

# Data warehouse

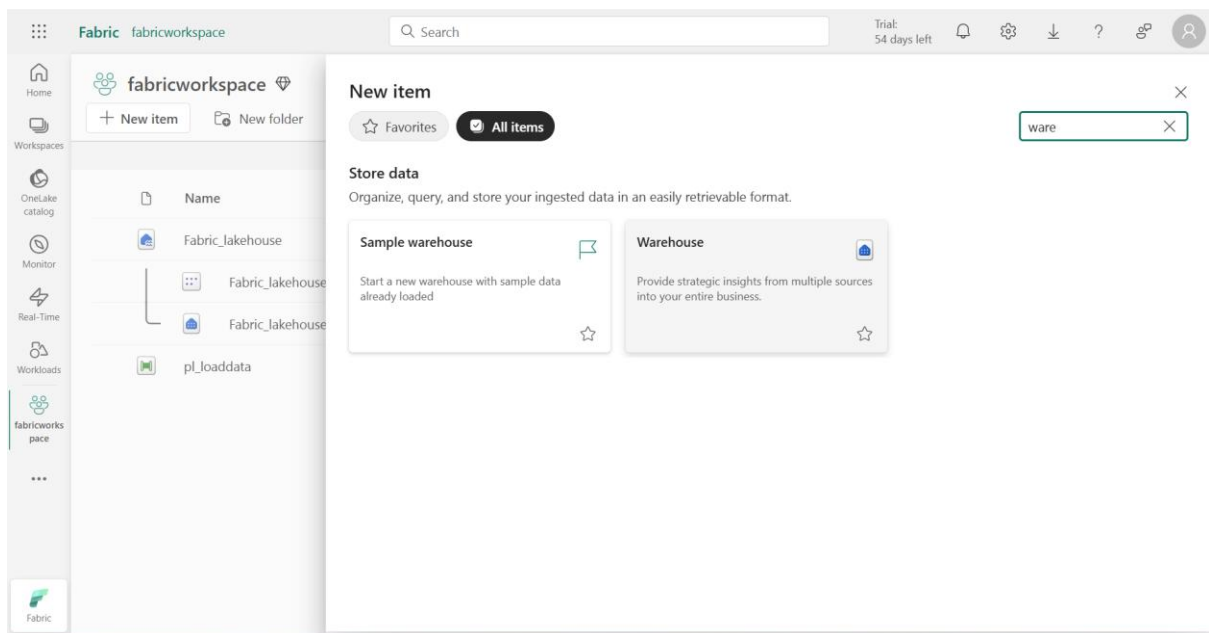
## Load data dynamically from warehouse to lake house

### Data Warehouse

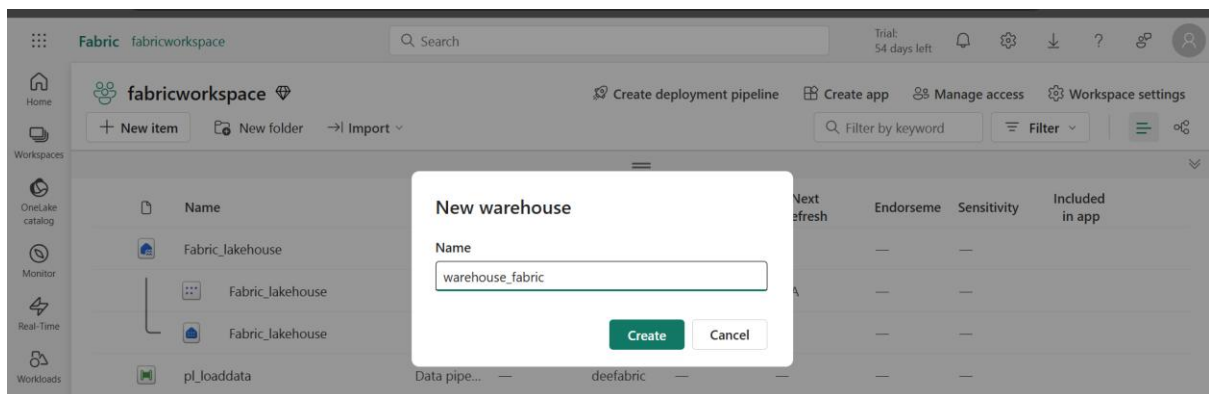
A **data warehouse** is a centralized repository that stores large volumes of structured data from multiple sources. It is designed to support **reporting, analytics, business intelligence (BI), and decision-making** by enabling fast query performance and historical data analysis. Data in a warehouse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

### Create Data warehouse in fabric

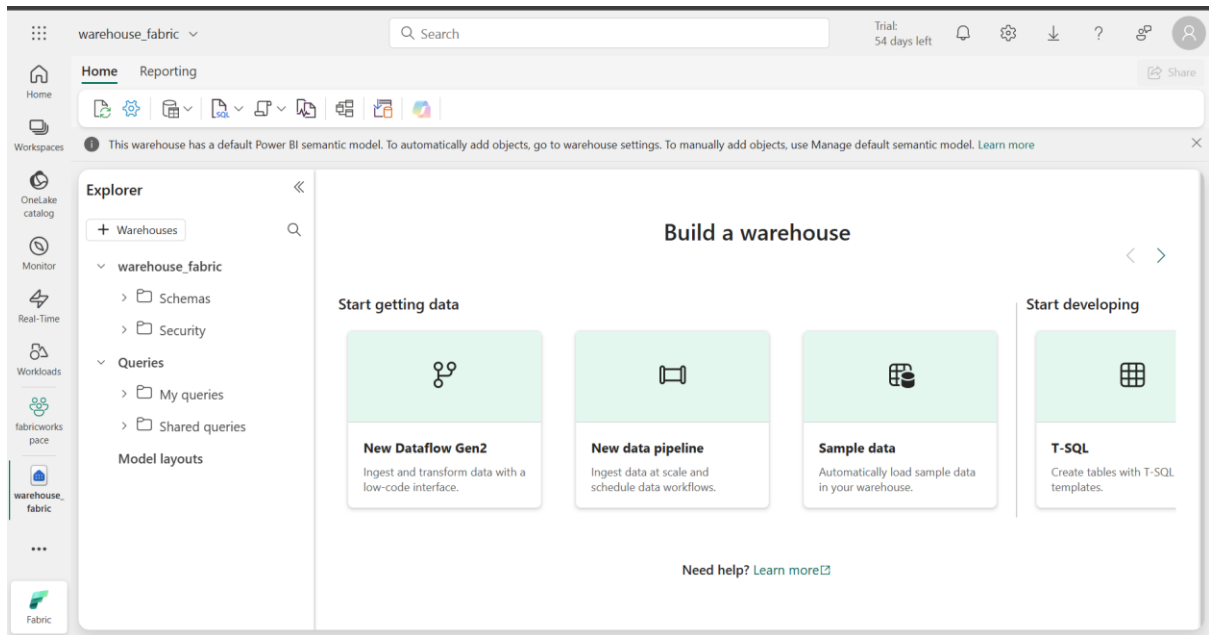
Go to workspace -> new item -> search for warehouse and create



### Select warehouse

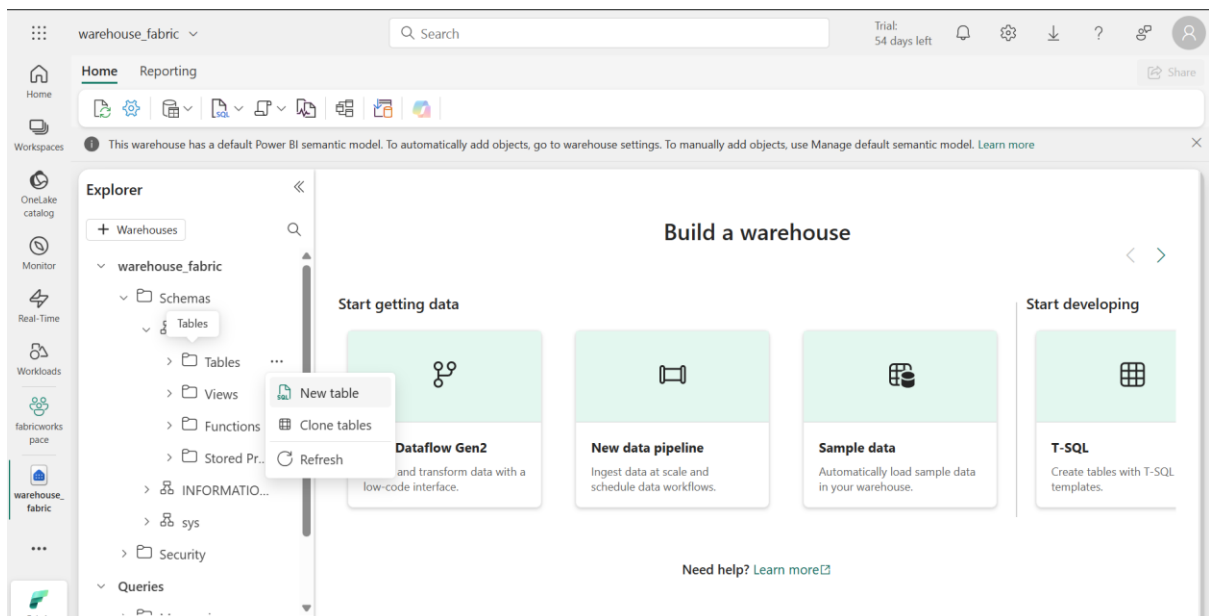


Warehouse created.



