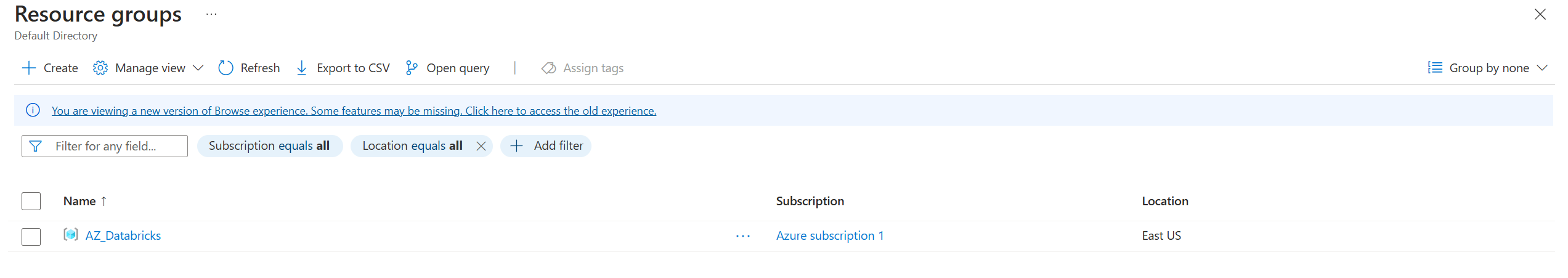
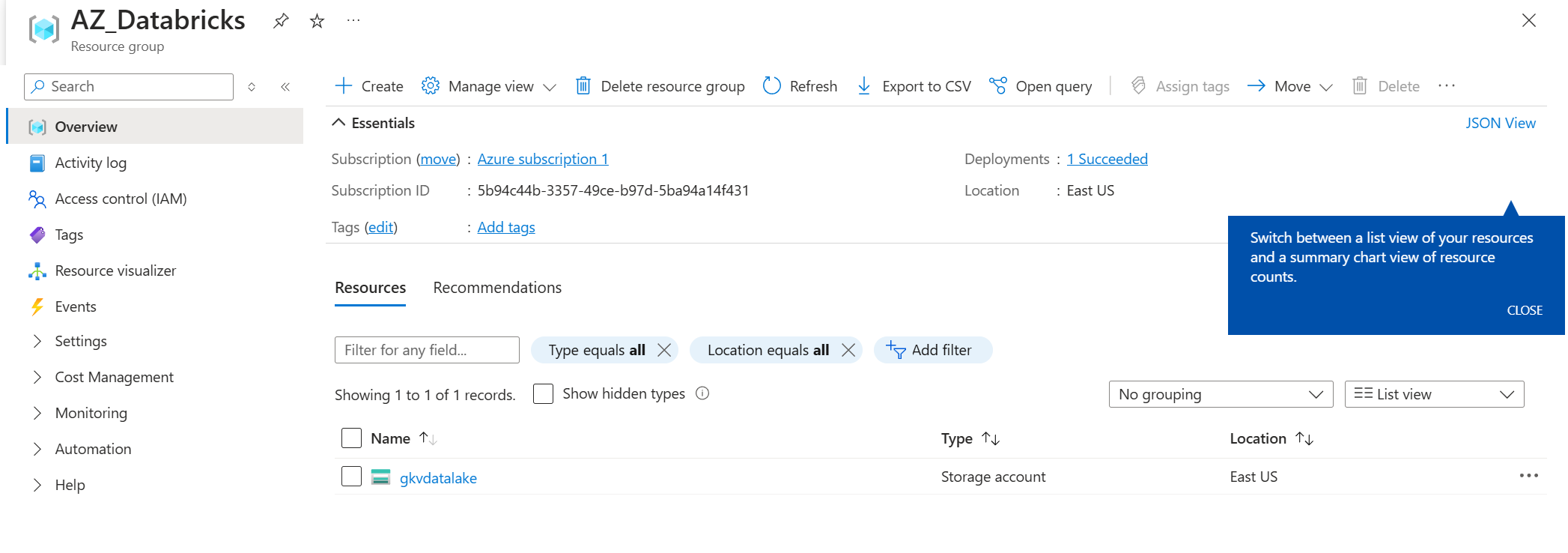
**Databricks in Azure**

Step 1: Create a Resource Group.



