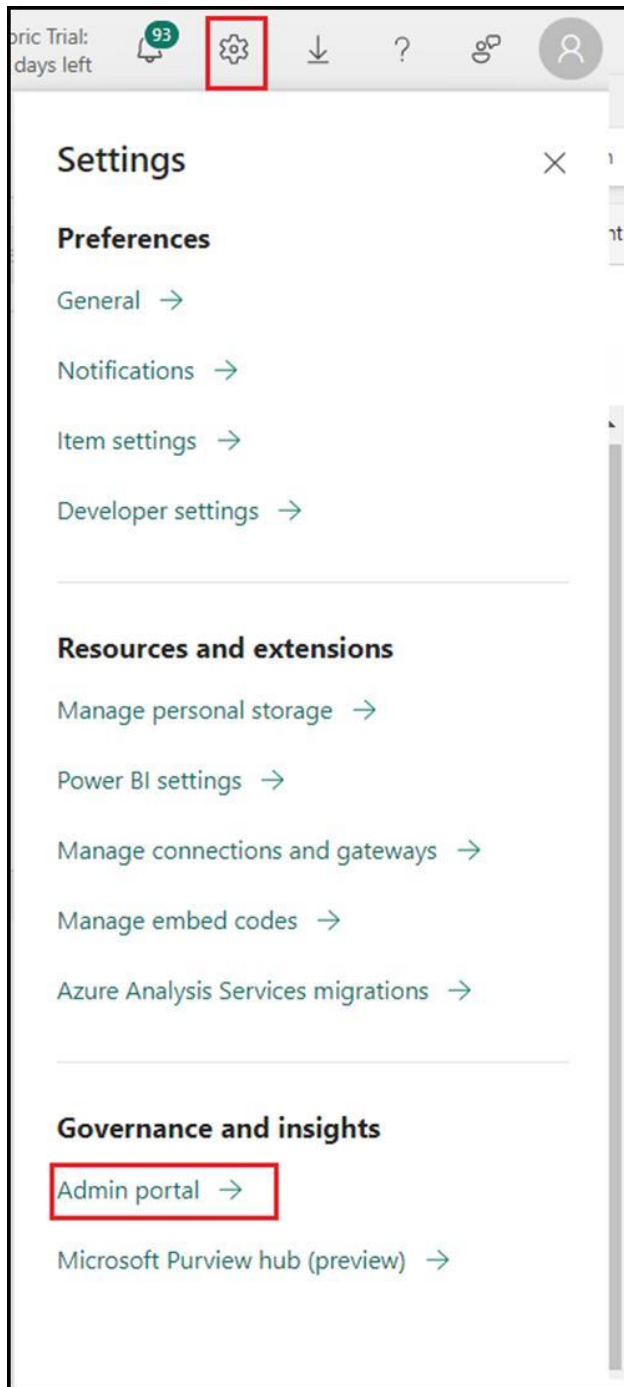


Microsoft Fabric Data Agent (Preview)

With the Microsoft Fabric AI skill, you can make data more accessible to your colleagues. You can configure a generative AI system to generate queries that answer questions about your data. After you configure the AI skill, you can share it with your colleagues, who can then ask their questions in plain English. Based on their questions, the AI generates queries over your data that answer those questions.

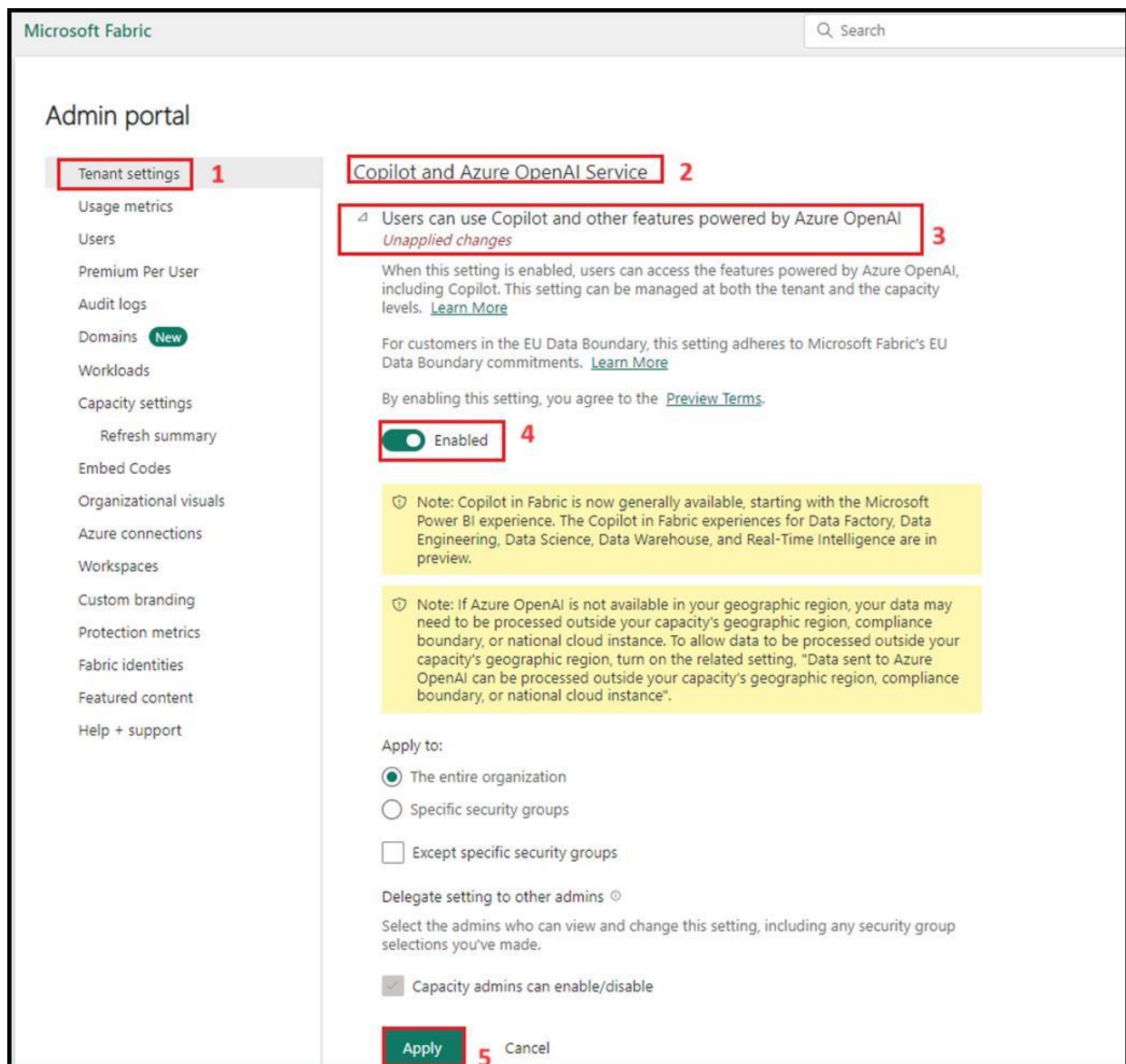
Task 1: Copilot tenant settings (For reference only)

1. On right side of Power BI home page, click on the **Settings** icon.
2. In **Settings** pane, scroll down to **Governance and insights**, then click on **Admin portal**.