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
```

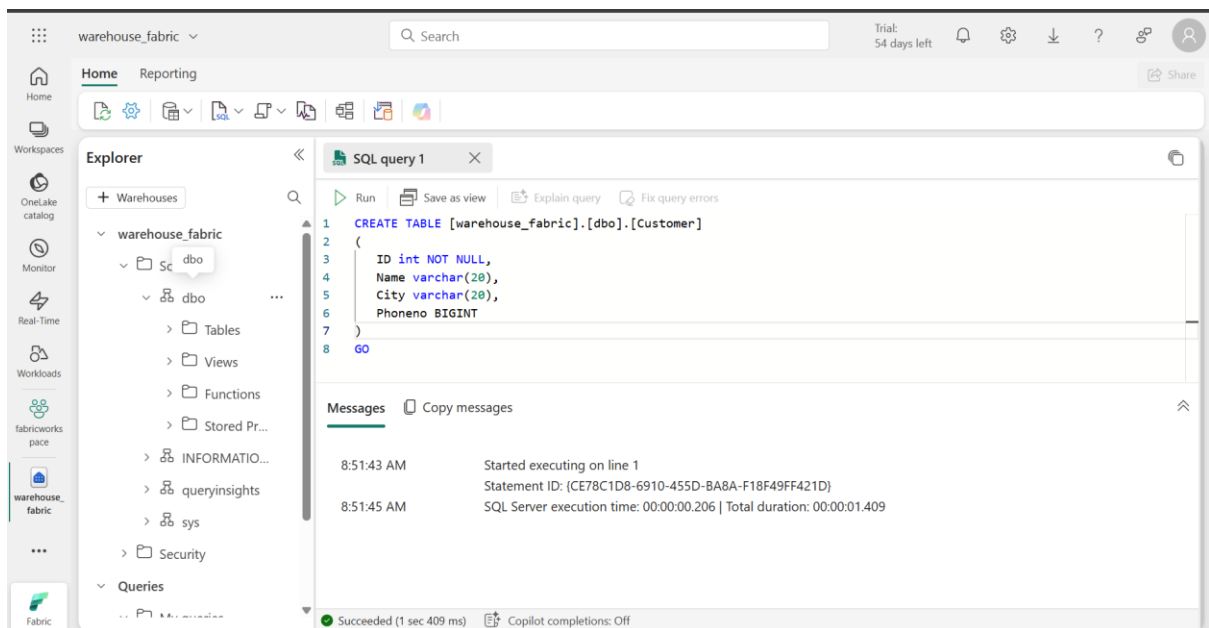
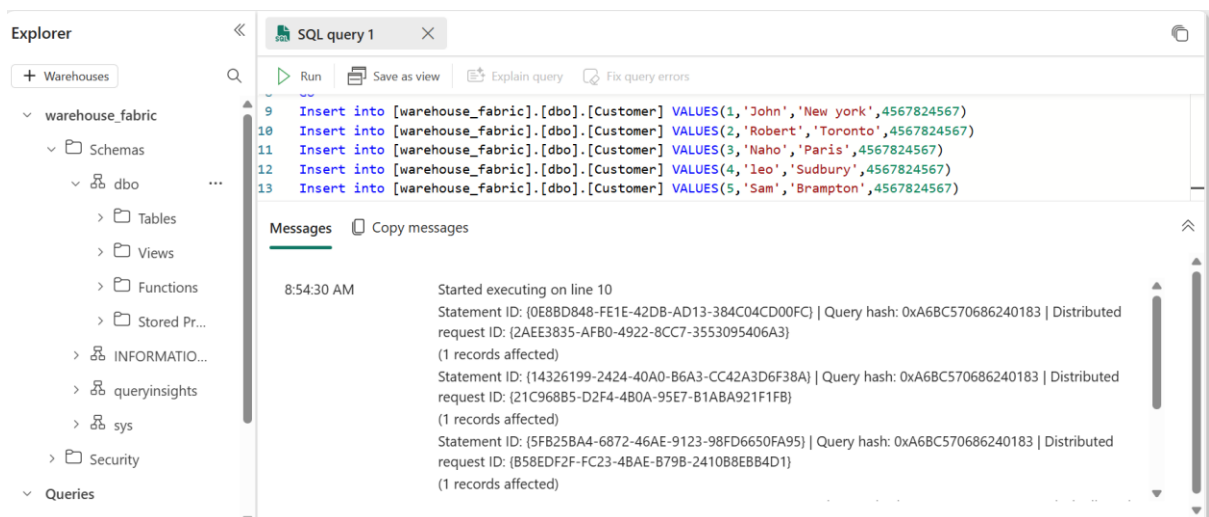
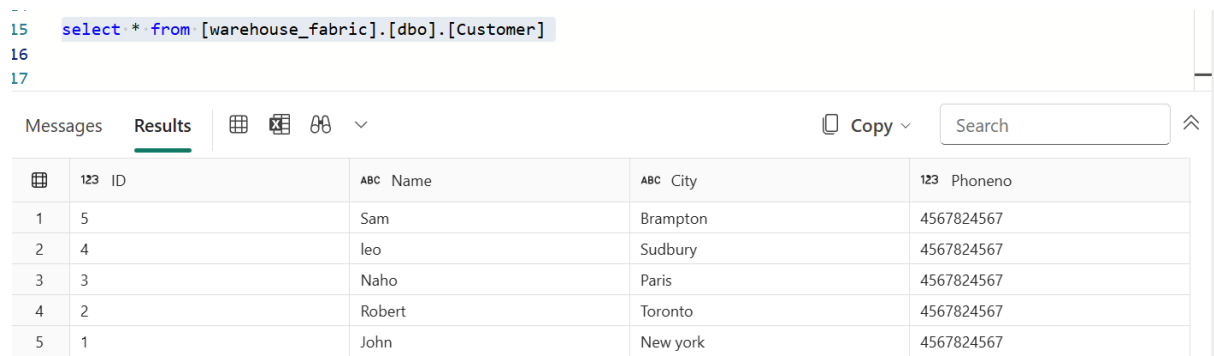


Table created successfully

Insert data into table



To check data in table use below query.

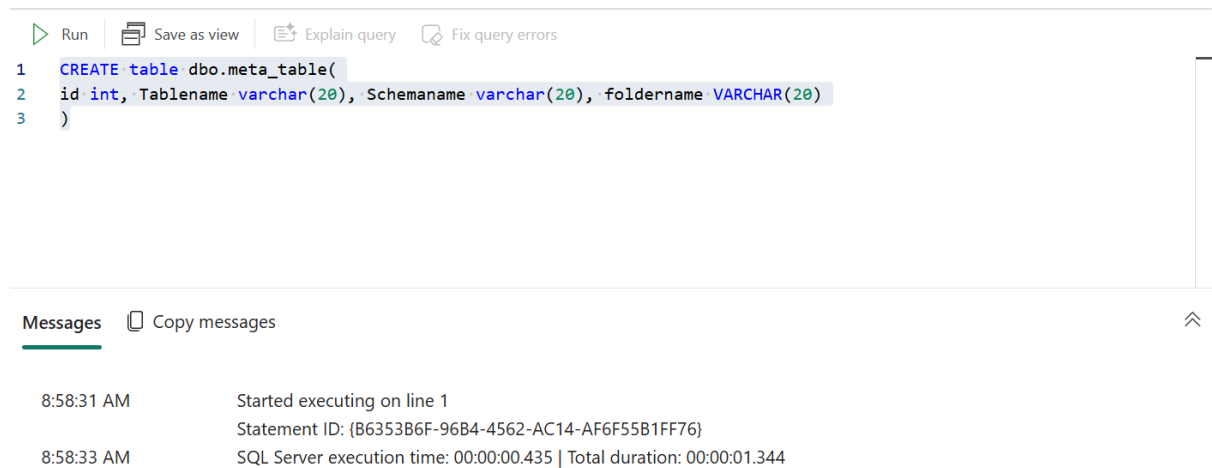


Load this data into lake house using pipeline dynamically.

Create a meta\_data table in warehouse

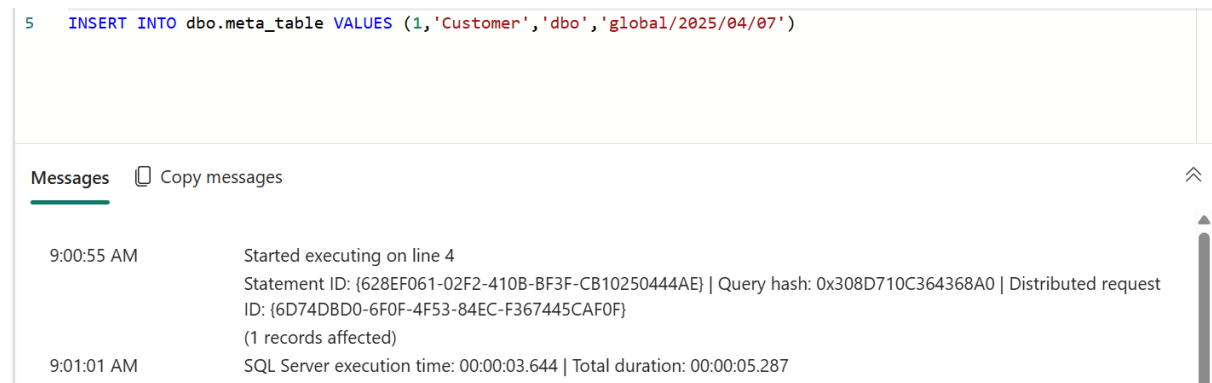
This table will store details about all other tables which are in warehouse, by using this table we can get all those table data dynamically.

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



The screenshot shows a SQL query execution window. The query is: `CREATE table dbo.meta_table(id int, Tablename varchar(20), Schemaname varchar(20), foldername VARCHAR(20))`. The execution messages show it started at 8:58:31 AM and completed at 8:58:33 AM with a total duration of 00:00:01.344.