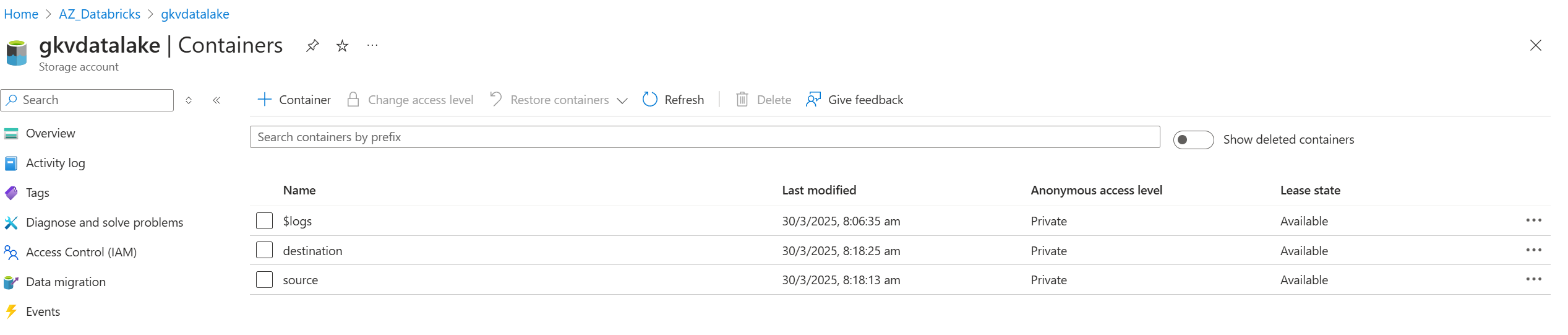
Step 2: Create a storage system “Azure Lake”.

Note: By default, Azure creates Blob storage. To convert to a Data lake while creating a storage container, enable a Hierarchical namespace.