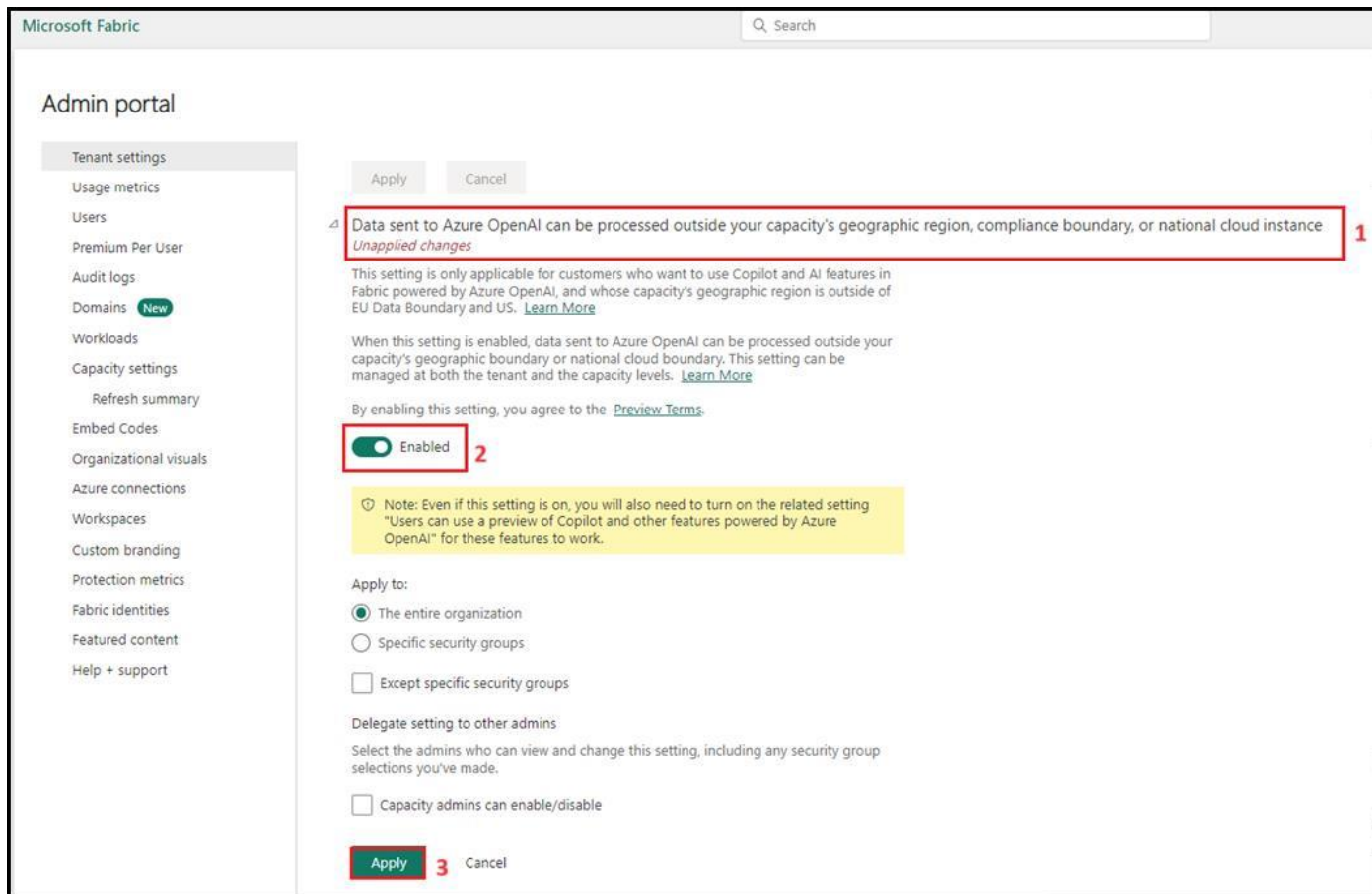


3. In **Admin portal** pane, select **Tenant settings**, scroll down to **Copilot and Azure OpenAI Service** section, click on **Users can use Copilot and other features powered by Azure OpenAI**, then enable it using

the **toggle** button. After **Users can use Copilot and other features powered by Azure OpenAI** were Enabled, click on the **Apply** button.

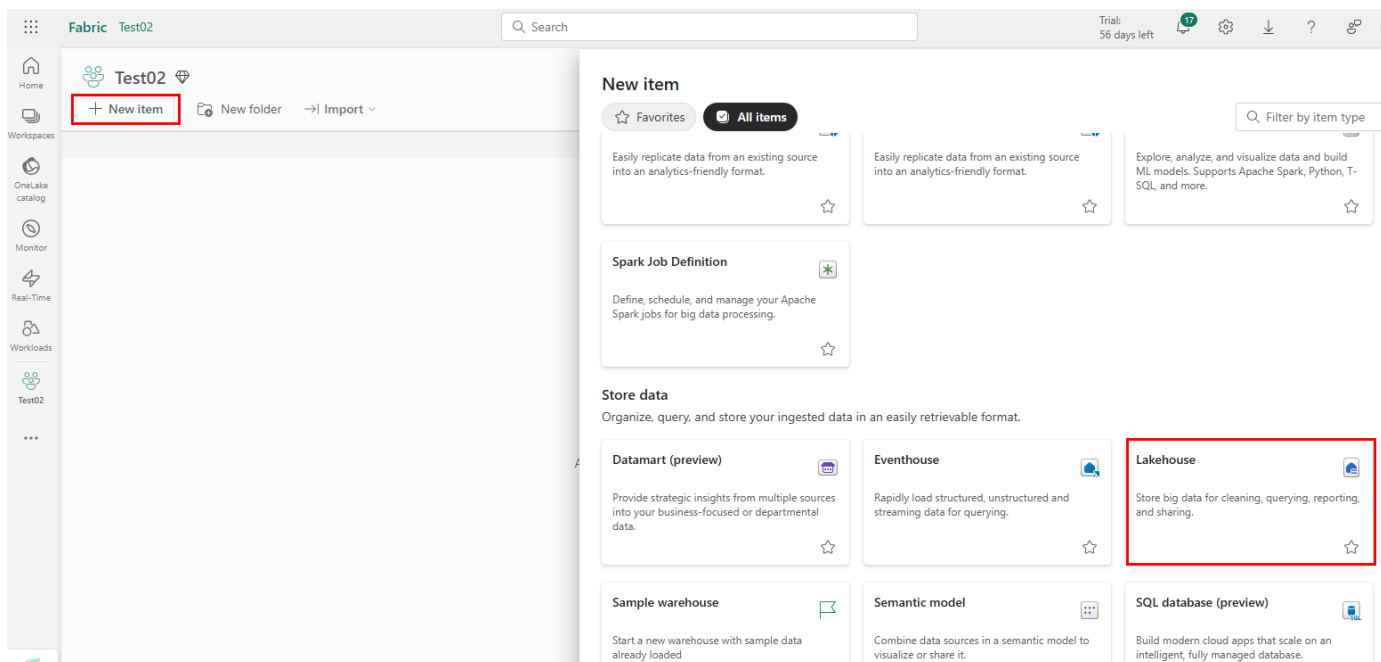


4. In **Admin portal** pane, select **Tenant settings**, scroll down to **Copilot and Azure OpenAI Service** section, click on **Data sent to Azure OpenAI can be processed outside your capacity\'s geographic region, compliance boundary, or national cloud instance**, then enable it using the **toggle** button. After **Data sent to Azure OpenAI can be processed outside your capacity\'s geographic region, compliance boundary, or national cloud instance** were Enabled, click on the **Apply** button.



