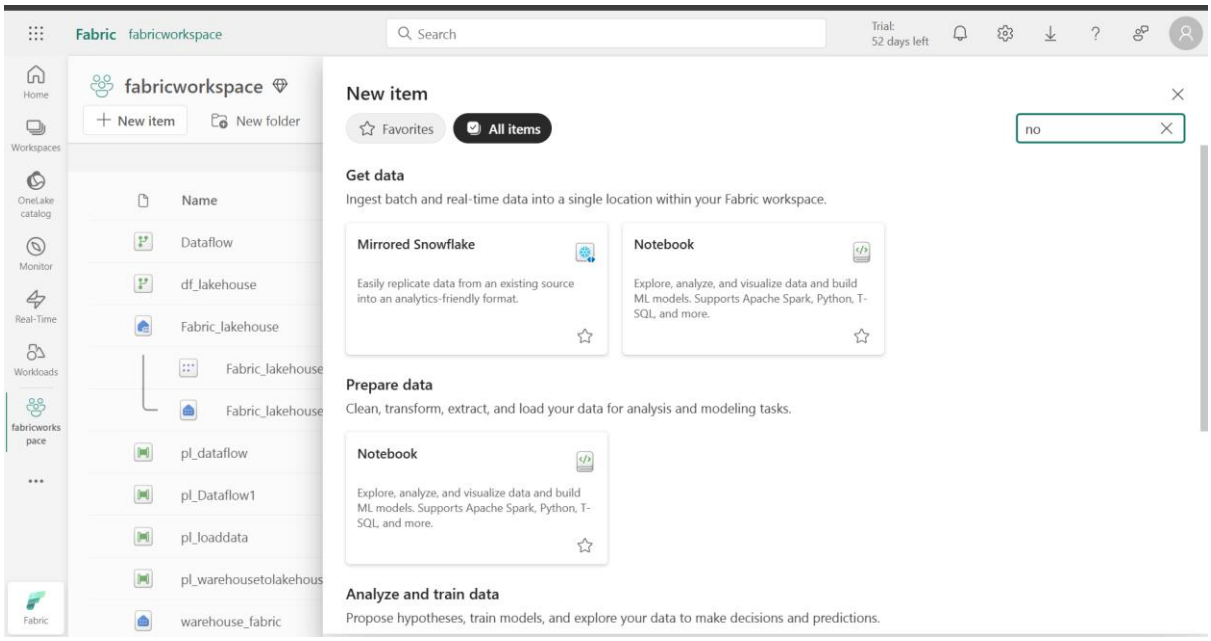
Fabric Notebook

Explore the fabric Notebooks

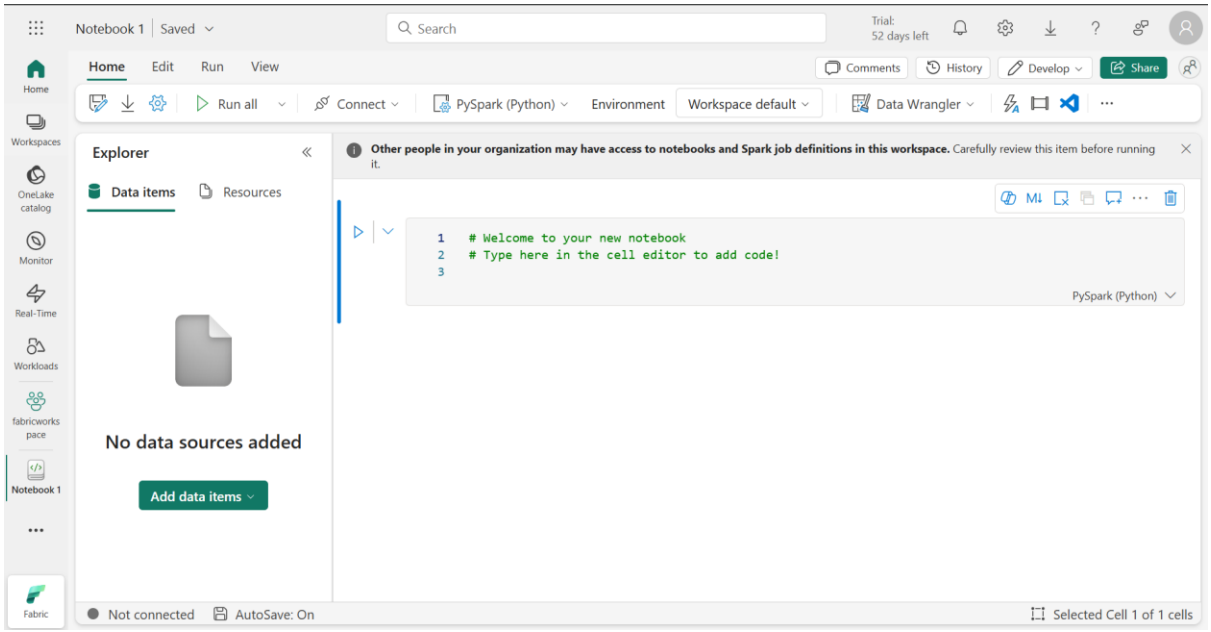
Notebook:

A **Fabric Notebook** is a feature within Microsoft Fabric that enables users to perform data exploration, transformation, and analysis using interactive code cells. It supports multiple programming languages, including **PySpark, SQL, and SparkR**, and runs on **Apache Spark** clusters provided by Fabric.