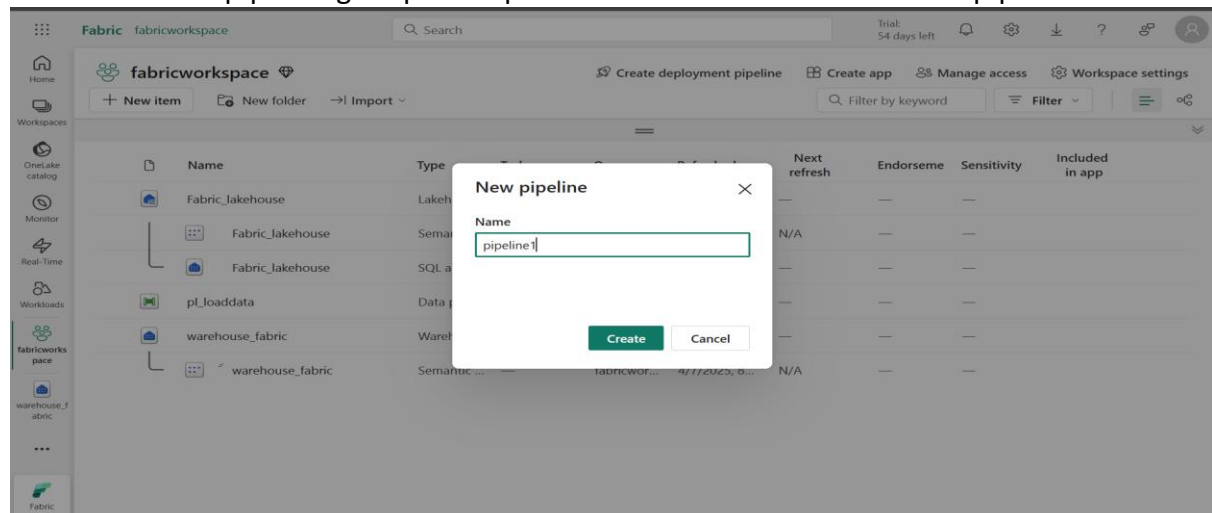
Insert values into this table



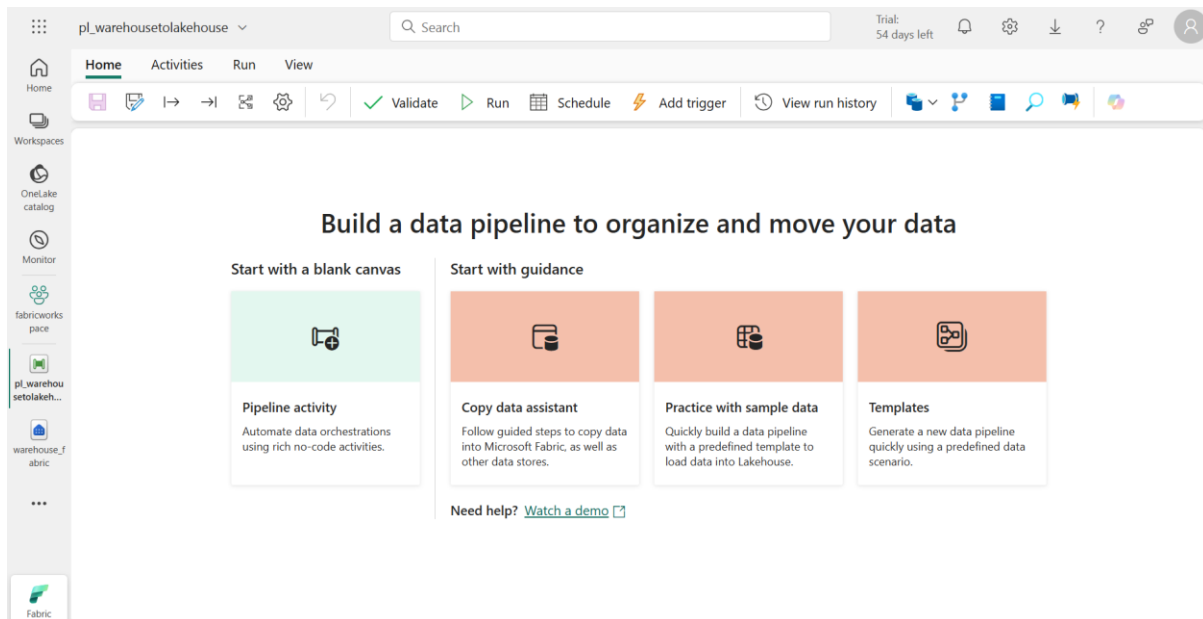
The screenshot shows a SQL query execution window. The query is: `INSERT INTO dbo.meta_table VALUES (1,'Customer','dbo','global/2025/04/07')`. The execution messages show it started at 9:00:55 AM and completed at 9:01:01 AM with a total duration of 00:00:05.287. It also indicates that 1 record was affected.

As we have only one table in warehouse only one record is inserted into meta\_data table

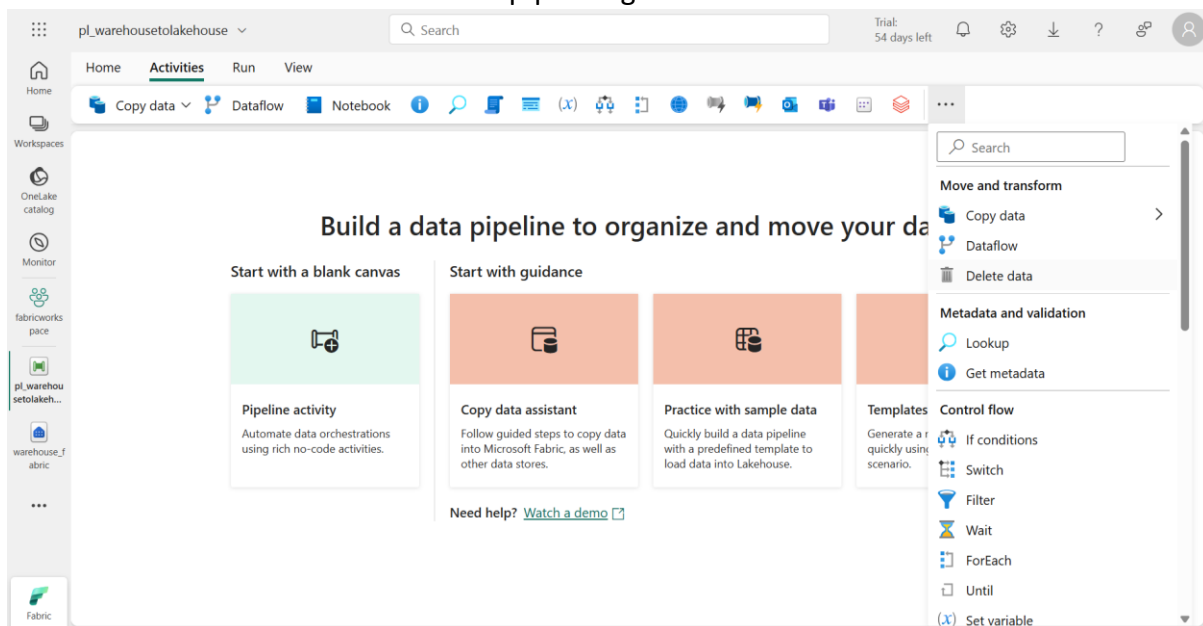
Now to create a pipeline go top workspace -> new item -> search for data pipeline -> create



The screenshot shows the Microsoft Fabric workspace interface. A 'New pipeline' dialog box is open, prompting the user to enter a name for the new pipeline. The name 'pipeline1' is entered in the text field. The background shows a list of items in the workspace, including 'Fabric\_lakehouse' and 'warehouse\_fabric'.



To check what all activities we have in pipeline go to activities tab.



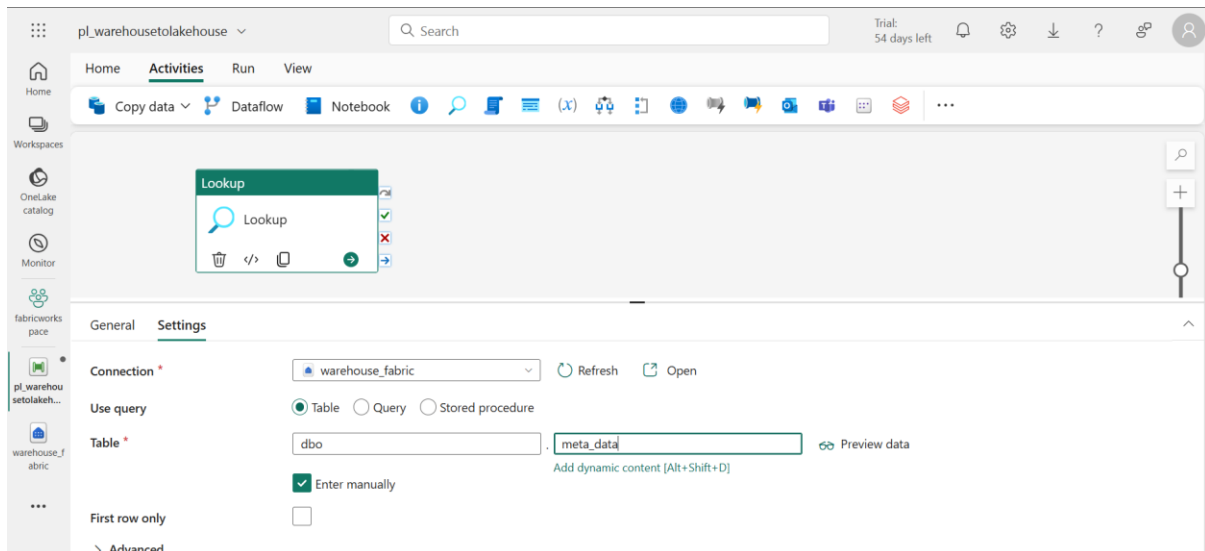
As we are building dynamic pipeline, we need lookup activity, foreach activity and copy activity.

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.