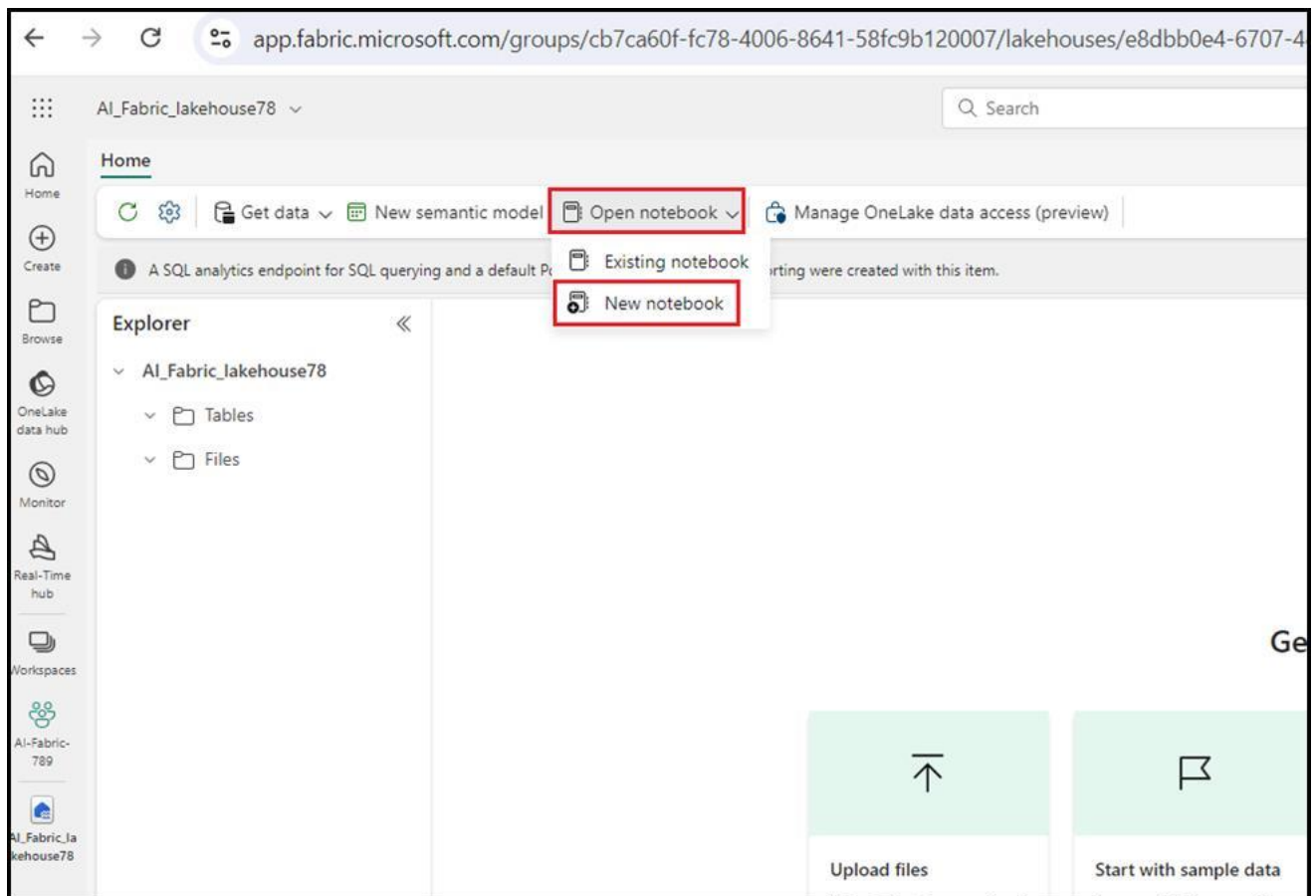
Task 2: Create a lakehouse

1. In the **Your Fabric Workspace** page
2. Click **+ New item**, select **Lakehouse** to create a lakehouse.



3. In the **New lakehouse** dialog box, enter **!!AILHX!!** in the **Name** field, click on the **Create** button and open the new lakehouse.
4. You will see a notification stating **Successfully created SQL endpoint**.

5. Next, create a new notebook to query the table. In the **Home** ribbon, select the drop down for **Open notebook** and choose **New notebook**.



Task 3: Upload AdventureWorksDW data into lakehouse

First, create a lakehouse and populate it with the necessary data.

In the query editor, copy and paste the following code. Select the **Run all** button to execute the query. After the query is completed, you will see the results.

```
import pandas as pd

from tqdm.auto import tqdm

base = "https://synapseaisolutionsa.blob.core.windows.net/public/AdventureWorks"

# load list of tables

df_tables = pd.read_csv(f"{base}/adventureworks.csv", names=["table"])

for table in (pbar := tqdm(df_tables['table'].values)):

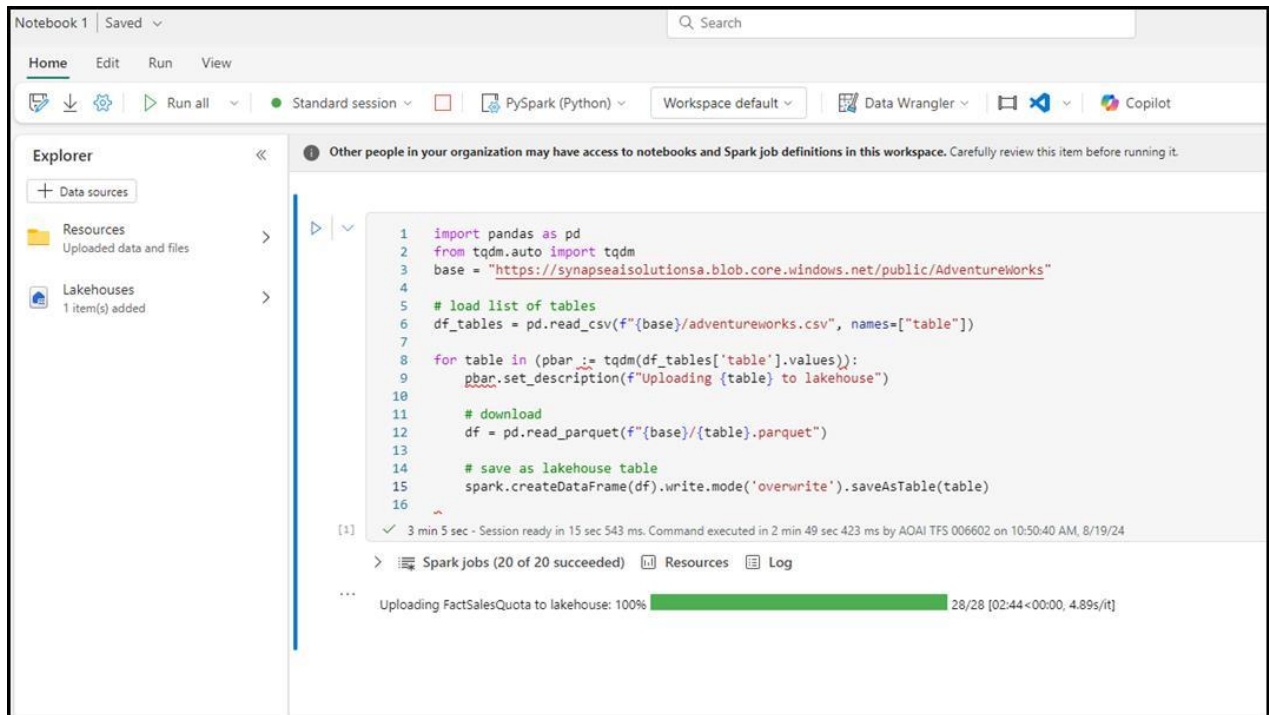
    pbar.set_description(f"Uploading {table} to lakehouse")

    # download

    df = pd.read_parquet(f"{base}/{table}.parquet")

    # save as lakehouse table

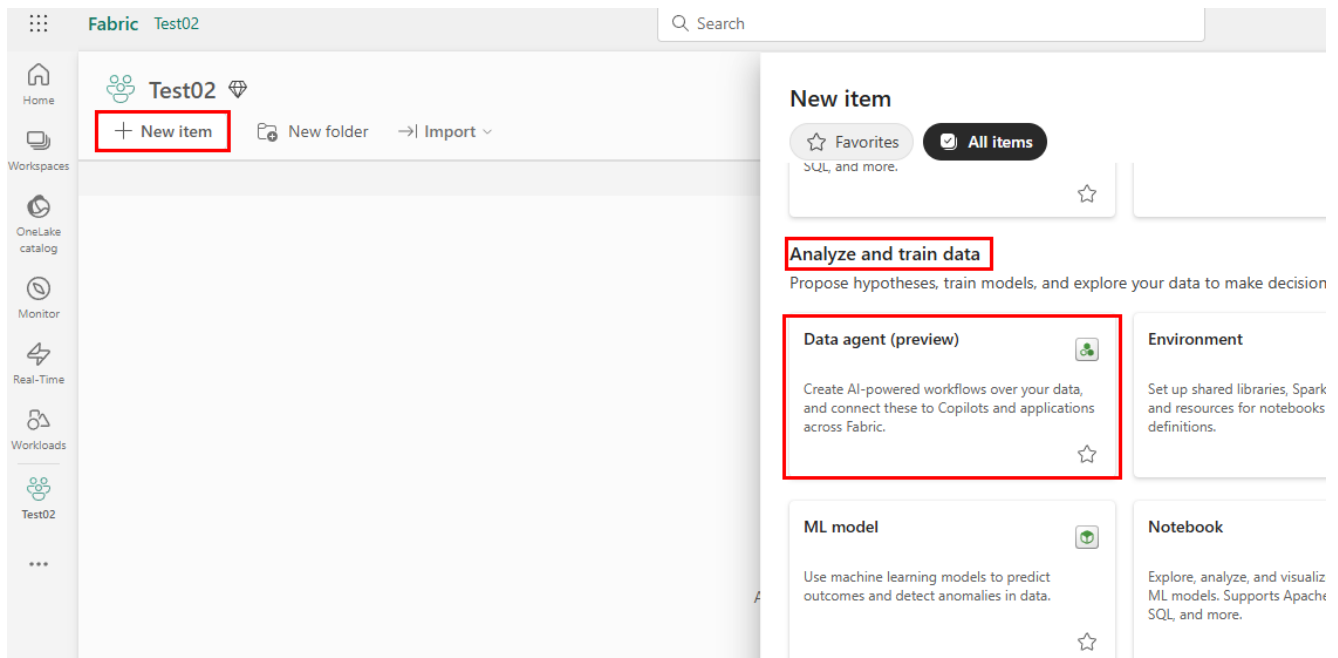
    spark.createDataFrame(df).write.mode('overwrite').saveAsTable(table)
```



After a few minutes, the lakehouse is populated with the necessary data.