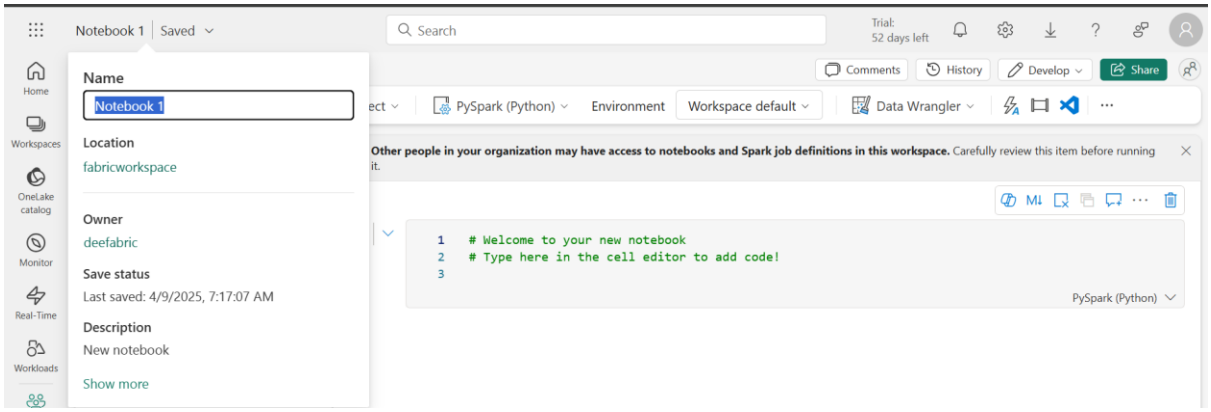
To create a notebook in fabric we need to go to workspace-> new item -> notebook

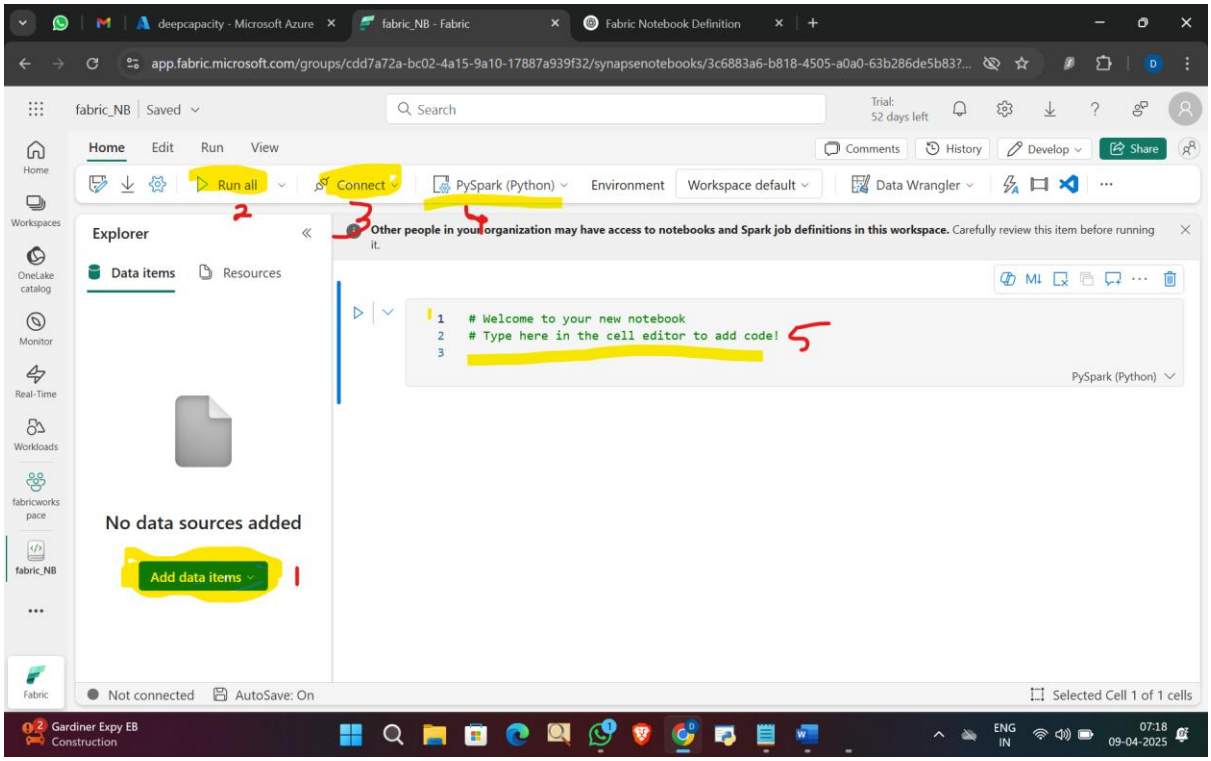


Notebook is created.



Rename the notebook from top left corner.





1. Add data items, here we will add our source
2. BY using Run all button we can run all code cell.
3. Connect, here we will connect our pool, it can be a starter pool(default) or customized pool.
4. Here we can select language as of now we are using Pyspark.
5. This is where we write code.

