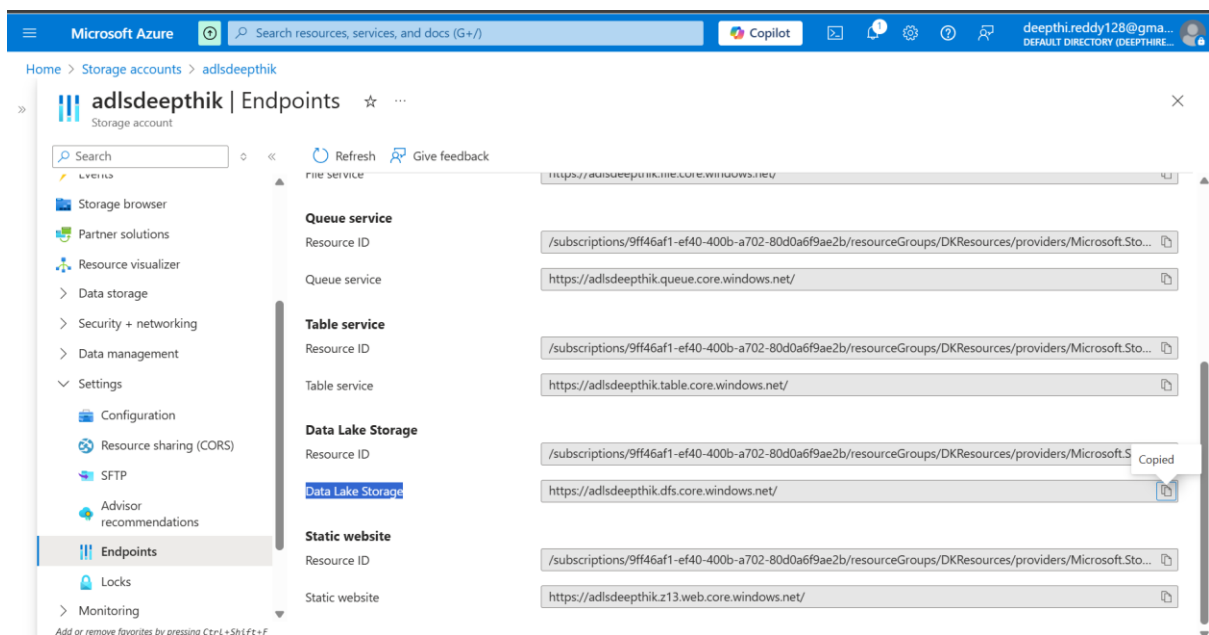


Copy that link and go to fabric
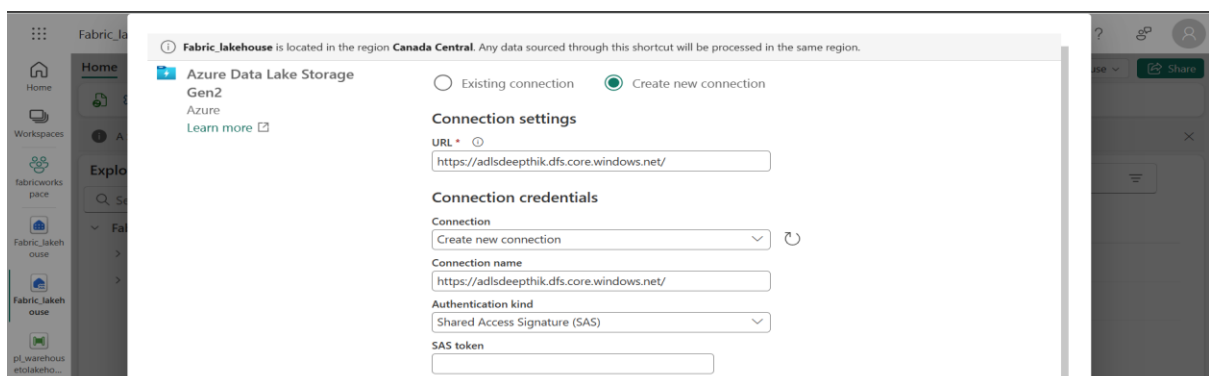
As for Authentication kind choose SAS and provide SAS token from Azure data lake gen 2 storage.