Task 4: Create a Data Agent

1. To create a new Data Agent, go to **your Workspace**, Click **+ New item** and select **Data Agent(preview)**



2. In the **Create Data Agent** dialog box, enter **AIDataAgent** in the **Name** field, click on the **Create** button.
3. To add a Data source, select Data Source Option, then select your Lakehouse **AdventureWorksLH**

Add a data source

Filter by keyword **Filter** All domains

Clear all

Type

- ☐ KQL Database
- ☐ Lakehouse
- ☐ Semantic model
- ☐ Warehouse

Name	Type	Owner	Refreshed	En
Sales_Report	Semantic model		11/14/24, 3:29:05 P...	
Data Science Report	Semantic model		8/21/23, 2:32:36 PM	
AdventureWorksLH	Lakehouse		—	
Fabric WinWires	Semantic model		2/8/25, 6:07:45 AM	Confidential\Mi...
AdventureWorksLH	Semantic model		10/17/24, 8:27:48 A...	Confidential\Mi...
RAGTest2	Semantic model		10/15/24, 10:14:55 ...	Confidential\Mi...
RAGTest	Semantic model		10/14/24, 4:21:08 P...	Confidential\Mi...
churnProbelm	Semantic model		2/18/25, 6:09:33 PM	General
ChurnModel	Semantic model		6/22/23, 1:27:15 PM	Confidential\Mi...
MachineTest	Semantic model		9/10/24, 10:50:05 A...	Confidential\Mi...
ForecastDemo	Semantic model		9/2/24, 10:21:05 PM	Confidential\Mi...

Add Cancel

4. Once your data source is added, expand the AdventureWorksLH.

Home

Data agent instructions Example queries Clear chat Publish Revert to published version

Explorer

+ Data source

AdventureWorksLH

- dbo
 - ☐ dimaccount
 - ☐ dimcurrency
 - ☒ dimcustomer
 - ☒ dimdate
 - ☐ dimdepartmentgroup
 - ☐ dimemployee
 - ☒ dimgeography
 - ☐ dimorganization
 - ☒ dimproduct
 - ☒ dimproductcategory
 - ☐ dimproductsubcategory
 - ☒ dimpromotion
 - ☒ dimreseller
 - ☐ dimsalesreason
 - ☒ dimsalesterritory
 - ☐ dimscenario
 - ☐ factadditionalinternationalproductdescription
 - ☐ factcalcenter
 - ☐ factcurrencyrate
 - ☐ factfinance
 - ☒ factinternetsales
 - ☐ factinternetsalesreason
 - ☐ factproductinventory

AdventureSales-DataAgent-Scenario

What are the historical trends across all my data? Analyze recent data for any outliers Show me the details for a specific subcategory of data

Sample questions

Who are the top 5 customers by yearly income, and what are their state or province names?

Created with AI. Mistakes are possible. Review terms

5. You must then select the tables for which you want the AI skill to have available access.

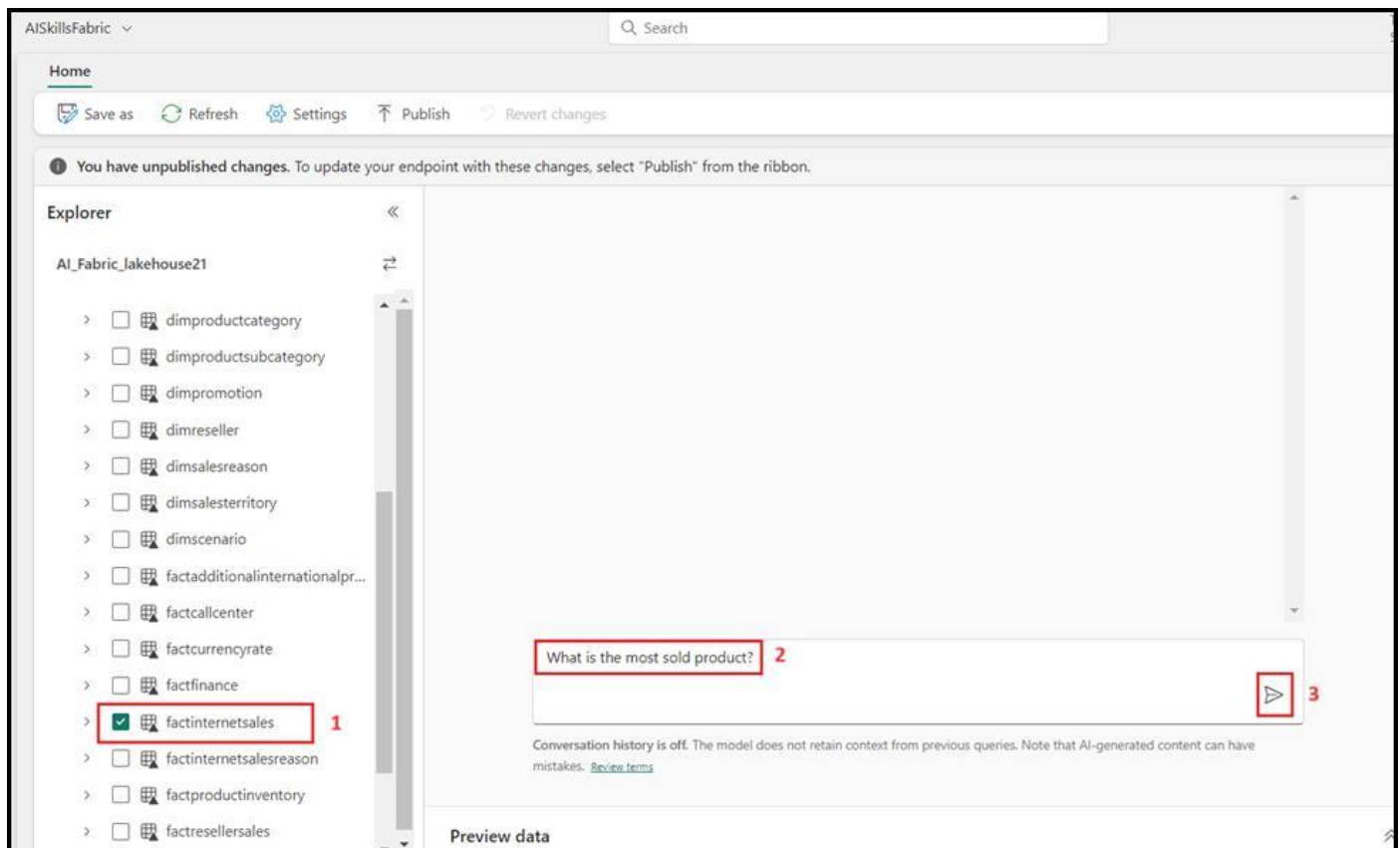
This lab uses these tables:

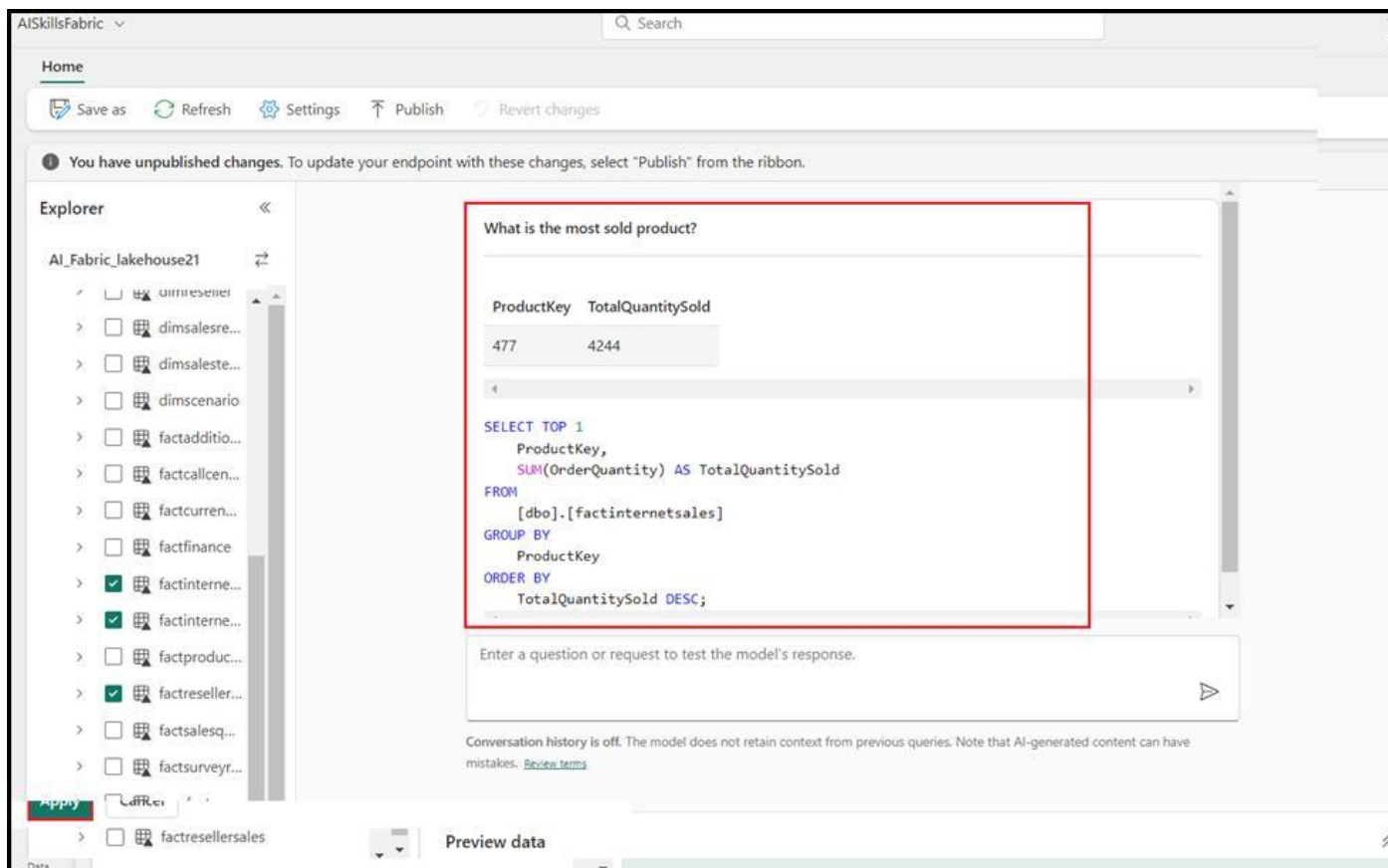
- DimCustomer
- DimDate
- DimGeography

- DimProduct
- DimProductCategory
- DimPromotion
- DimReseller
- DimSalesTerritory
- FactInternetSales
- FactResellerSales

Task 5: Provide instructions

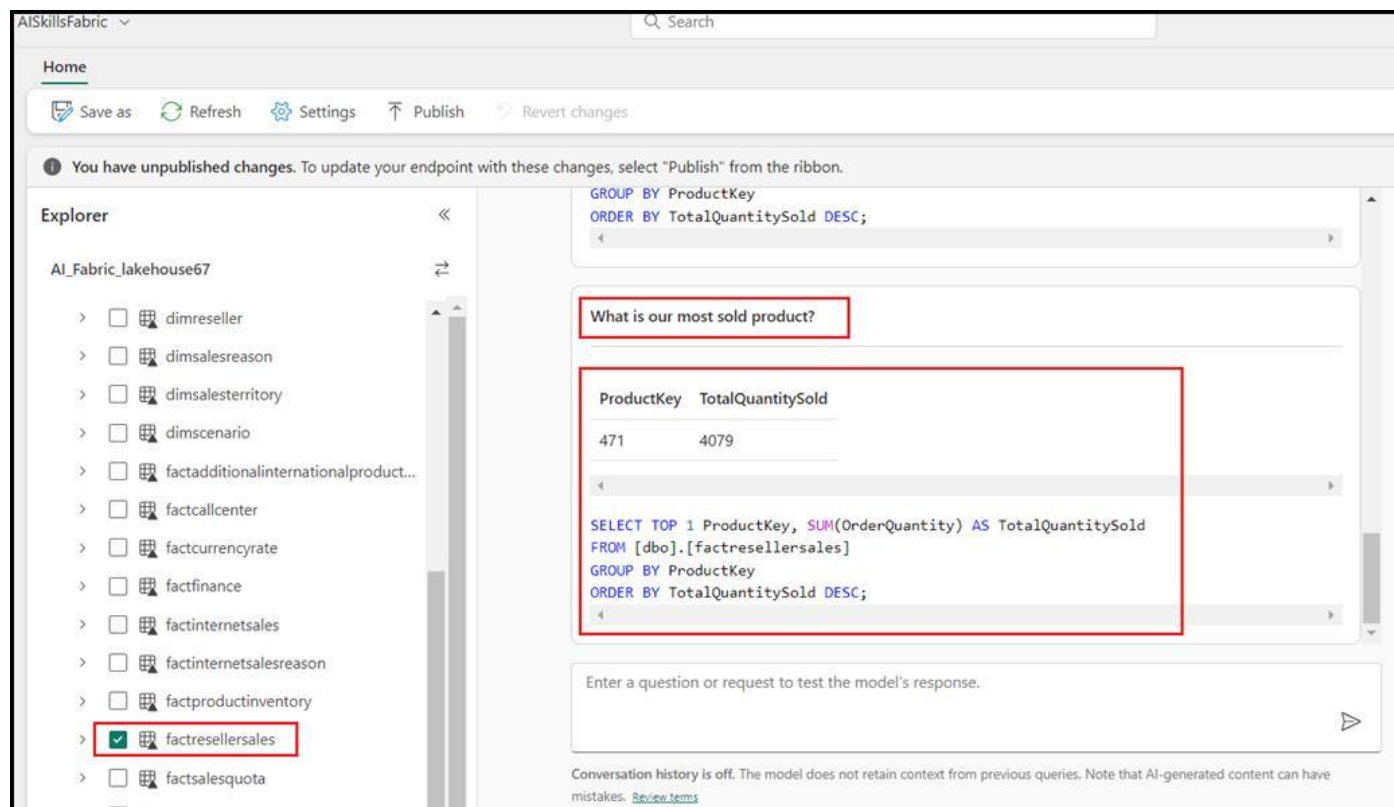
1. When you first ask the AI skill questions with the listed tables select **factinternetsales**, the AI skill answers them fairly well.
2. For instance, for the question **What is the most sold product?**, the AI skill returns:





3. Copy the all question and SQL queries and paste them in a notepad and then Save the notepad to use the information in the upcoming tasks
4. Select FactResellerSales and enter the following text and click on the **Submit icon** as shown in the below image.

What is our most sold product?



As you continue to experiment with queries, you should add more instructions.

5. Select the **dimcustomer** , enter the following text and click on the **Submit icon**

how many active customers did we have June 1st, 2013?

The screenshot shows the AI Skills Fabric interface. On the left, the 'Explorer' pane lists tables in the 'AI_Fabric_lakehouse67' database. The 'dimcustomer' table is selected and highlighted with a red box. The main pane displays a SQL query: `FROM [dbo].[factresellersales] GROUP BY ProductKey ORDER BY TotalQuantitySold DESC;`. Below the query, a text input field contains the question 'how many active customers did we have June 1st, 2010?'. The output area shows a table with one row: 'ActiveCustomers' with the value '0'. A second text input field at the bottom contains the question 'how many active customers did we have June 1st, 2013?' and is highlighted with a red box. A red arrow points to the 'Submit' icon (a right-pointing triangle) next to this input field. A message at the top states: 'You have unpublished changes. To update your endpoint with these changes, select "Publish" from the ribbon.'

The screenshot shows the AI Skills Fabric interface. On the left, the 'Explorer' pane lists tables in the 'AI_Fabric_lakehouse67' database. The 'dimcustomer' table is selected and highlighted with a green checkmark. The main pane displays a SQL query: `SELECT COUNT(CustomerKey) AS ActiveCustomers FROM [dbo].[dimcustomer] WHERE DateFirstPurchase <= '2010-06-01'`. Below the query, a text input field contains the question 'how many active customers did we have June 1st, 2013?'. The output area shows a table with one row: 'ActiveCustomers' with the value '10292'. A second text input field at the bottom contains the placeholder text 'Enter a question or request to test the model's response.' and is highlighted with a red box. A red arrow points to the 'Submit' icon (a right-pointing triangle) next to this input field. A message at the top states: 'You have unpublished changes. To update your endpoint with these changes, select "Publish" from the ribbon.'

6. Select the **dimdate**, **FactInternetSales** , enter the following text and click on the **Submit icon**

what are the monthly sales trends for the last year?

AI Skills Fabric

Home

Save as Refresh Settings Publish Revert changes

You have unpublished changes. To update your endpoint with these changes, select "Publish" from the ribbon.

Explorer

AI_Fabric_lakehouse67

Tables

- dimaccount
- dimcurrency
- dimcustomer
- dimdate**
- dimdepartmentgroup
- dimemployee
- dimgeography
- dimorganization
- dimproduct
- dimproductcategory
- dimproductsubcategory
- dimpromotion
- dimreseller

what are the monthly sales trends for the last year?

SQL result empty

```

SELECT
    YEAR(OrderDate) AS Year,
    MONTH(OrderDate) AS Month,
    SUM(SalesAmount) AS TotalSales
FROM
    [dbo].[factinternetsales]
WHERE
    OrderDate >= DATEADD(YEAR, -1, GETDATE())
GROUP BY
    YEAR(OrderDate),
    MONTH(OrderDate)
ORDER BY
    YEAR(OrderDate) DESC,
    MONTH(OrderDate) DESC;

```

what are the monthly sales trends for the last year?