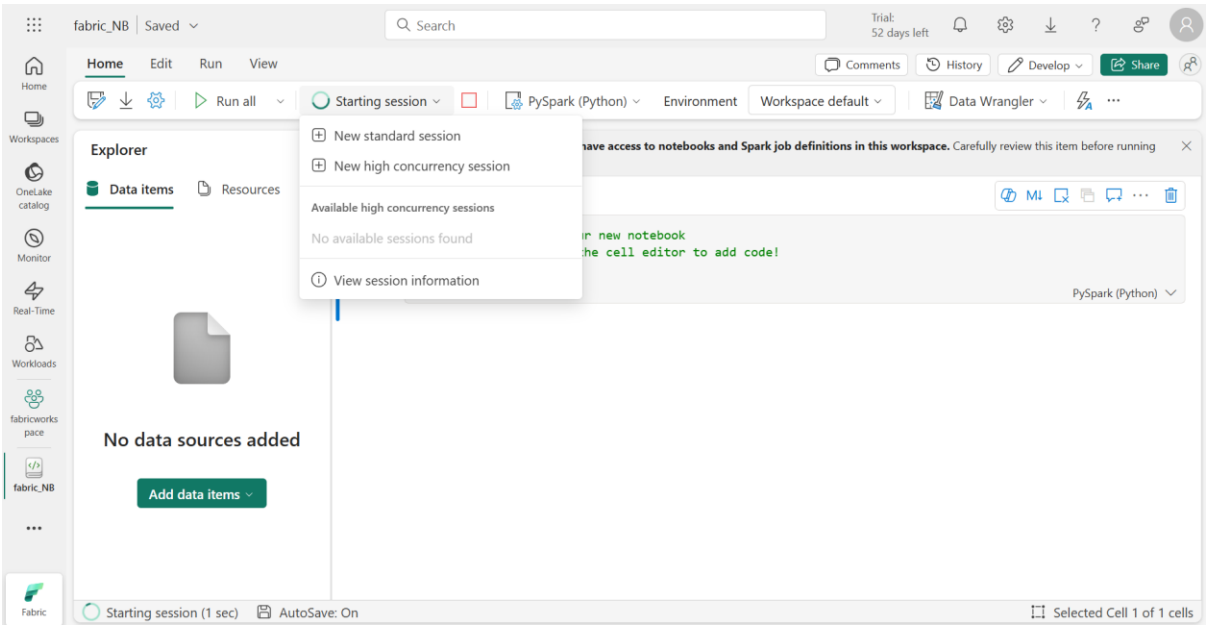
Starter pool:

A **Starter Pool** in Microsoft Fabric is a **pre-warmed Apache Spark compute cluster** that is **automatically available** to users for quick execution of notebooks without the need to create and manage custom Spark pools.

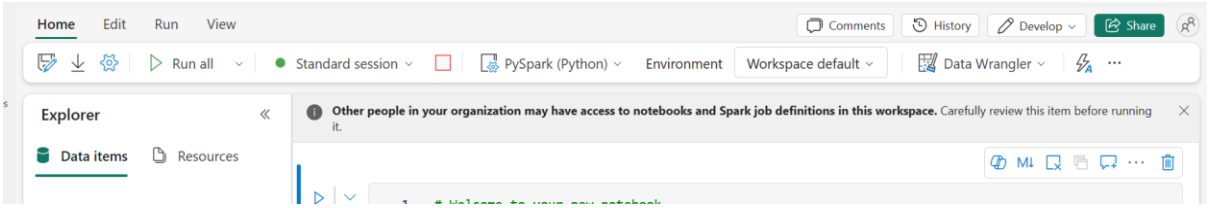
Key Characteristics:

- **No manual setup required** – It's ready to use by default in Microsoft Fabric.
- **Optimized for speed** – Because it's pre-warmed, notebooks start faster compared to cold Spark clusters.

- **Cost-efficient** – Good for small to medium workloads and exploratory tasks.
- **Auto-scaling** – Adjusts compute resources as needed for the job.