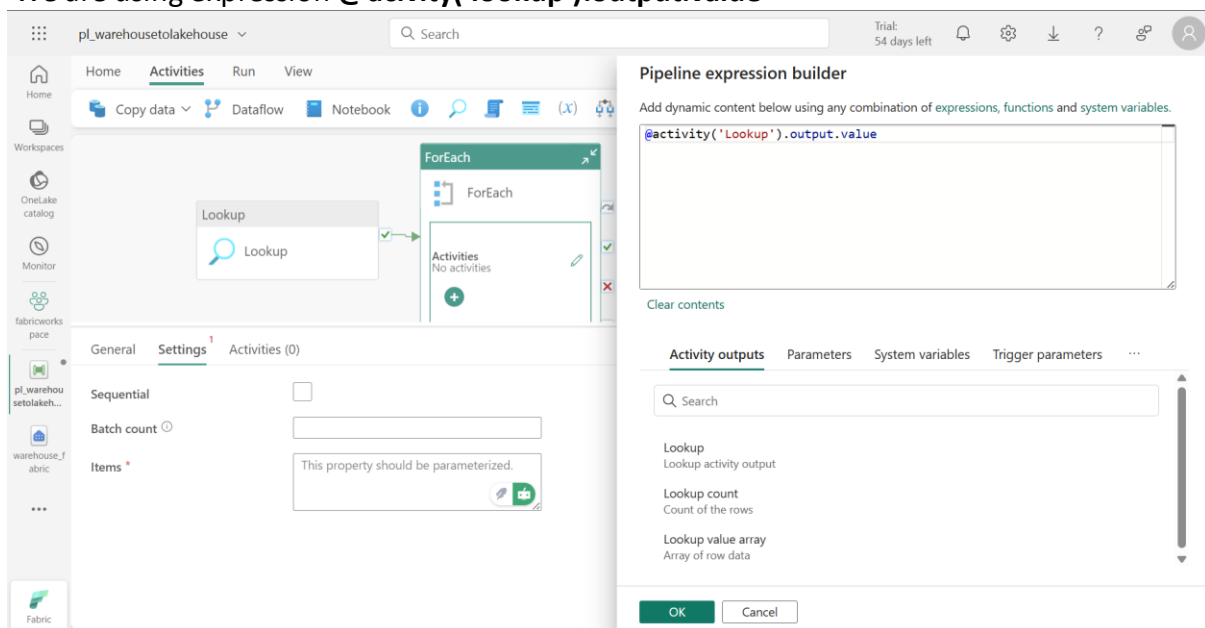


Select foreach activity

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 **@activity('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

pl\_warehouse1lakehouse

Home Activities Run View

Copy data Dataflow Notebook

Main canvas > ForEach

Copy data

Copy data

General Source Destination Mapping Settings

Connection \* warehouse\_fabric Refresh

Use query Table Query Stored procedure

Table \* schema name table name

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

pl\_warehouse1lakehouse

Home Activities Run View

Copy data Dataflow Notebook

Main canvas > ForEach

Copy data

Copy data

General Source Destination Mapping Settings

Connection \* warehouse\_fabric Refresh

Use query Table Query Stored procedure

Table \* @item().Schemaname table name

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

pl\_warehouse1lakehouse

Home Activities Run View

Copy data Dataflow Notebook

Main canvas > ForEach

Copy data

Copy data

General Source Destination Mapping Settings

Connection \* warehouse\_fabric Refresh Open

Use query Table Query Stored procedure

Table \* @item().Schemaname @item().Tablename Preview data

Enter manually

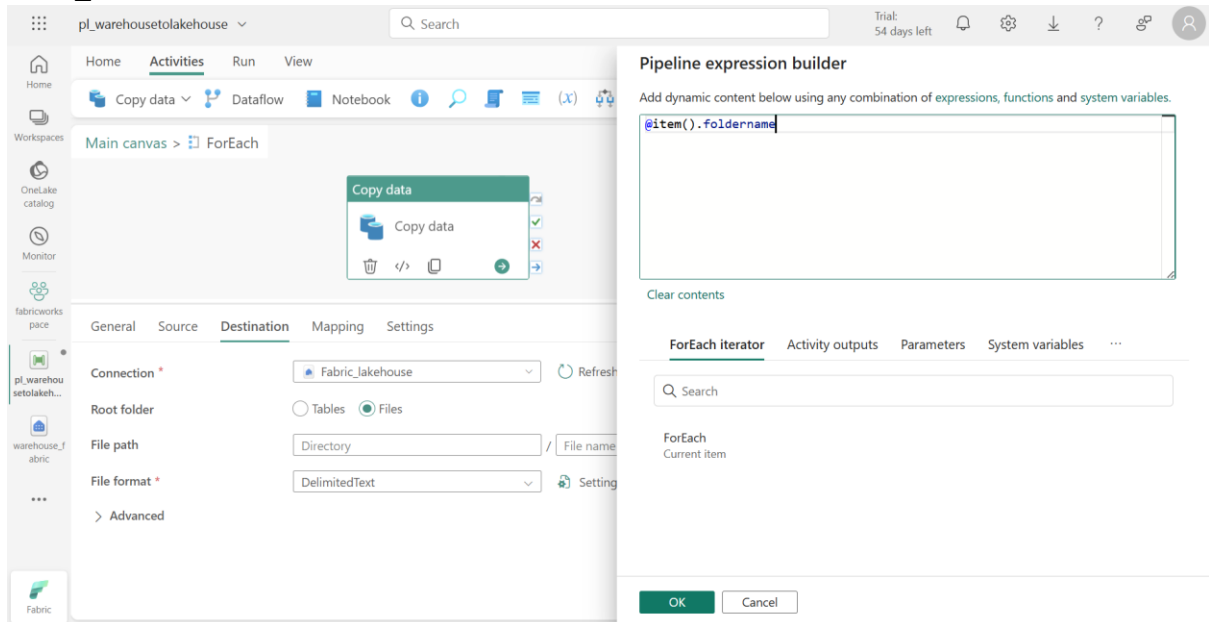
Advanced

Go to destination

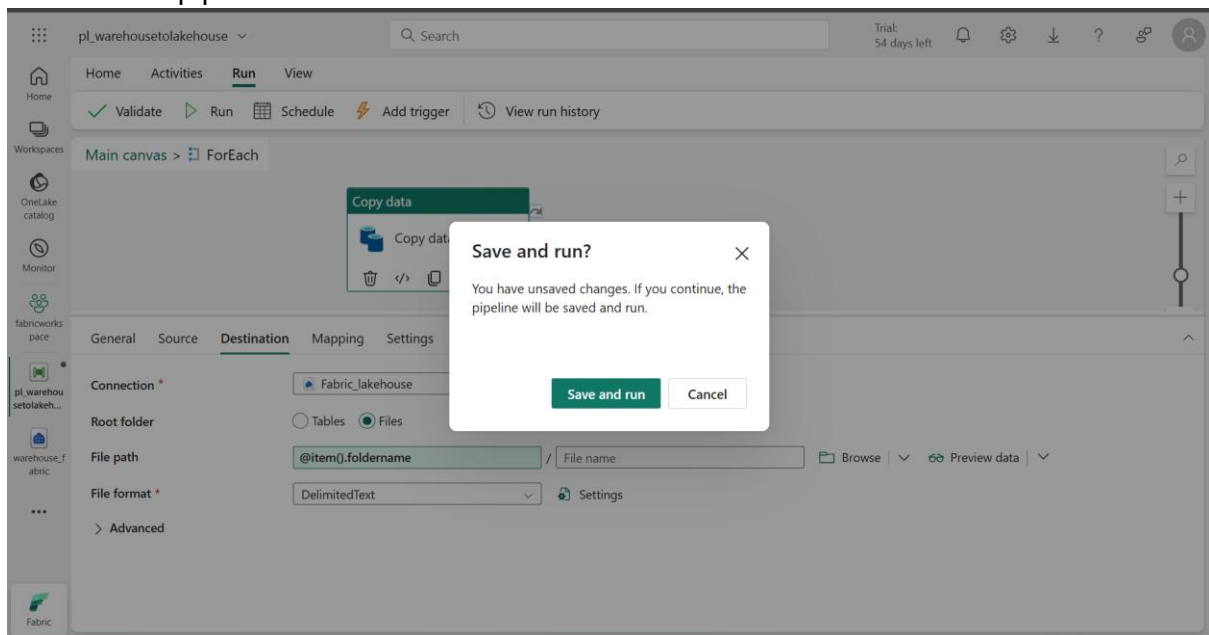
Select lakehouse in connection

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



pl\_warehouseetolakehouse

Home Activities **Run** View

Validate Run Schedule Add trigger View run history

Run Succeeded  
Successfully ran pl\_warehouseetolakehouse...

Lookup  
ForEach

Parameters Variables Settings **Output**

Pipeline run ID: 7ea72a0a-ceaf-470d-8976-a0d1ce4410d0 Pipeline status: Succeeded View run detail Export to CSV

Filter by keyword Showing 3 items

Activity name	Activity status	Run start	Duration	Input	Output
Lookup	Succeeded	4/7/2025, 9:25:45 AM	21s	->	->
ForEach	Succeeded	4/7/2025, 9:26:07 AM	29s	->	->
Copy data	Succeeded	4/7/2025, 9:26:08 AM	25s	->	->

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

Fabric\_lakehouse

Home

Get data New semantic model Open notebook Manage OneLake data access (preview)

A SQL analytics endpoint for SQL querying and a default Power BI semantic model for reporting were created with this item.

Explorer

Search tables

Fabric\_lakehouse

- Tables
- Files
  - global
    - 2025
      - 04
        - 07

Files > global > 2025 > 04 > 07

Search files

Name	Date modified	Type	Size
dbo.Customer.csv	4/7/2025, 9:26:...	csv	180 B

Data is stored successfully

Fabric\_lakehouse

Home

Get data New semantic model Open notebook Manage OneLake data access (preview)

A SQL analytics endpoint for SQL querying and a default Power BI semantic model for reporting were created with this item.