Conversation history is off. The model does not retain context from previous queries. Note that AI-generated content can have mistakes. [Review terms](#)

AI Skills Fabric

Home

Save as Refresh Settings Publish Revert changes

You have unpublished changes. To update your endpoint with these changes, select "Publish" from the ribbon.

Explorer

AI_Fabric_lakehouse67

Tables

- dimaccount
- dimcurrency
- dimcustomer
- dimdate**
- dimdepartmentgroup
- dimemployee
- dimgeography
- dimorganization
- dimproduct
- dimproductcategory
- dimproductsubcategory
- dimpromotion
- dimreseller

what are the monthly sales trends for the last year?

SQL result empty

```

SELECT
    DATEPART(YEAR, d.FullDateAlternateKey) AS Year,
    DATEPART(MONTH, d.FullDateAlternateKey) AS Month,
    SUM(s.SalesAmount) AS TotalSalesAmount
FROM
    dbo.dimdate d
JOIN
    dbo.factinternetsales s ON d.DateKey = s.OrderDateKey
WHERE
    d.FullDateAlternateKey >= DATEADD(YEAR, -1, GETDATE())
GROUP BY
    DATEPART(YEAR, d.FullDateAlternateKey),

```

Enter a question or request to test the model's response.

Conversation history is off. The model does not retain context from previous queries. Note that AI-generated content can have mistakes. [Review terms](#)

7. Select the **dimproduct**, **FactInternetSales**, enter the following text and click on the **Submit** icon

which product category had the highest average sales price?

Home

Save as Refresh Settings Publish Revert changes

You have unpublished changes. To update your endpoint with these changes, select "Publish" from the ribbon.

Explorer

AI_Fabric_lakehouse67

- ☐ dimorganization
- ☒ dimproduct
- ☐ dimproductcategory
- ☐ dimproductsubcategory
- ☐ dimpromotion
- ☐ dimreseller
- ☐ dimsalesreason
- ☐ dimsalesterritory
- ☐ dimscenario
- ☐ factadditionalinternationalproduct...
- ☐ factcallcenter
- ☐ factcurrencyrate
- ☐ factfinance
- ☒ factinternetsales
- ☐ factinternetsalesreason

SQL result empty

```
SELECT
    DATEPART(YEAR, d.FullDateAlternateKey) AS Year,
    DATEPART(MONTH, d.FullDateAlternateKey) AS Month,
    SUM(s.SalesAmount) AS TotalSalesAmount
FROM
    dbo.dimdate d
JOIN
    dbo.factinternetsales s ON d.DateKey = s.OrderDateKey
WHERE
    d.FullDateAlternateKey >= DATEADD(YEAR, -1, GETDATE())
GROUP BY
    DATEPART(YEAR, d.FullDateAlternateKey),
    DATEPART(MONTH, d.FullDateAlternateKey)
ORDER BY
    DATEPART(YEAR, d.FullDateAlternateKey),
    DATEPART(MONTH, d.FullDateAlternateKey);
```

which product category had the highest average sales price?

Conversation history is off. The model does not retain context from previous queries. Note that AI-generated content can have mistakes. [Review terms](#)

Preview data

Home

Save as Refresh Settings Publish Revert changes

You have unpublished changes. To update your endpoint with these changes, select "Publish" from the ribbon.

Explorer

AI_Fabric_lakehouse67

- ☐ dimorganization
- ☒ dimproduct
- ☐ dimproductcategory
- ☐ dimproductsubcategory
- ☐ dimpromotion
- ☐ dimreseller
- ☐ dimsalesreason
- ☐ dimsalesterritory
- ☐ dimscenario
- ☐ factadditionalinternationalproduct...
- ☐ factcallcenter
- ☐ factcurrencyrate
- ☐ factfinance
- ☒ factinternetsales
- ☐ factinternetsalesreason

DATEPART(MONTH, d.FullDateAlternateKey);

which product category had the highest average sales price?

ProductSubcategoryKey	AvgSalesPrice
1	2002.567559

```
SELECT TOP 1 dp.ProductSubcategoryKey, AVG(fis.SalesAmount) AS AvgSalesPrice
FROM dbo.dimproduct dp
JOIN dbo.factinternetsales fis ON dp.ProductKey = fis.ProductKey
GROUP BY dp.ProductSubcategoryKey
ORDER BY AvgSalesPrice DESC;
```

Enter a question or request to test the model's response.

Conversation history is off. The model does not retain context from previous queries. Note that AI-generated content can have mistakes. [Review terms](#)

Preview data

Part of the problem is that "active customer" doesn't have a formal definition. More instructions in the notes to the model text box might help, but users might frequently ask this question. You need to make sure that the AI handles the question correctly

- The relevant query is moderately complex, so provide an example by selecting the edit button.
- In the Example SQL queries tab, select the **+Add example**.

Example SQL queries

Add examples of how a natural language question would become a SQL query for your data. The AI skill will automatically select up to three valid examples to share with the model each time it handles a user question. [Learn more](#)

+ Add example

Import from .json

Download all as .json

Delete

Showing 0 items (0 selected)



No examples added yet

Add an example manually or import from a .json file. [Learn more](#)

10. You can manually add examples, but you can also upload them from a JSON file. Providing examples from a file is helpful when you have many SQL queries that you want to upload all at once, instead of manually uploading the queries one by one.
11. Add all the queries and SQL queries that you have saved in Notepad, and then click on 'Download all as .json'.

Example SQL queries

Add examples of how a natural language question would become a SQL query for your data. The AI skill will automatically select up to three valid examples to share with the model each time it handles a user question. [Learn more](#)

+ Add example

1 Import from .json

Download all as .json

3 Delete

Filter by keyword

Showing 3 items (3 selected)

Question

What is the most sold product?

SQL query

```
SELECT TOP 1 ProductKey, SUM(OrderQuantity) AS TotalQuantitySold
FROM [dbo].[factinternetsales]
GROUP BY ProductKey
ORDER BY TotalQuantitySold DESC;
```

30/500 characters used

171/1000 characters used

What is our most sold product?

```
-- Enter SQL query
SELECT TOP 1 ProductKey, SUM(OrderQuantity) AS TotalQuantitySold
FROM [dbo].[factresellersales]
GROUP BY ProductKey
ORDER BY TotalQuantitySold DESC;
```

30/500 characters used

171/1000 characters used

how many active customers did we have June 1st, 2013?

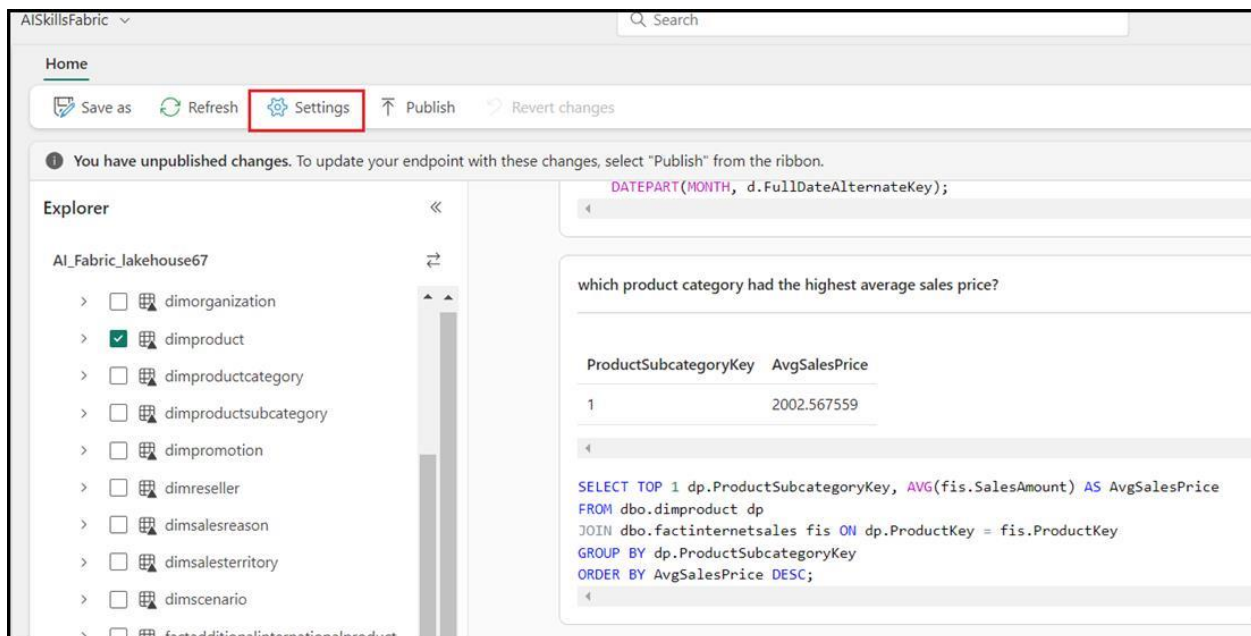
```
FROM [dbo].[dimcustomer]
WHERE DateFirstPurchase <= '2013-06-01'
```