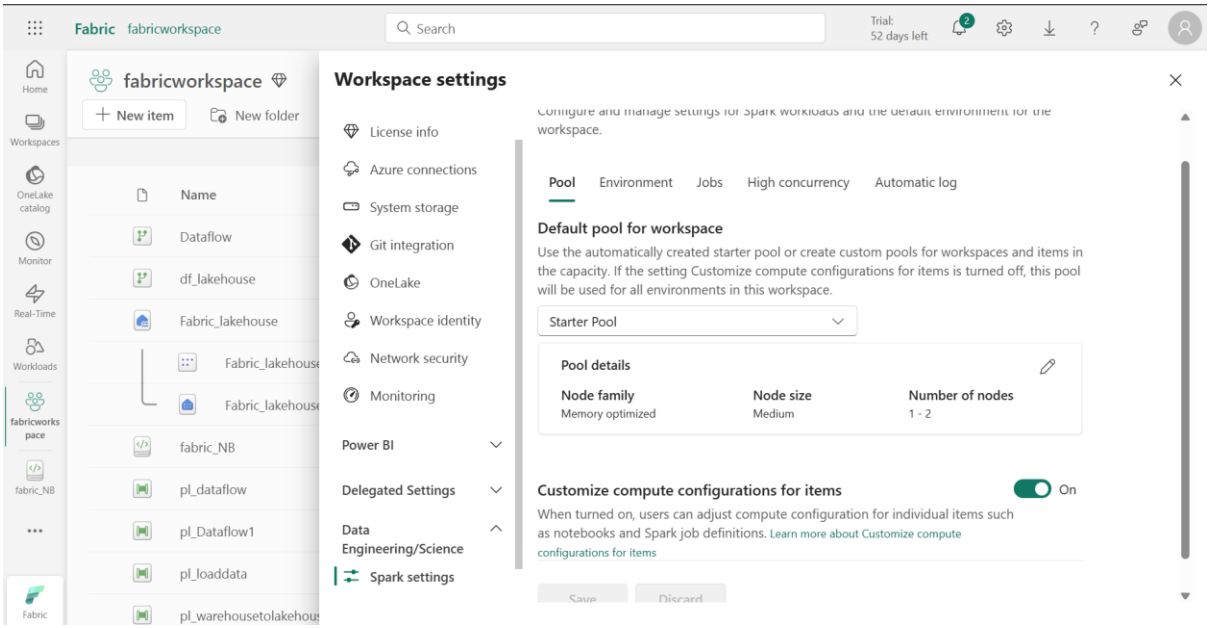


Click on New Standard Session to start the Starter pool.

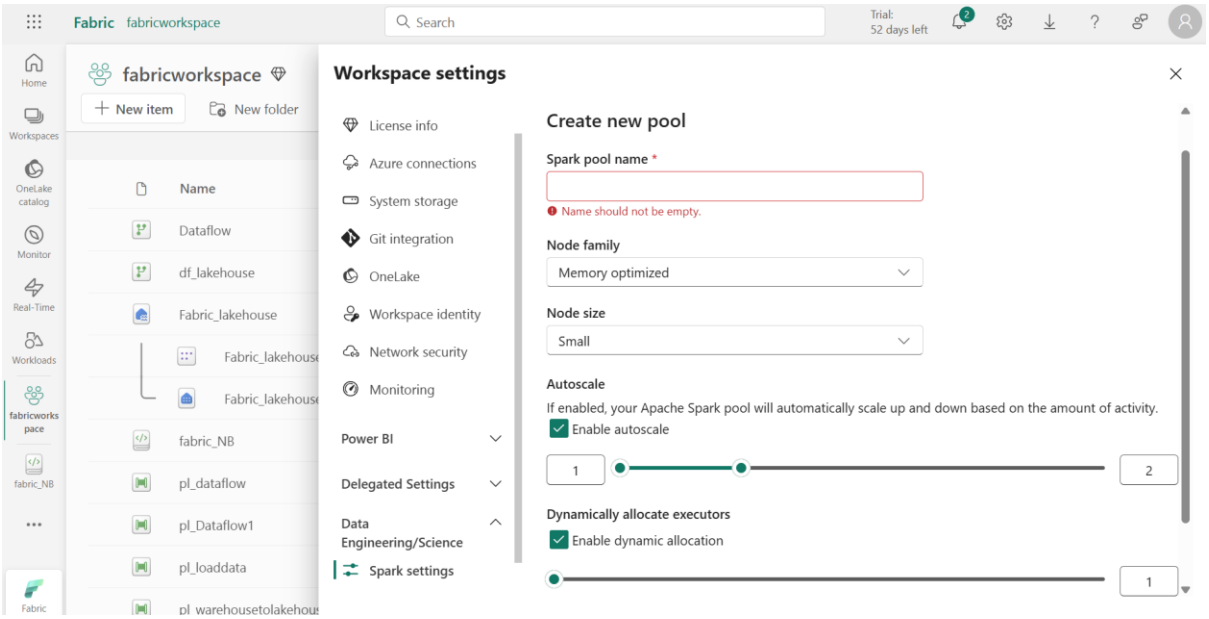


It is started, it takes only few seconds for starter pool to be up and running, and once it starts running billing will start.

To create a custom pool, we need to go to fabric workspace -> workspace settings-> Data engineering -> Spark pool

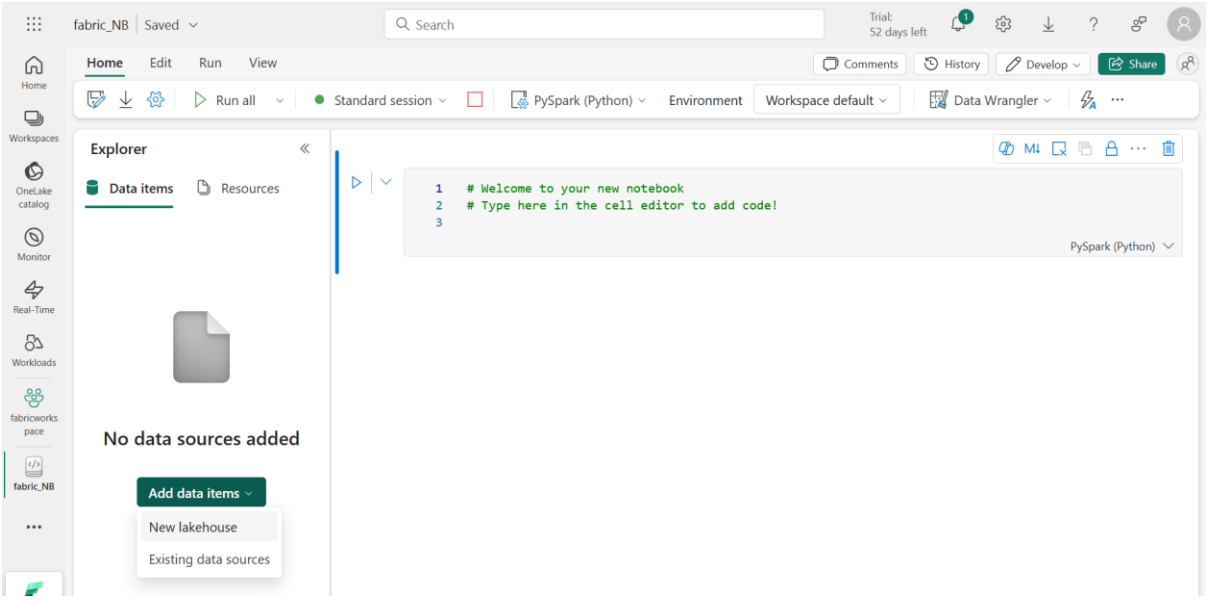


Here we can add our custom pool.