Explorer

Search tables

Fabric\_lakehouse

- Tables
- Files
  - global
    - 2025
      - 04
        - 07

... > global > 2025 > 04 > 07 > dbo.Customer.csv (preview)

```
1 ID,Name,City,Phoneno
2 5,"Sam","Brampton",4567824567
3 4,"leo","Sudbury",4567824567
4 3,"Naho","Paris",4567824567
5 2,"Robert","Toronto",4567824567
6 1,"John","New york",4567824567
7
```

## Explore the short cuts feature

### Shortcut in Microsoft Fabric

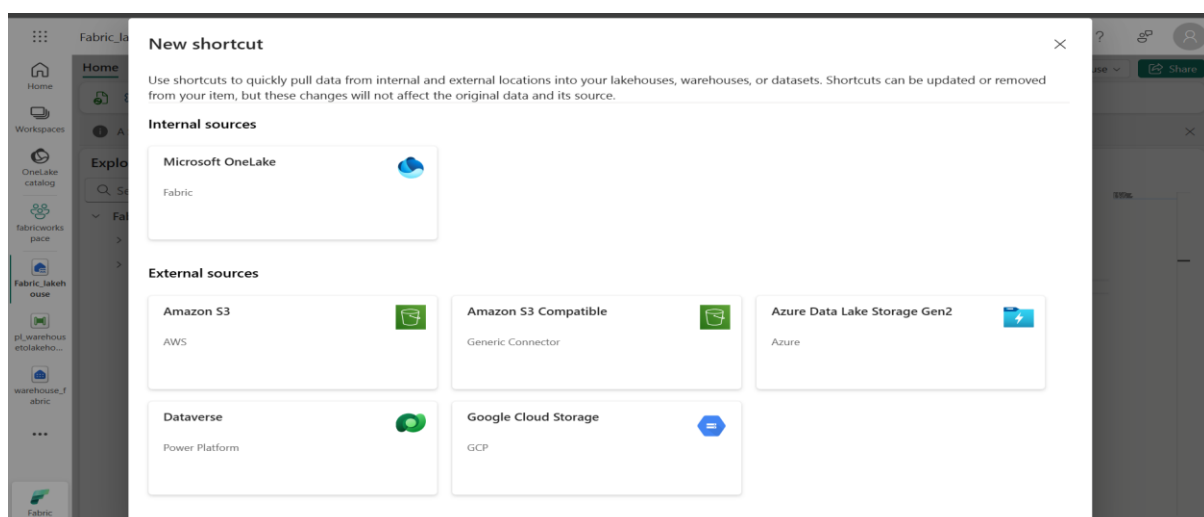
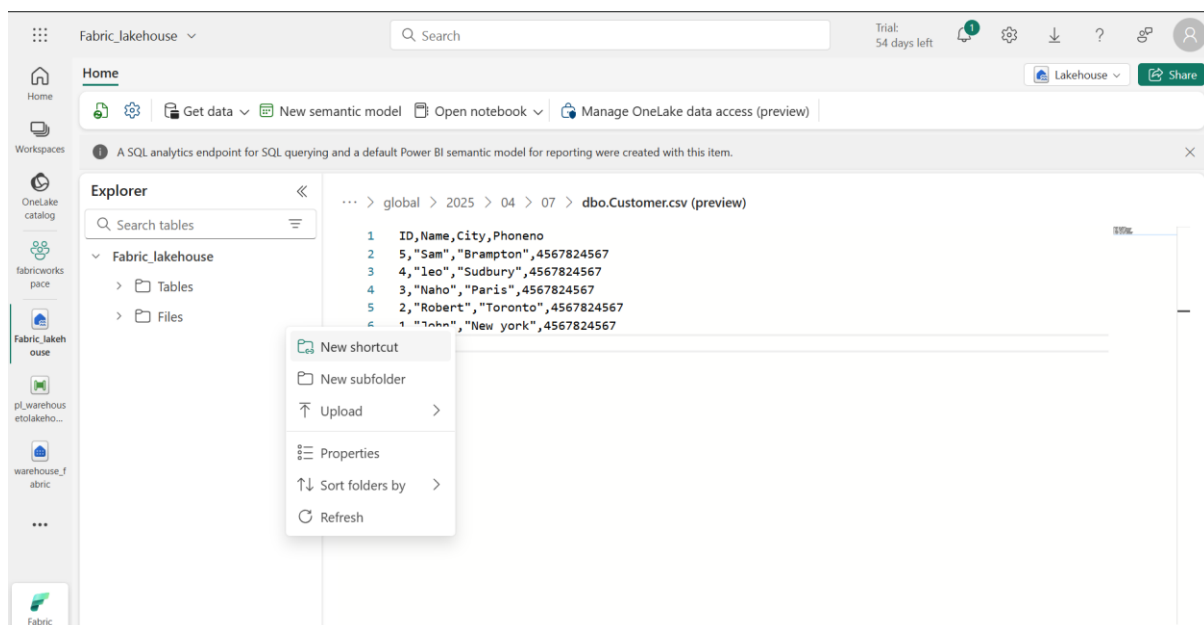
A *Shortcut* in Microsoft Fabric is a reference pointer that allows users to access data stored in other OneLake locations, such as different workspaces, without physically duplicating the data. It enables seamless data sharing across domains, improves collaboration, and optimizes storage by creating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

#### Key Benefits:

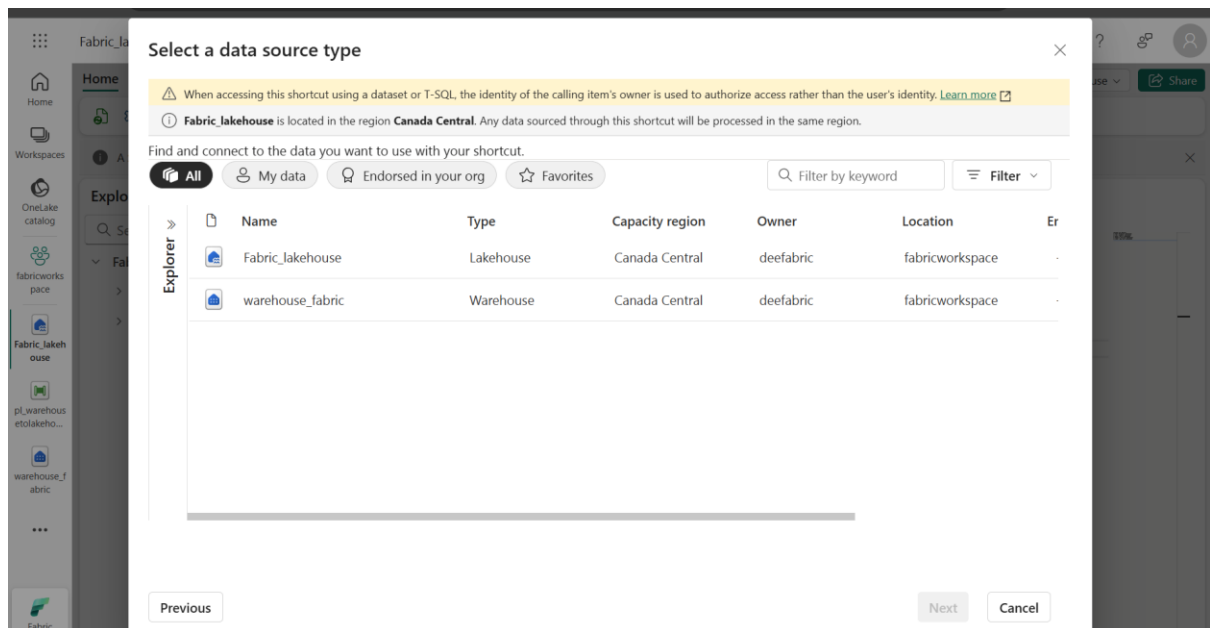
- Avoids data duplication
- Simplifies data access across workspaces
- Enables unified analytics across domains
- Enhances governance and control by maintaining a single source of truth

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

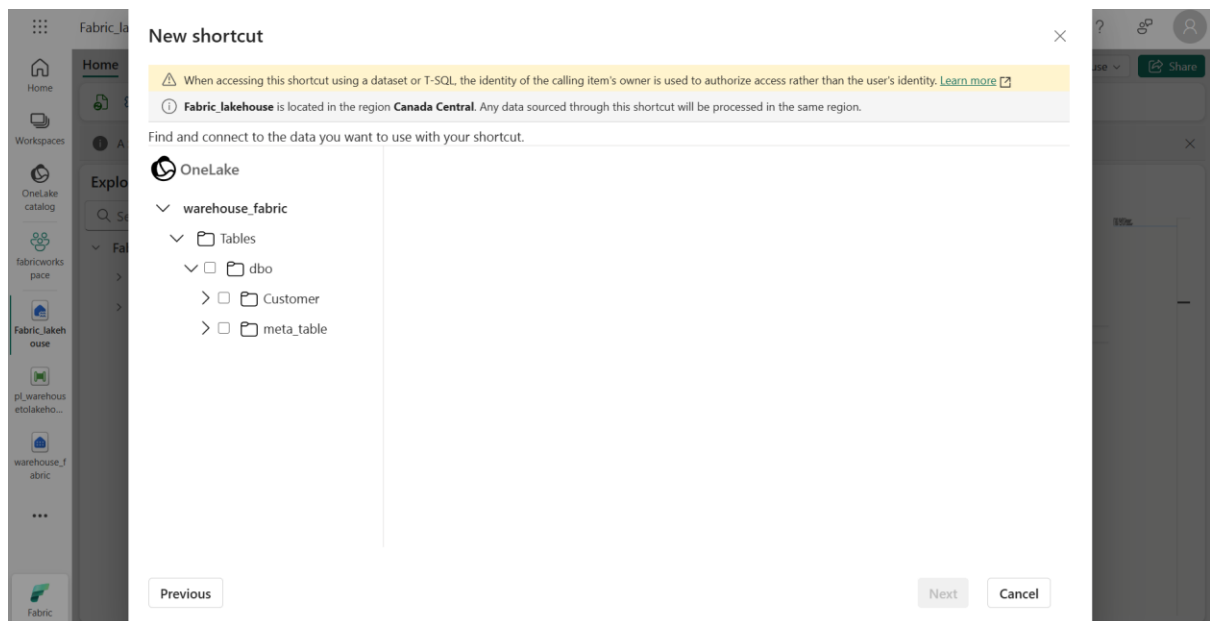
Go to lakehouse -> files /Tables -> 3 dots -> New shortcut



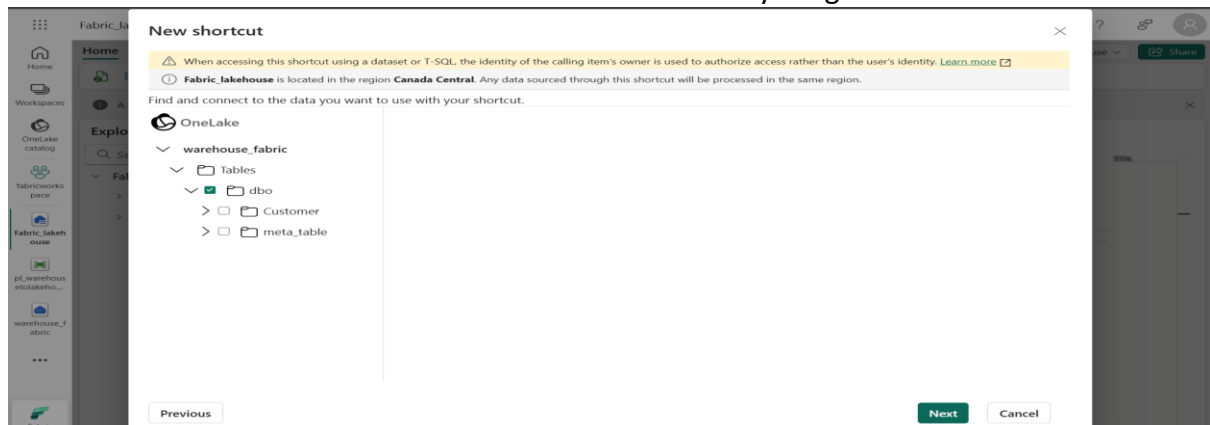
We can create a connection to everything in Onelake and to external sources as well.  
Now as we want to create a connection to warehouse select Onelake.