Task 9: Use the AI skill programmatically

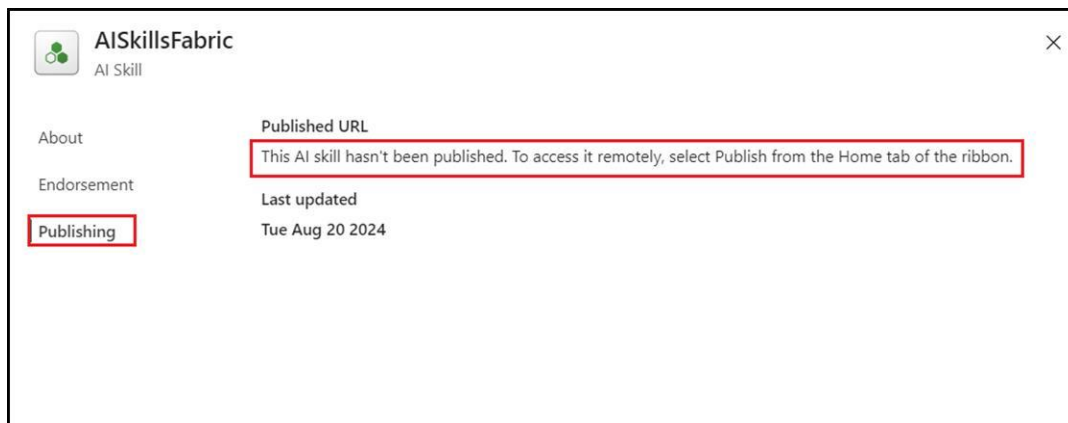
Both instructions and examples were added to the AI skill. As testing proceeds, more examples and instructions can improve the AI skill even further. Work with your colleagues to see if you provided examples and instructions that cover the kinds of questions they want to ask.

You can use the AI skill programmatically within a Fabric notebook. To determine whether or not the AI skill has a published URL value.

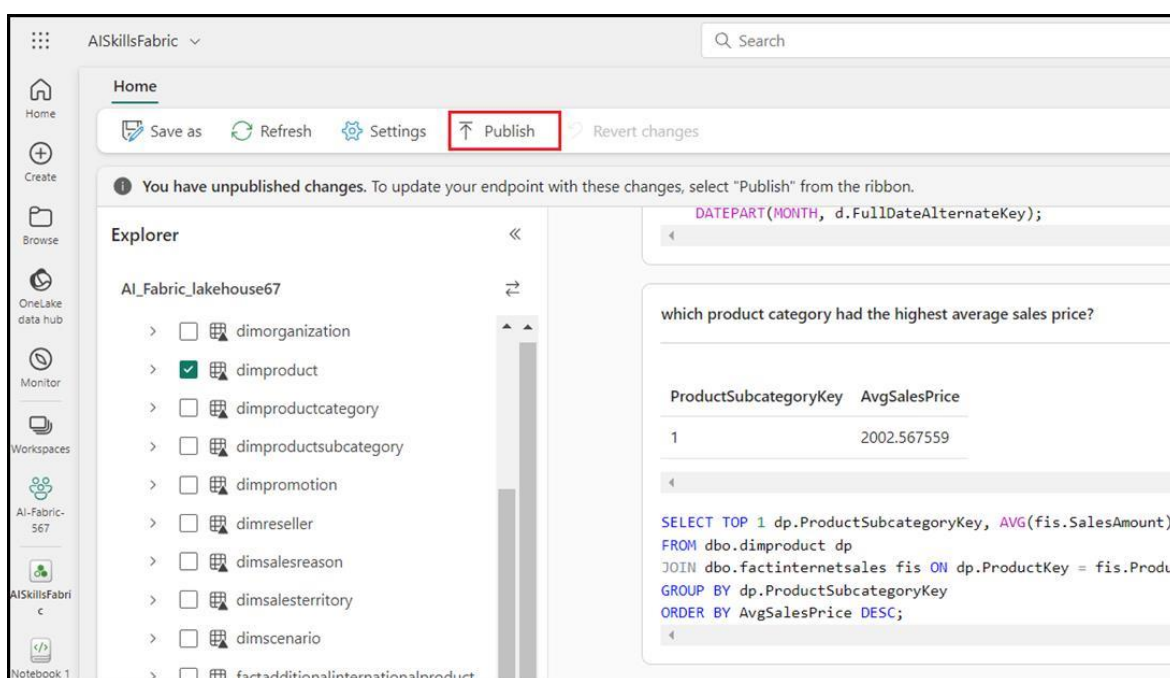
1. In the AISkillFabric page, in the **Home** ribbon select the **Settings**.



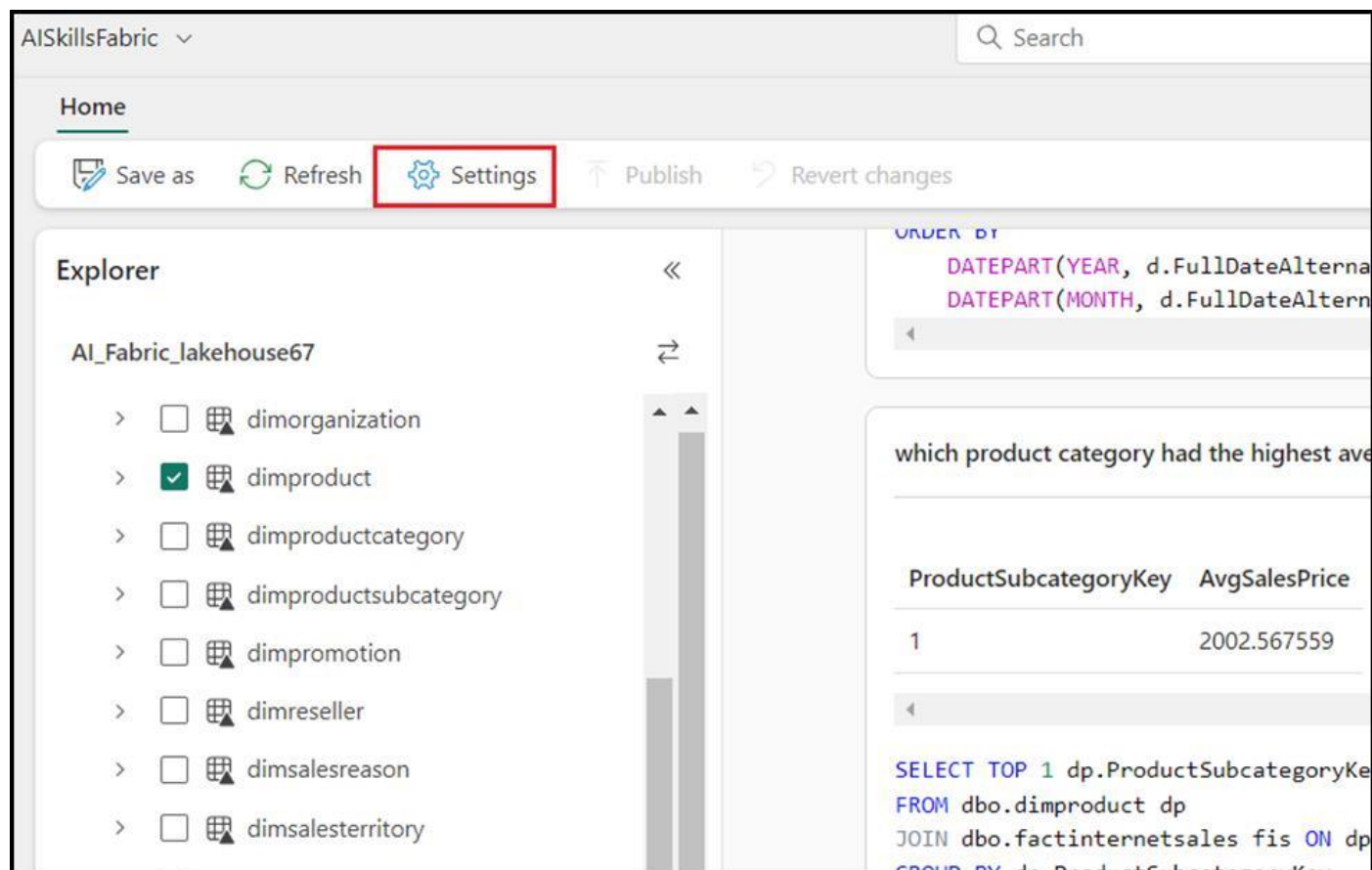
- Before you publish the AI skill, it doesn't have a published URL value, as shown in this screenshot.
- Close the AI Skill setting.