To generate SAS token, go to Azure data lake gen 2 -> Security + Networking -> Shared access signature -> Allowed resource types -> check all Service, Container and Object and generate SAS
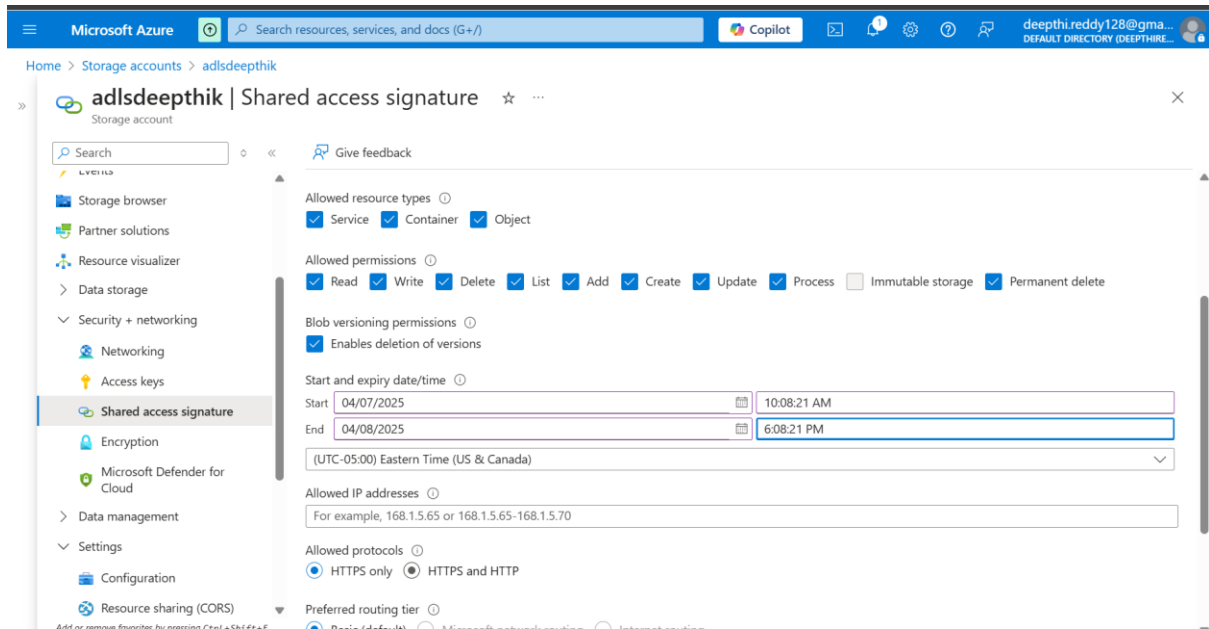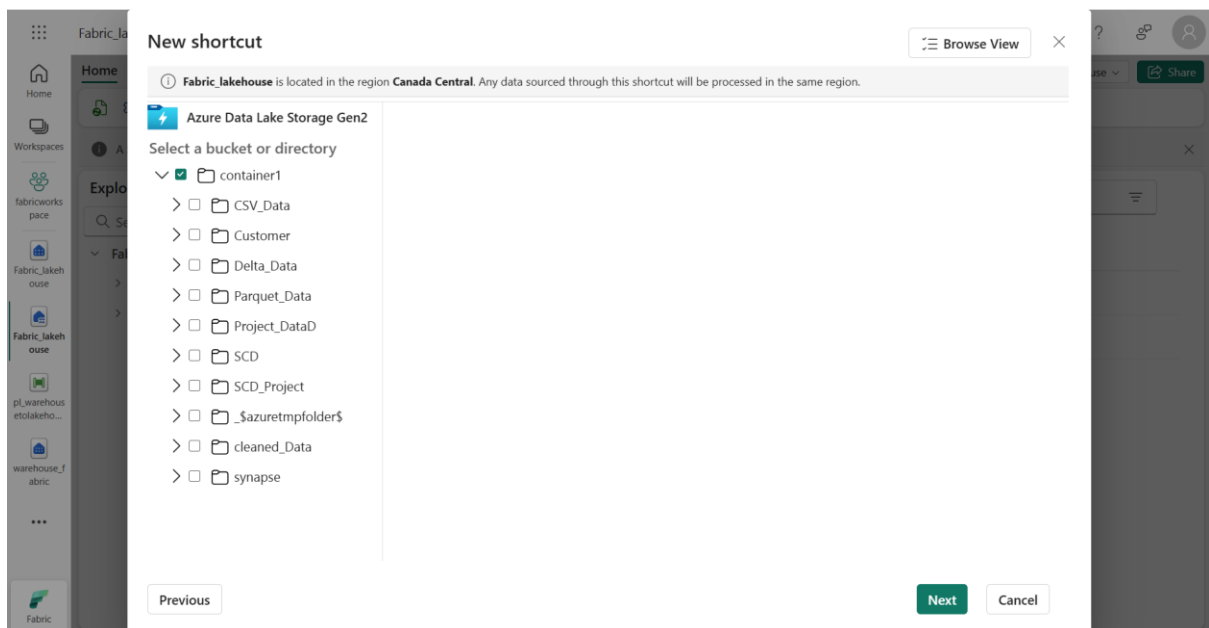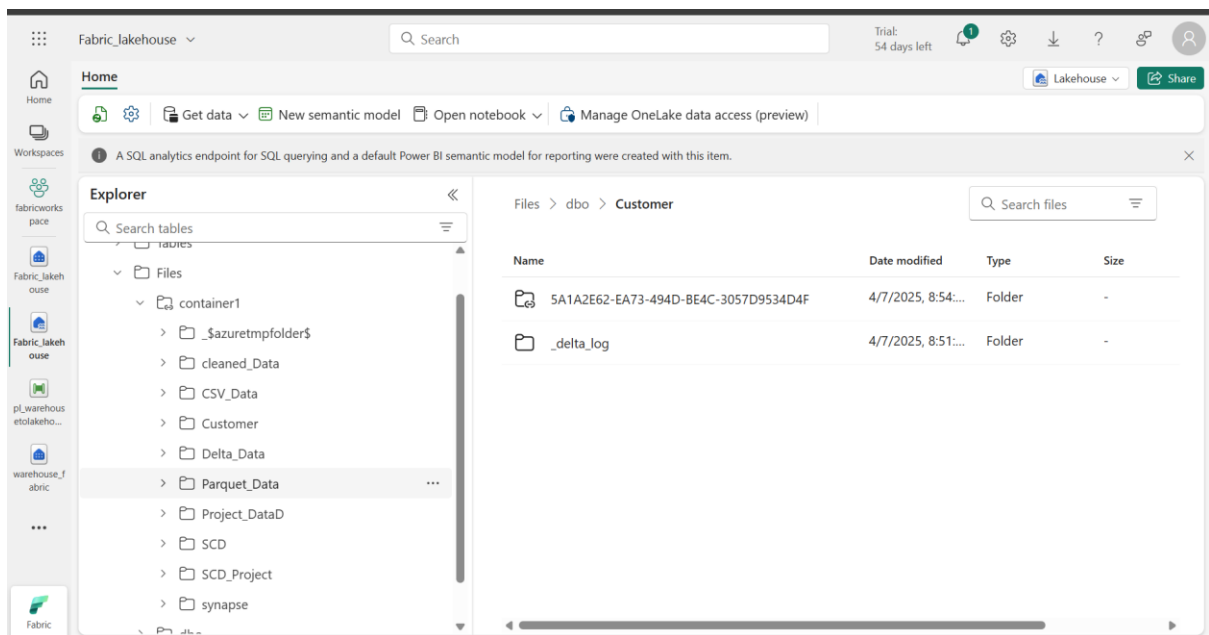