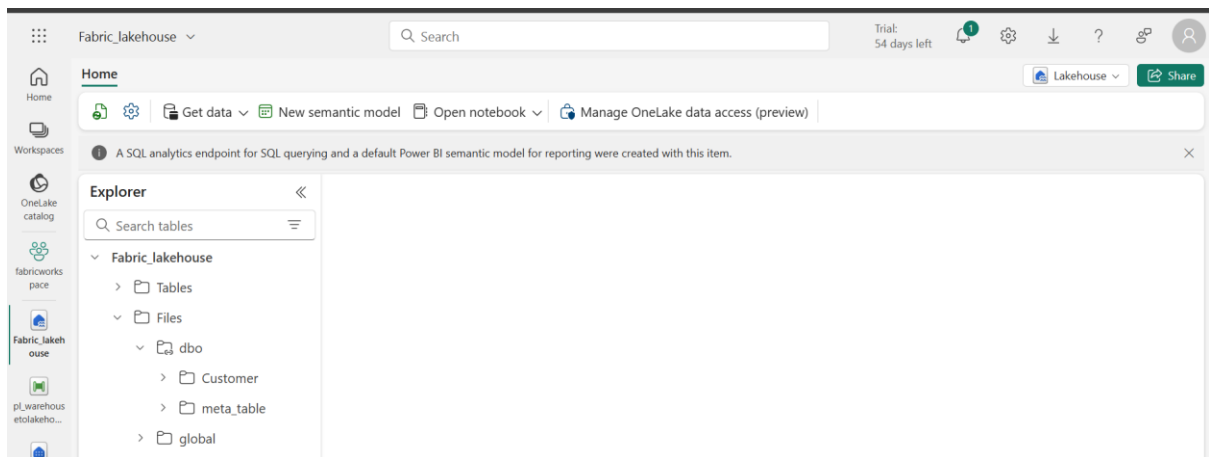


## Select warehouse



We can select what we want to connect one table or everything.





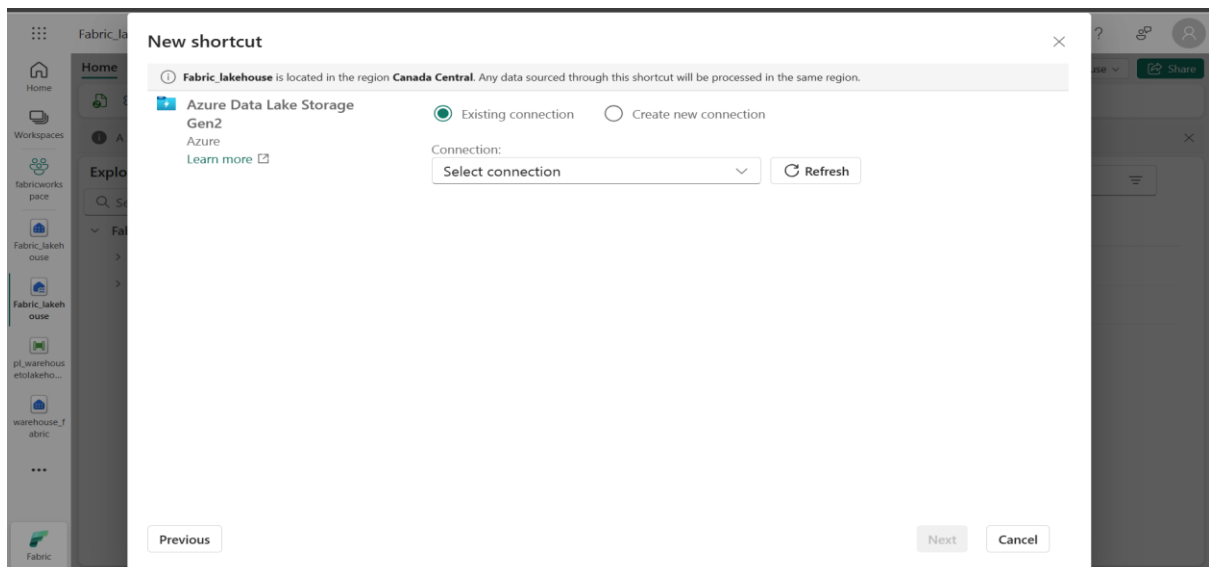
A link is created between lakehouse and Warehouse, we see tables in lakehouse without loading them manually.

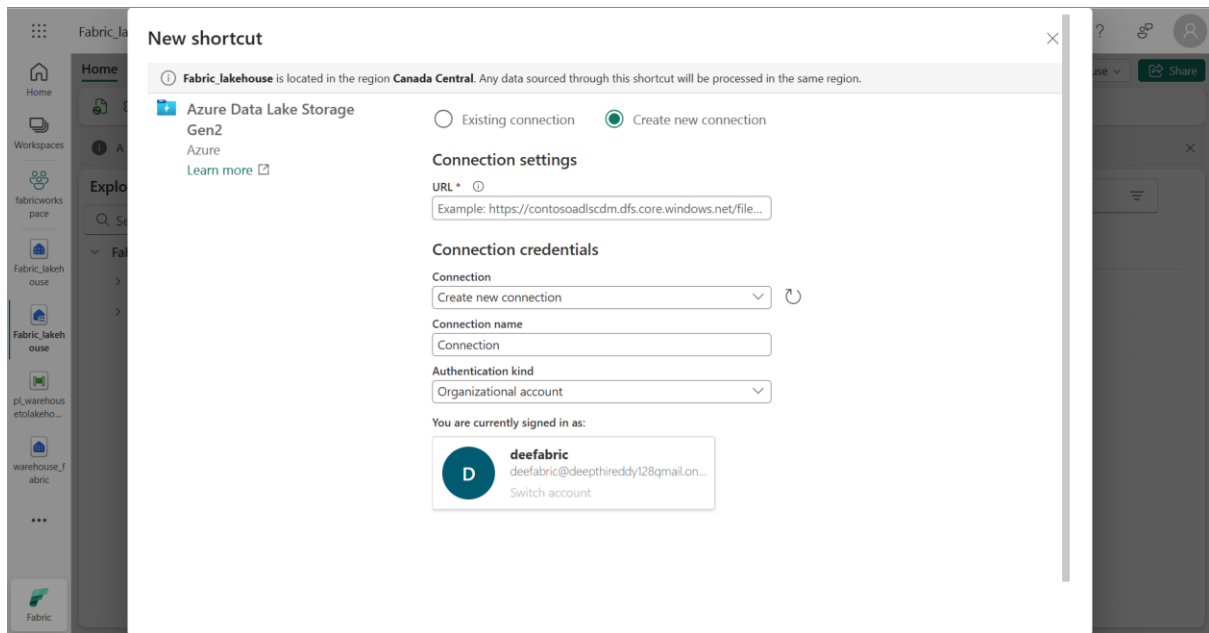
As these are tables in files, they are stored in delta format.

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

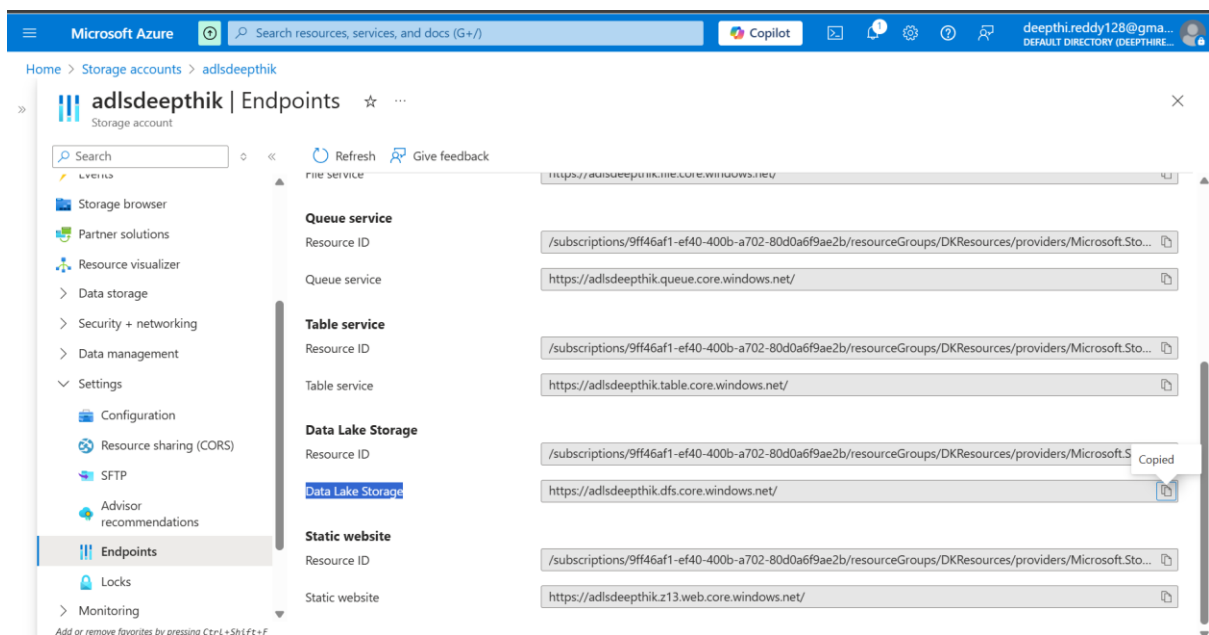
We can also create a short cut to external source, let create a shot cur between lakehouse and external Azure data lake gen 2.

As we don't have any existing connection, we need to create a new connection.

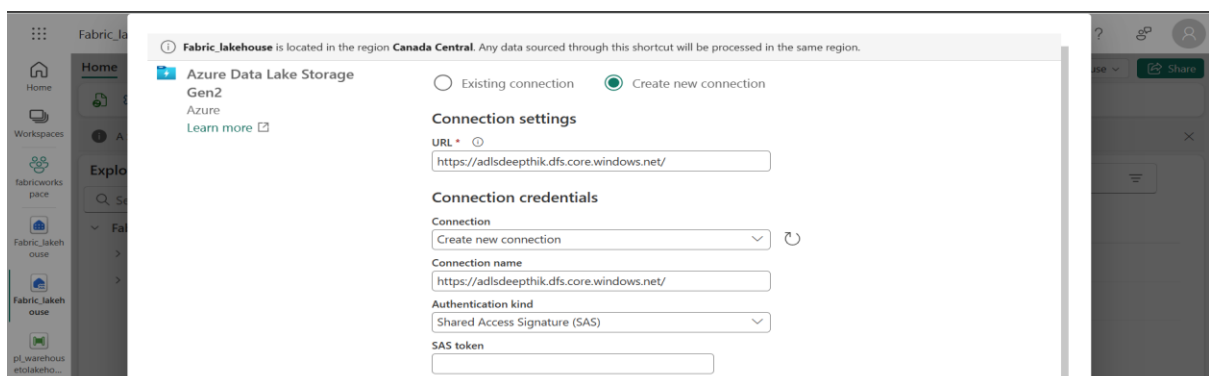




To get a connection string we need to go to Azure data lake gen 2 -> Settings-> Endpoints -> Data Lake Storage -> Data Lake Storage URL.