- In the **Home** ribbon, select the **Publish**.

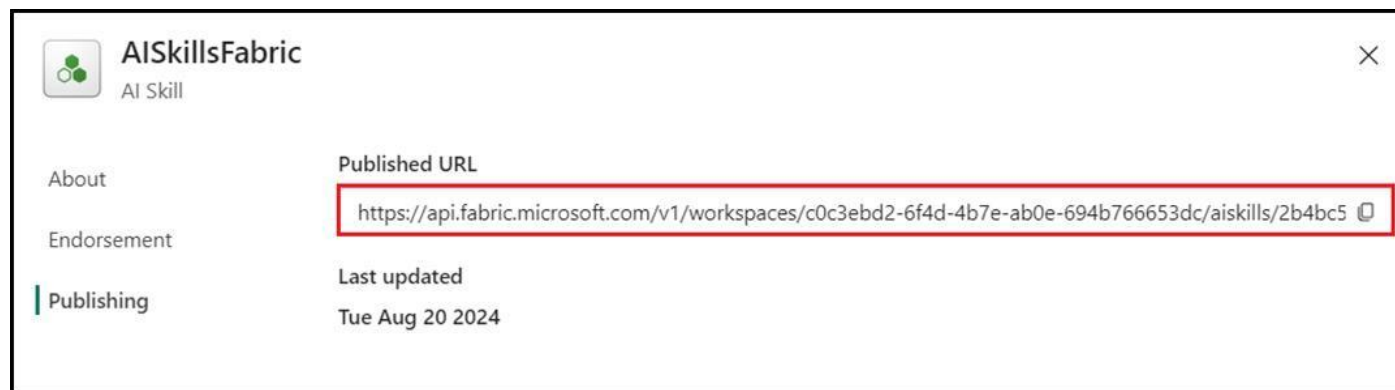


5. After publishing, select 'Settings' from the Home ribbon

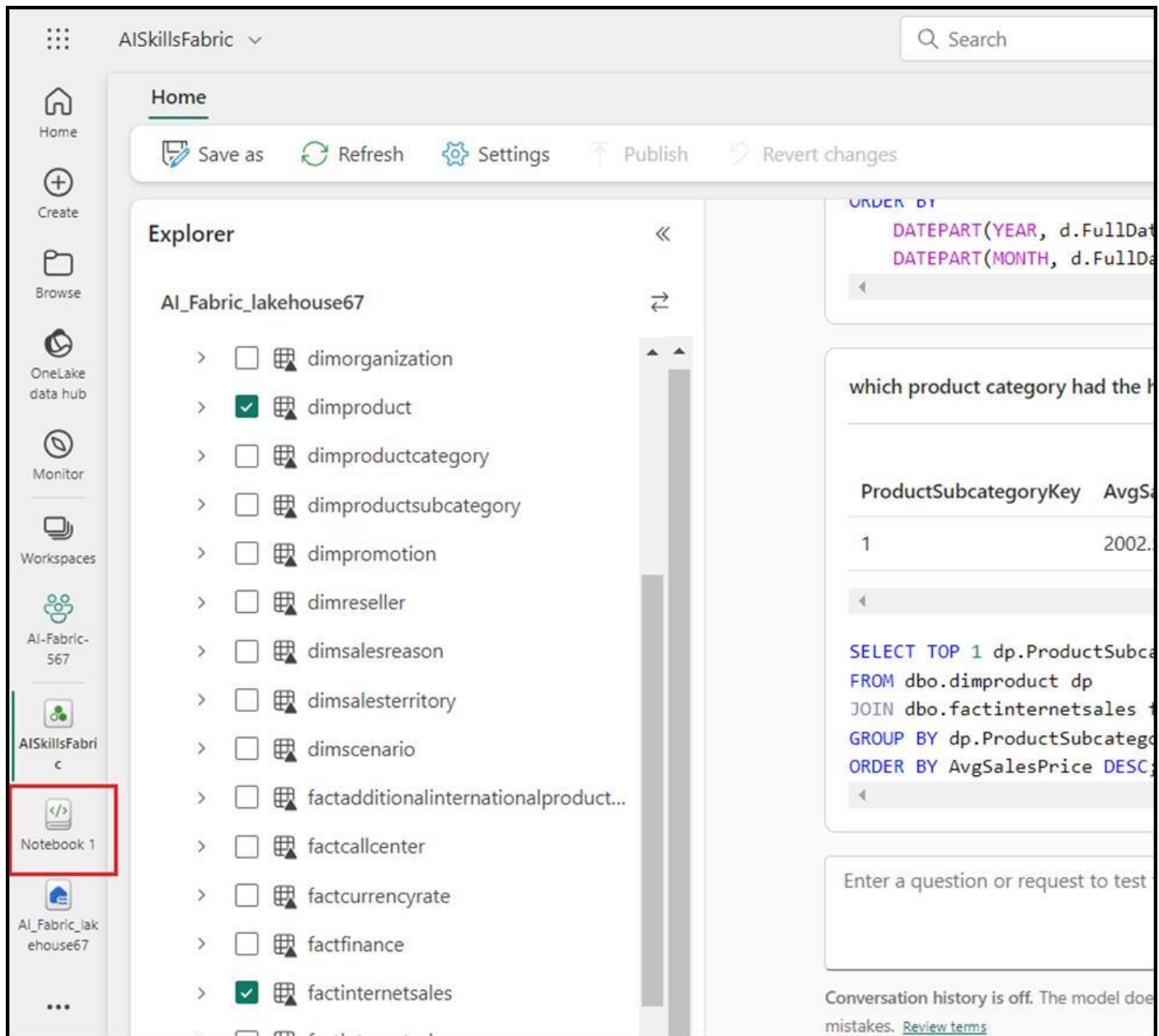


6. The published URL for the AI skill appears, as shown in this screenshot.

7. Copy the URL and paste that in a notepad and then Save the notepad to use the information in the upcoming



8. Select **Notebook1** in the left navigation pane.



9. Use the **+ Code** icon below the cell output to add a new code cell to the notebook, enter the following code in it and replace the **URL**. Click on **▶ Run** button and review the output

```
import requests
import json
import pprint
from synapse.ml.mlflow import get_mlflow_env_config
# the URL could change if the workspace is assigned to a different capacity
url = "https://<generic published URL value>"
configs = get_mlflow_env_config()
headers = {
    "Authorization": f"Bearer {configs.driver_aad_token}",
    "Content-Type": "application/json; charset=utf-8"
}
```



```

question = "{userQuestion: \"what is an example product?\"}"

response = requests.post(url, headers=headers, data = question)

print("RESPONSE: ", response)

print("")

response = json.loads(response.content)

print(response["result"])

```

```

1 t requests
2 t json
3 t pprint
4 synapse.ml.mflow import get_mflow_env_config
5
6
7 URL could change if the workspace is assigned to a different capacity
8 "https://api.fabric.microsoft.com/v1/workspaces/c0c3ebd2-6f4d-4b7e-ab0e-694b766653dc/aikills/2b4bc5c8-5523-408f-af7e-3f89b9296332/query/deployment"
9
10 gs = get_mflow_env_config()
11
12 rs = {
13   Authorization: f"Bearer {configs.driver_aad_token}",
14   Content-Type: "application/json; charset=utf-8"
15 }
16
17 ion = "{userQuestion: \"what is an example product?\"}"
18
19 nse = requests.post(url, headers=headers, data = question)
20
21 ("RESPONSE: ", response)
22
23 (""")
24
25 nse = json.loads(response.content)
26
27 (response["result"])

```

```

1 t json
2 t pprint
3 t pprint
4 synapse.ml.mflow import get_mflow_env_config
5
6
7 URL could change if the workspace is assigned to a different capacity
8 "https://api.fabric.microsoft.com/v1/workspaces/c0c3ebd2-6f4d-4b7e-ab0e-694b766653dc/aikills/2b4bc5c8-5523-408f-af7e-3f89b9296332/query/deployment"
9
10 gs = get_mflow_env_config()
11
12 rs = {
13   Authorization: f"Bearer {configs.driver_aad_token}",
14   Content-Type: "application/json; charset=utf-8"
15 }
16
17 ion = "{userQuestion: \"what is an example product?\"}"
18
19 nse = requests.post(url, headers=headers, data = question)
20
21 ("RESPONSE: ", response)
22
23 (""")
24
25 nse = json.loads(response.content)
26
27 (response["result"])

```

✓ 2 min 55 sec - Session ready in 2 min 30 sec 405 ms. Command executed in 25 sec 252 ms by MOD Administrator on 10:03:32 AM, 8/20/24

> Log

```

... RESPONSE: <Response [200]>
LL Road Frame - Black, 58

```