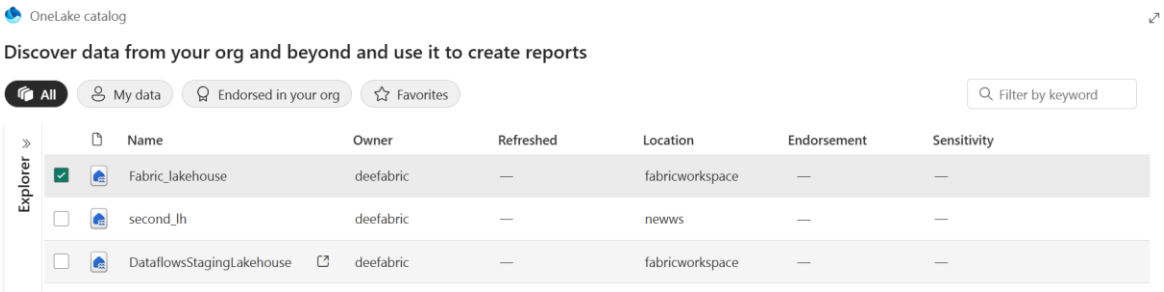


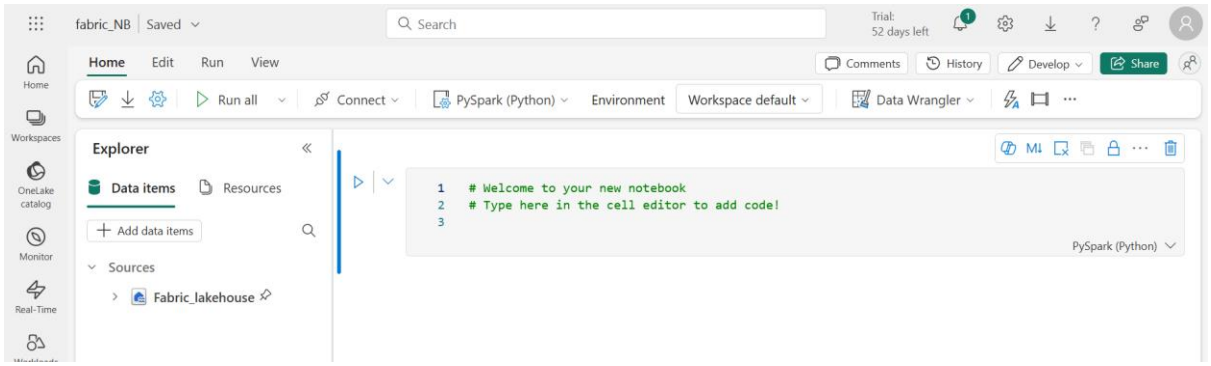
Load data from lake house to warehouse

First, we will add source as lake house to notebook.



Go to existing data source as we already have lake house in our workspace.