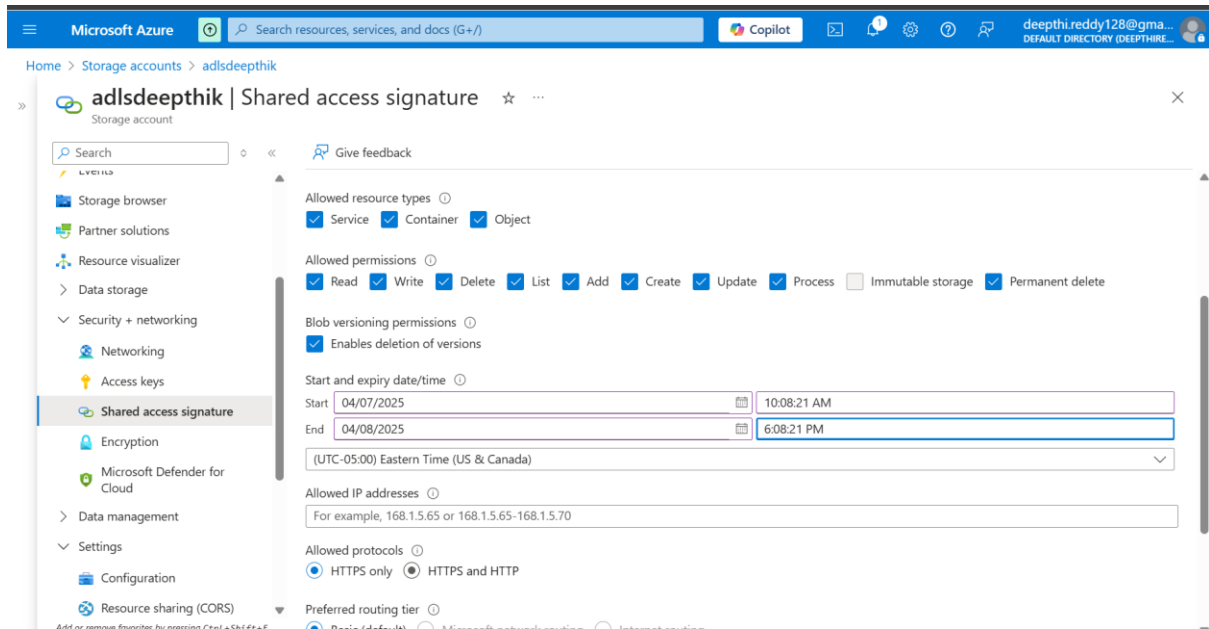


Copy that link and go to fabric

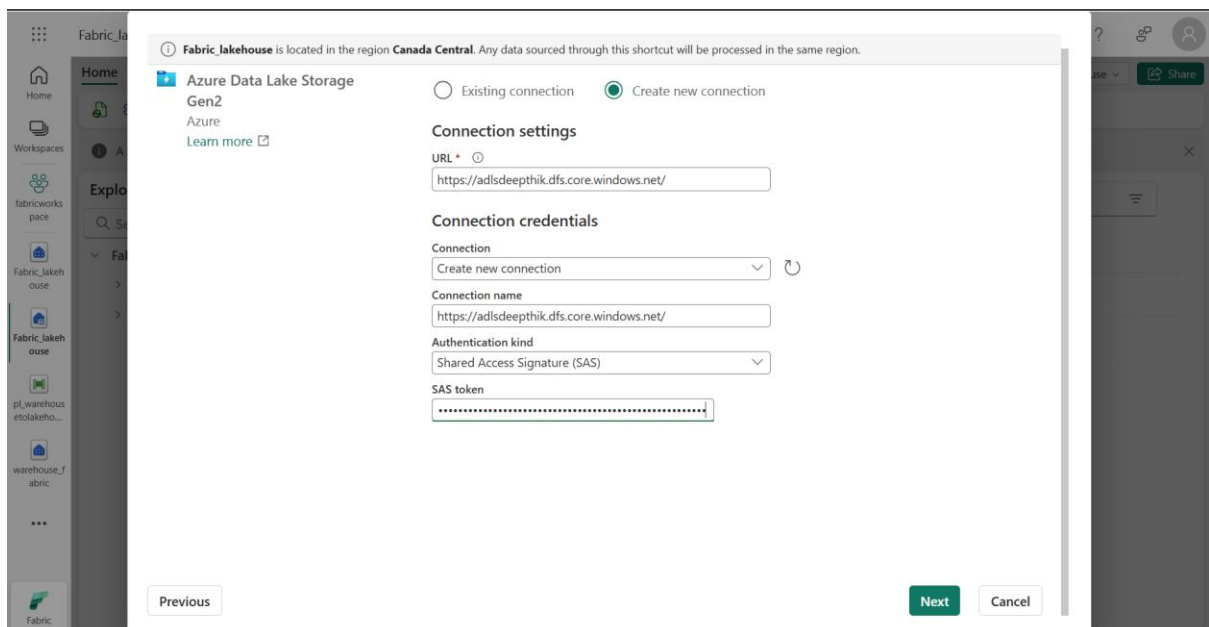


As for Authentication kind choose SAS and provide SAS token from Azure data lake gen 2 storage.

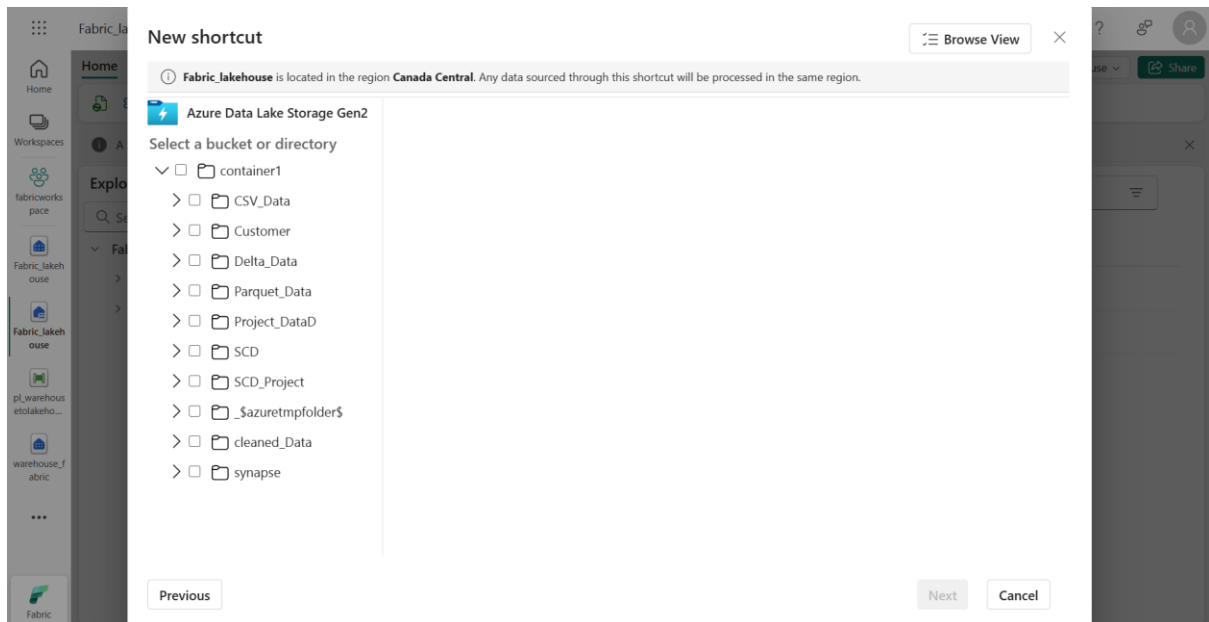
To generate SAS token, go to Azure data lake gen 2 -> Security + Networking -> Shared access signature -> Allowed resource types -> check all Service, Container and Object and generate SAS



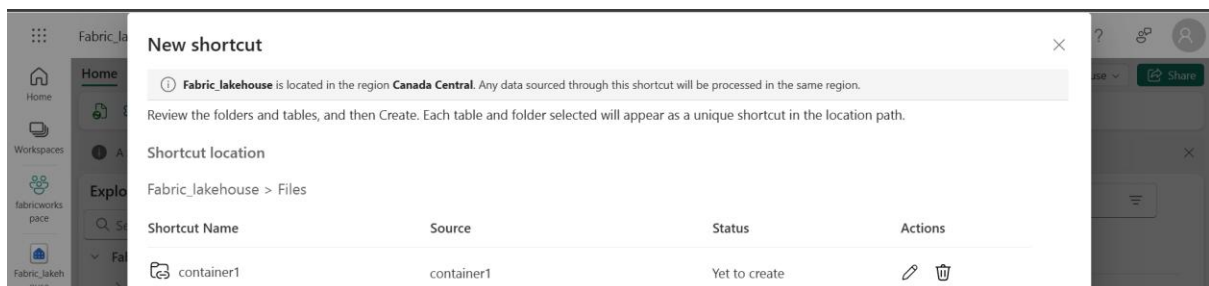
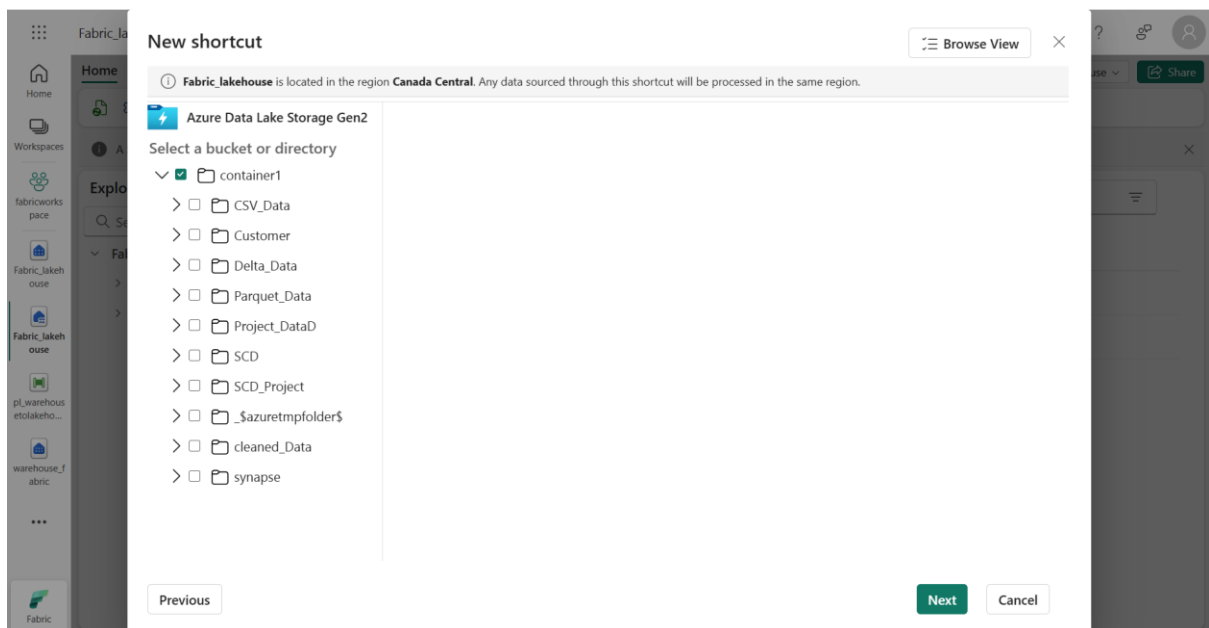
Generate token and copy in fabric.