Generate token and copy in fabric.
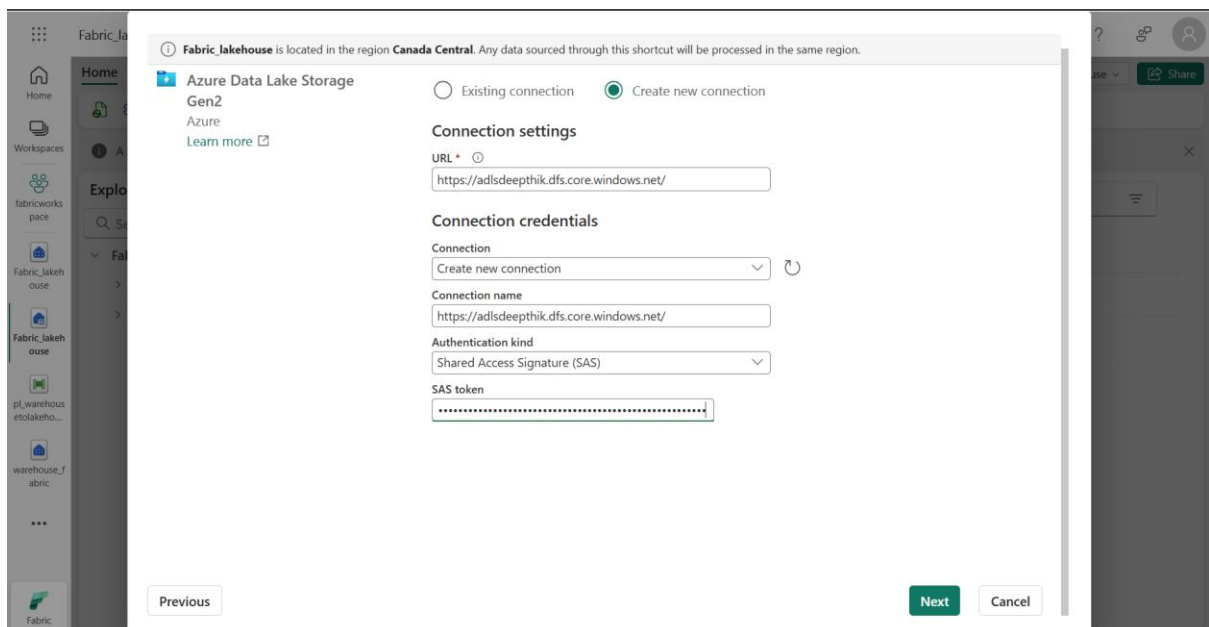


Select for what we need to create a link to container or to specific folder.
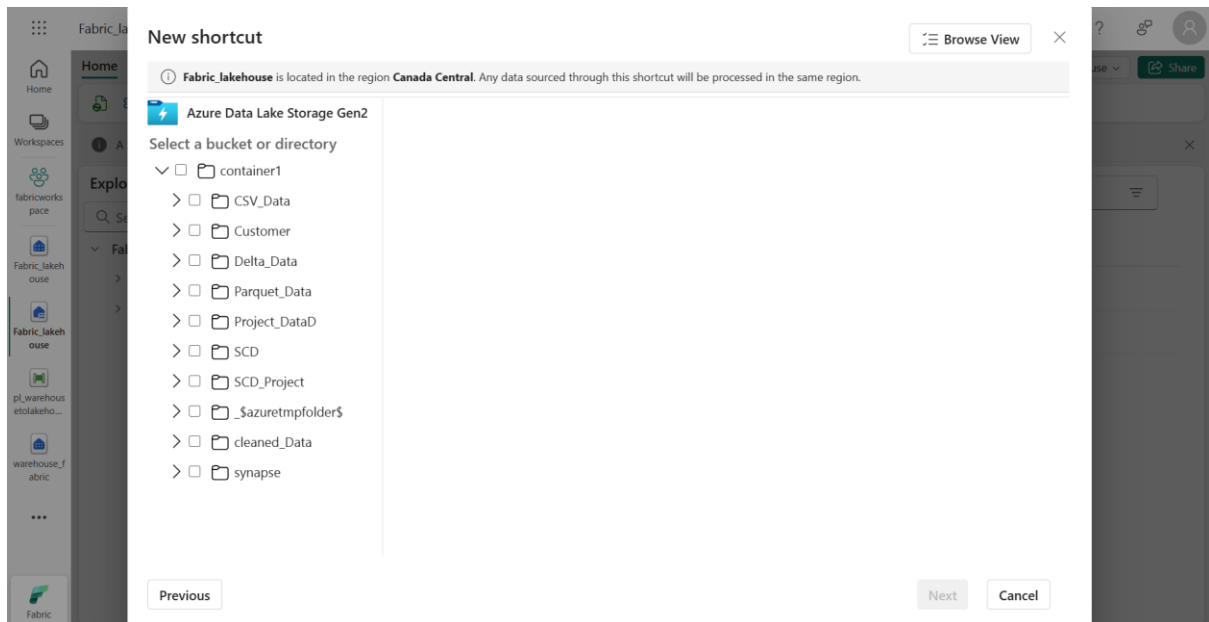
Creating link to container.
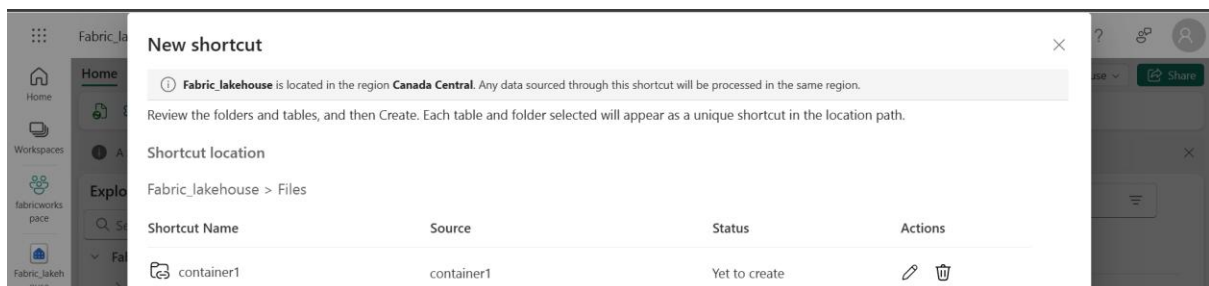