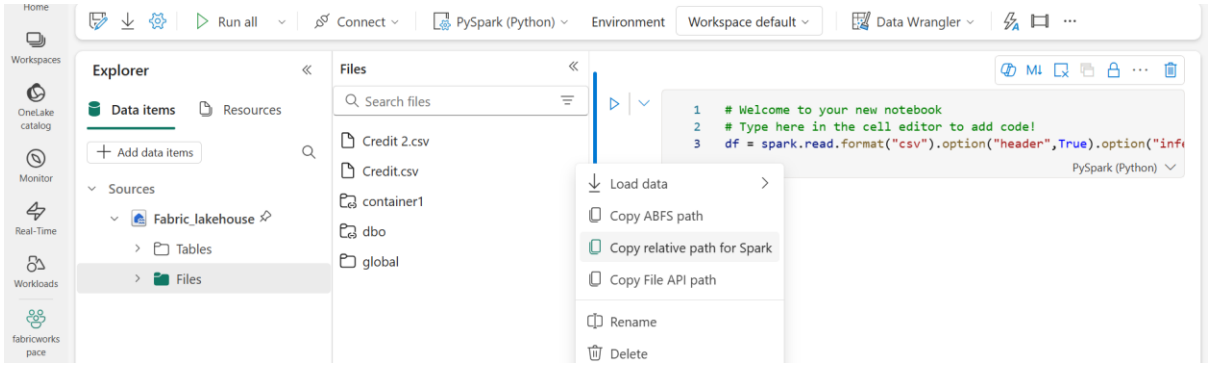


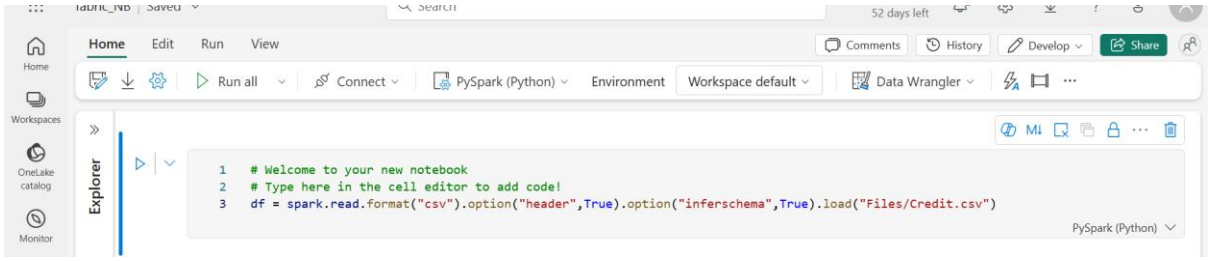
Source added.

Now we have to read data from lake house to a data frame, using below code.

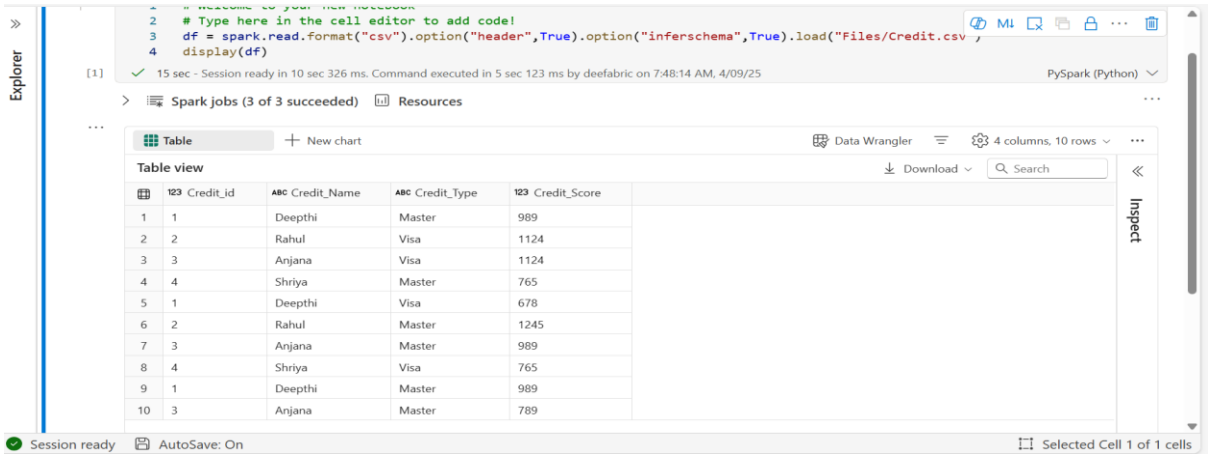
Df = spark.read.format(“File Format”).options(“Header”,True).options(“inferschema”,True).load(“Path where to load”)

To get file path go to file -> click on three dots -> Copy relative file path for spark.





Add display(df) code and run this cell.

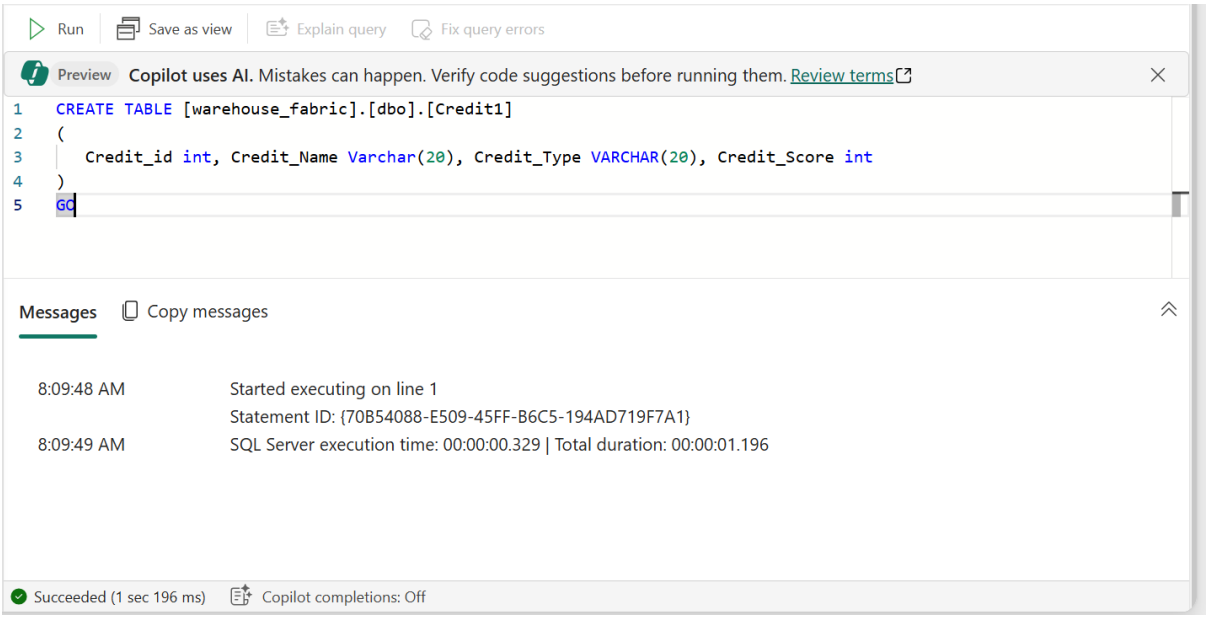


	Credit_id	Credit_Name	Credit_Type	Credit_Score
1	1	Deepthi	Master	989
2	2	Rahul	Visa	1124
3	3	Anjana	Visa	1124
4	4	Shriya	Master	765
5	1	Deepthi	Visa	678
6	2	Rahul	Master	1245
7	3	Anjana	Master	989
8	4	Shriya	Visa	765
9	1	Deepthi	Master	989
10	3	Anjana	Master	789