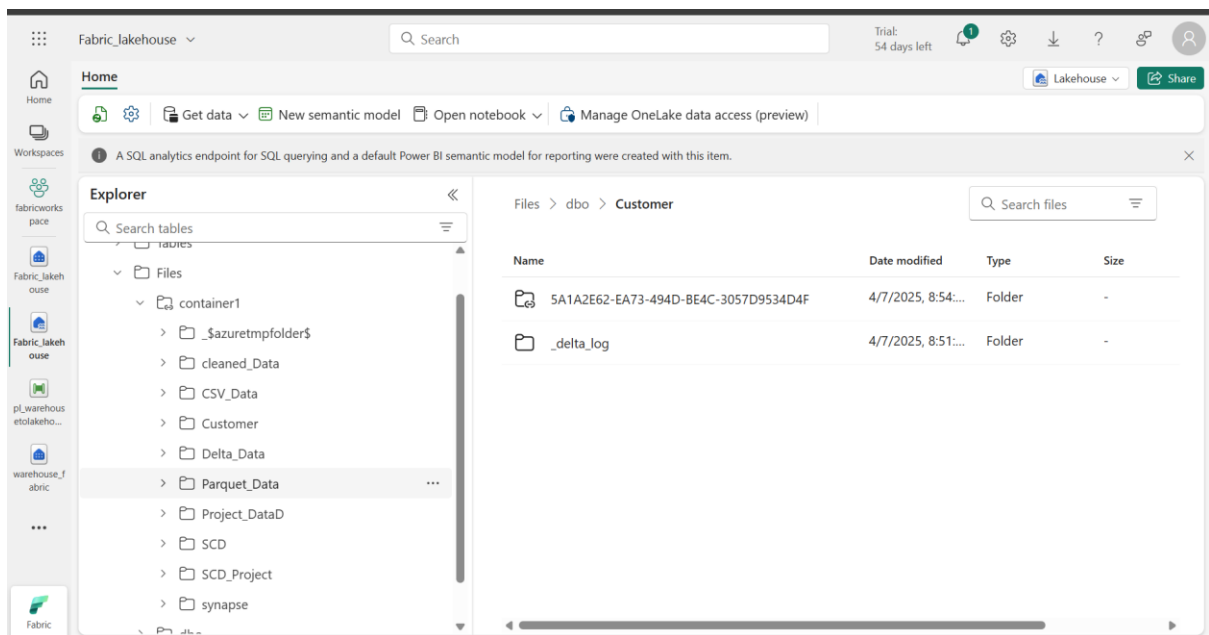
Select for what we need to create a link to container or to specific folder.



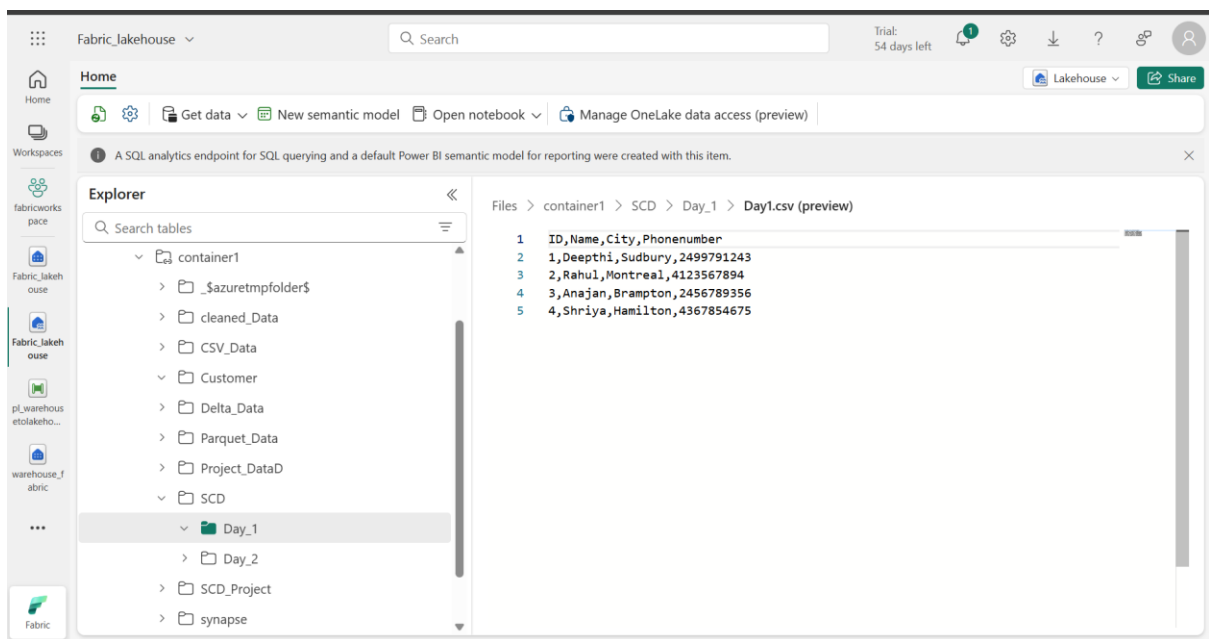
Creating link to container.



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

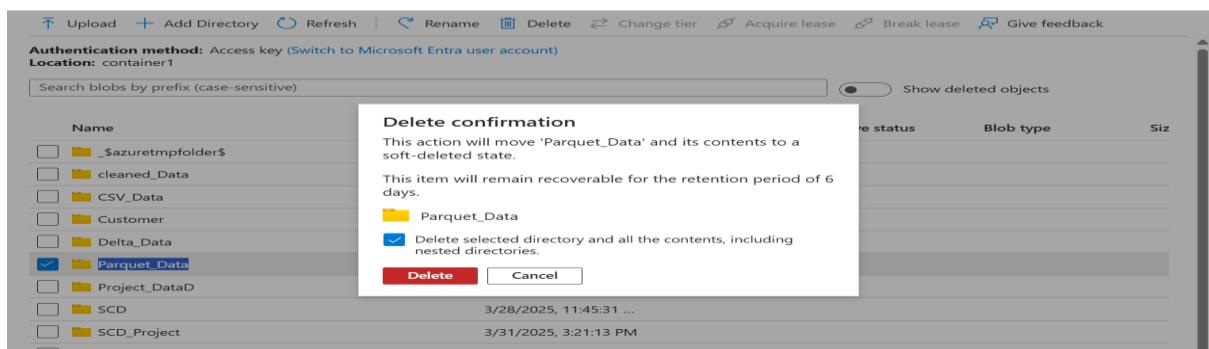


We can view the data



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.





Home > Storage accounts > adlsdeepthik | Containers >

container1  
Container

Search

Upload Add Directory Refresh Rename Delete Change tier Acquire lease Break lease Give feedback

Overview

Diagnose and solve problems

Access Control (IAM)

Settings

Authentication method: Access key (Switch to Microsoft Entra user account)  
Location: container1

Search blobs by prefix (case-sensitive)  Show deleted objects

Name	Modified	Access tier	Archive status	Blob type	Size
<input type="checkbox"/> _azuretmpfolder\$	3/31/2025, 2:36:24 PM				
<input type="checkbox"/> cleaned_Data	3/31/2025, 2:52:49 PM				
<input type="checkbox"/> CSV_Data	3/22/2025, 1:04:52 PM				
<input type="checkbox"/> Customer	3/28/2025, 12:04:19 ...				
<input type="checkbox"/> Delta_Data	3/22/2025, 1:05:07 PM				
<input type="checkbox"/> Project_DataD	3/31/2025, 2:42:56 PM				
<input type="checkbox"/> SCD	3/28/2025, 11:45:31 ...				
<input type="checkbox"/> SCD_Project	3/31/2025, 3:21:13 PM				
<input type="checkbox"/> synapse	3/31/2025, 11:18:06 ...				

Successfully deleted blob(s)  
Successfully deleted 1 blobs(s).

Folder is deleted in storage account.

Fabric\_lakehouse

Search

Trial: 54 days left

Home

Get data New semantic model Open notebook Manage OneLake data access (preview)

Workspaces

A SQL analytics endpoint for SQL querying and a default Power BI semantic model for reporting were created with this item.

Explorer

Search tables

Tables

Files

container1

\_azuretmpfolder\$

cleaned\_Data

CSV\_Data

Customer

Delta\_Data

Project\_DataD

SCD

SCD\_Project

synapse

dbo

Files > container1 > CSV\_Data

Search files

Name	Date modified	Type	Size
Credit.csv	3/26/2025, ...	csv	283 B
Employee_2025-03-11T17_07_01.8945...	3/25/2025, ...	csv	224 B
Employee_day_2.csv	3/22/2025, ...	csv	92 B

Folder is also got deleted in lakehouse.

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