A link is created between Lakehouse and Azure data lake gen 2 storage external source.

We can view the data
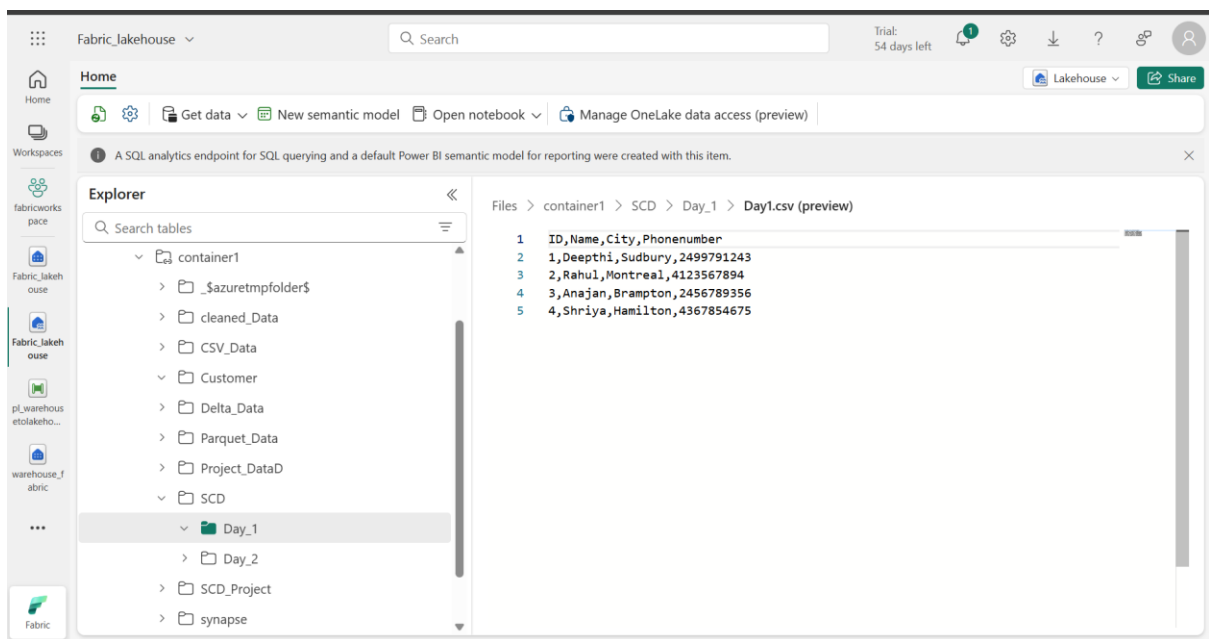


And if we do any changes in container, it will affect here in lake house.

If I delete Parquet_Data folder in Storage account it will affect in lakehosue as well.

Folder is deleted in storage account.



Folder is also got deleted in lakehouse.

Short cut is read only we cannot add any file or folder in lakehouse.