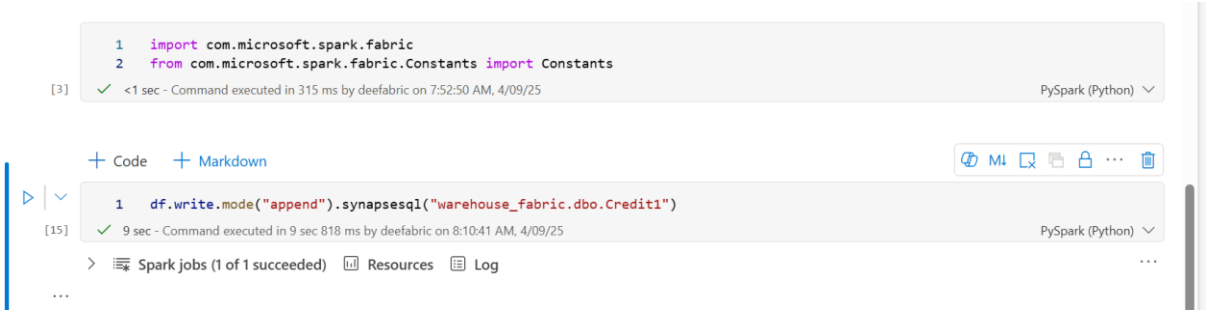
Code ran successfully.

Now write this data to datawarehouse.

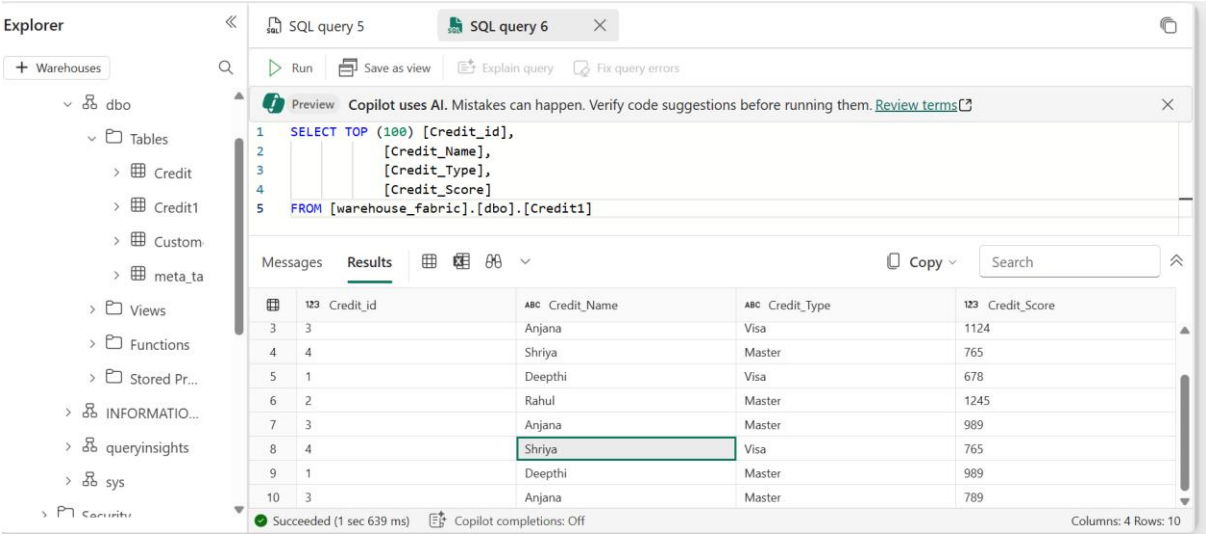
As we are loading data to warehouse we need to first create a table with same schema of lake house file.



Now we can write the data to warehouse, using below code



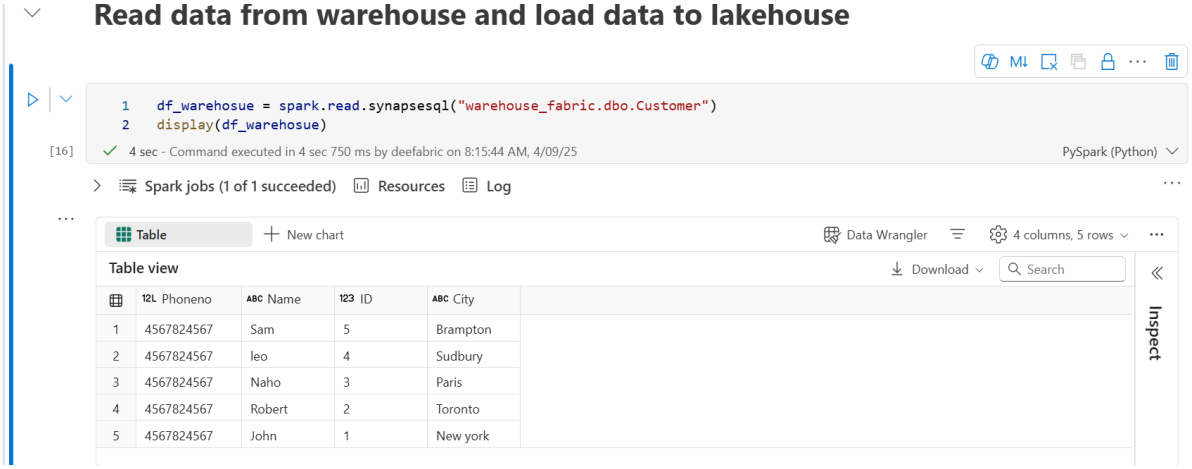
Now go to warehouse and check if the data is loaded correctly or not.



We see that data is loaded correctly to warehouse from Lake house.

Read data from warehouse and load data to lake house

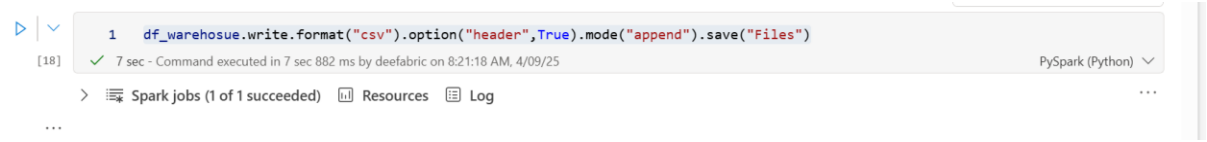
Now we have to read a table from warehouse, using below code.
Here I am reading table **warehouse_fabric.dbo.Customer**



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

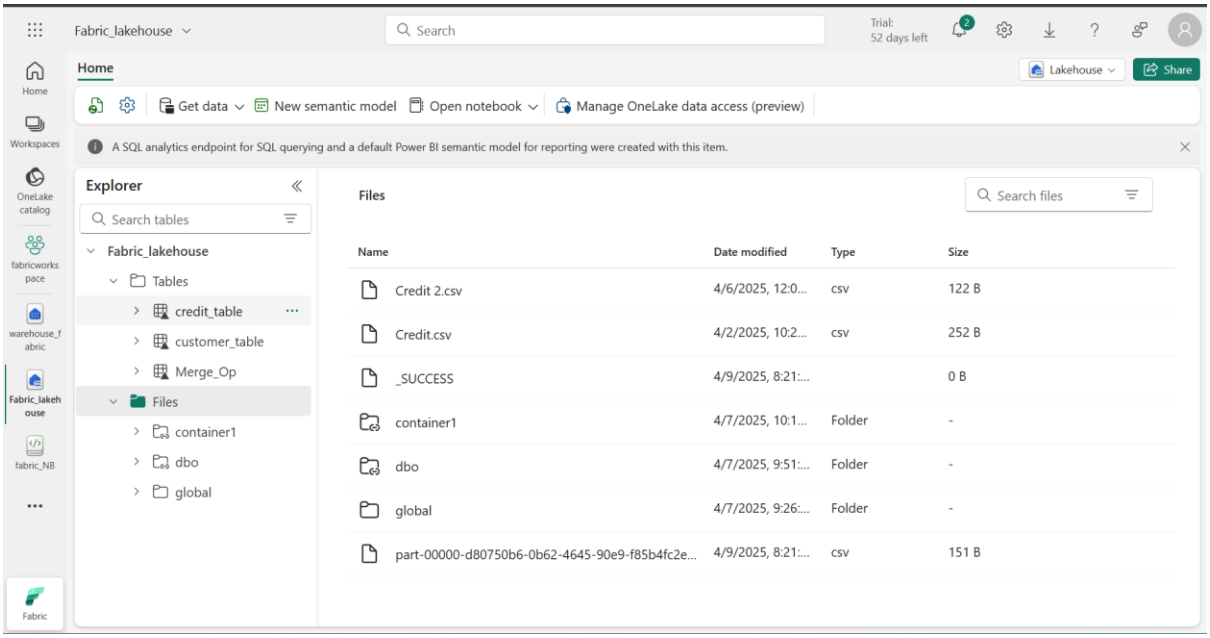
Data read successfully.

Now write this data to lakehouse, we need to use below code
df_warehosue.write.format("csv").option("header",True).mode("append").save("Files")
df_warehouse is a data frame which contains warehouse data
we are trying to save as csv file and saving this file to location file.



Code ran successfully.

Now check in lakehouse if file loaded correctly or not.



New file is loaded to lakehouse

Files > part-00000-d80750b6-0b62-4645-90e9-f85b4fc2e6ff-c000.csv (preview)

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

File data.

We can also specify the filename in path.

```
df_warehosue.write.format("csv").option("header",True).mode("append").save("Files/new_data.csv")
```

✓ 6 sec - Command executed in 6 sec 171 ms by deefabric on 8:23:23 AM, 4/09/25 PySpark (Python)

container1	4/7/2025, 10:1...	Folder	-
dbo	4/7/2025, 9:51:...	Folder	-
global	4/7/2025, 9:26:...	Folder	-
new_data.csv	4/9/2025, 8:23:...	Folder	-

It is created in lake house.