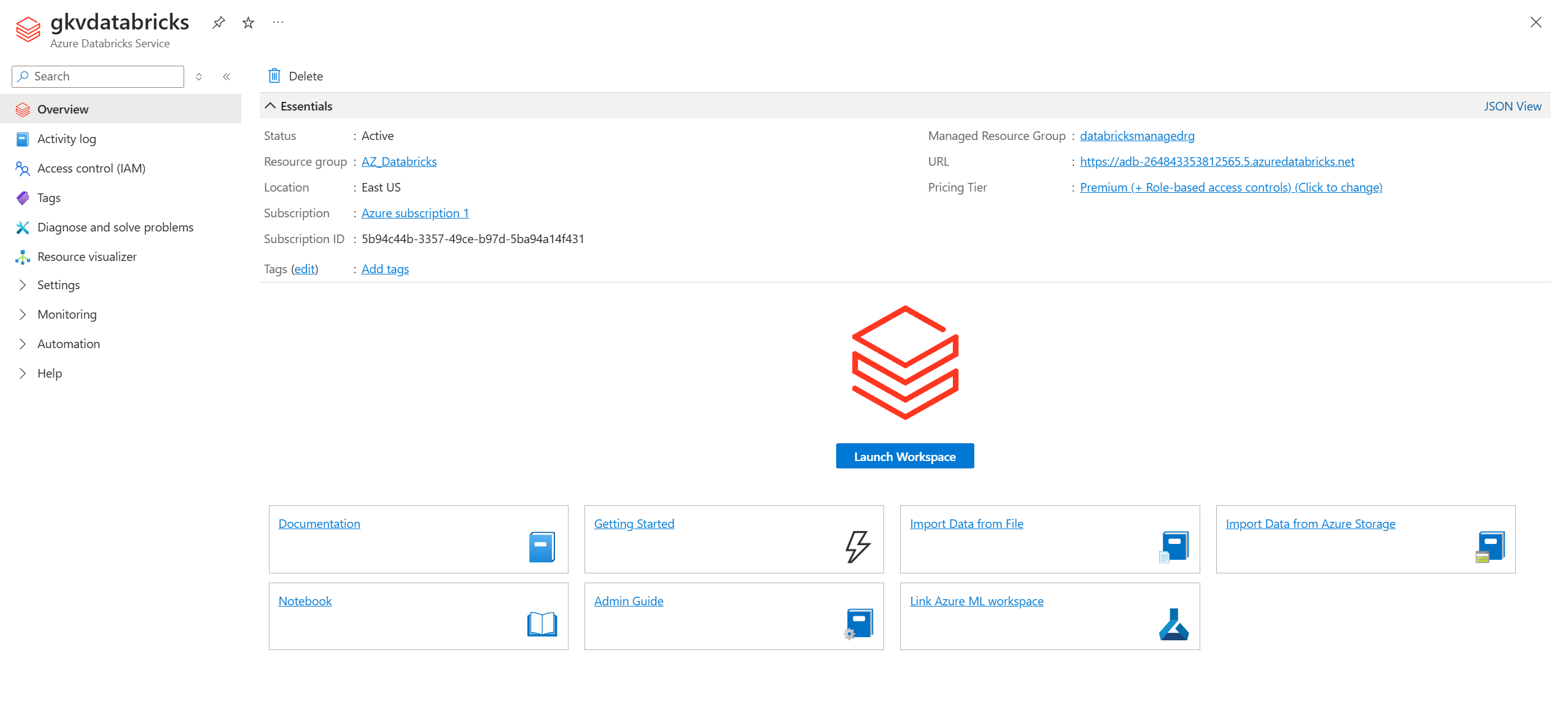
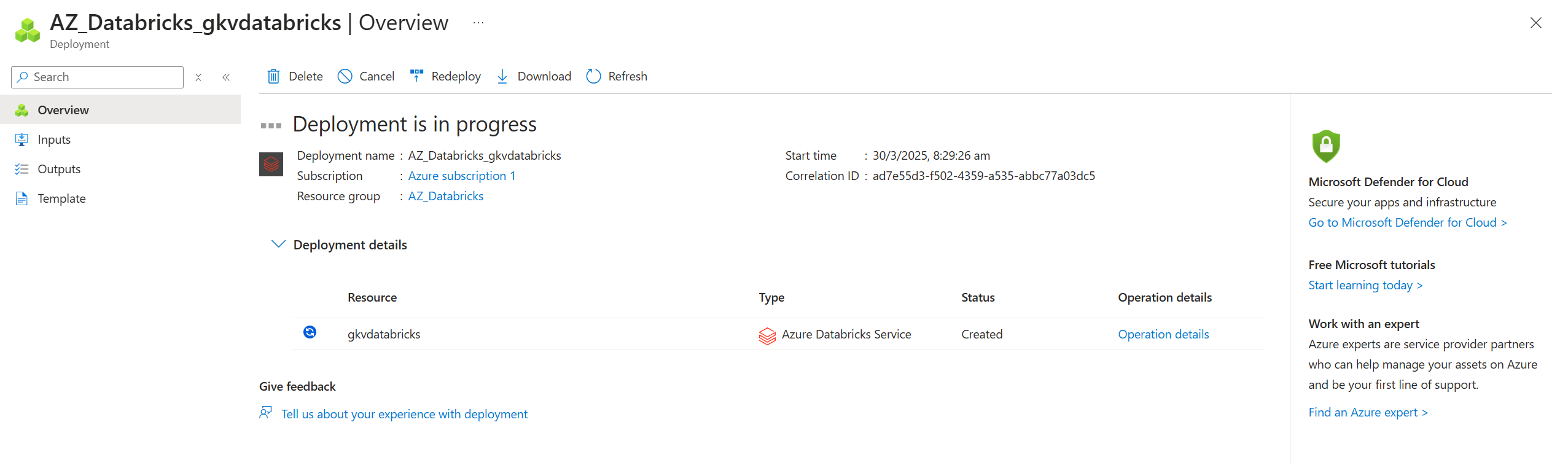


Step 3: In datalake created navigate to “Data Storage” and click on containers to create one.

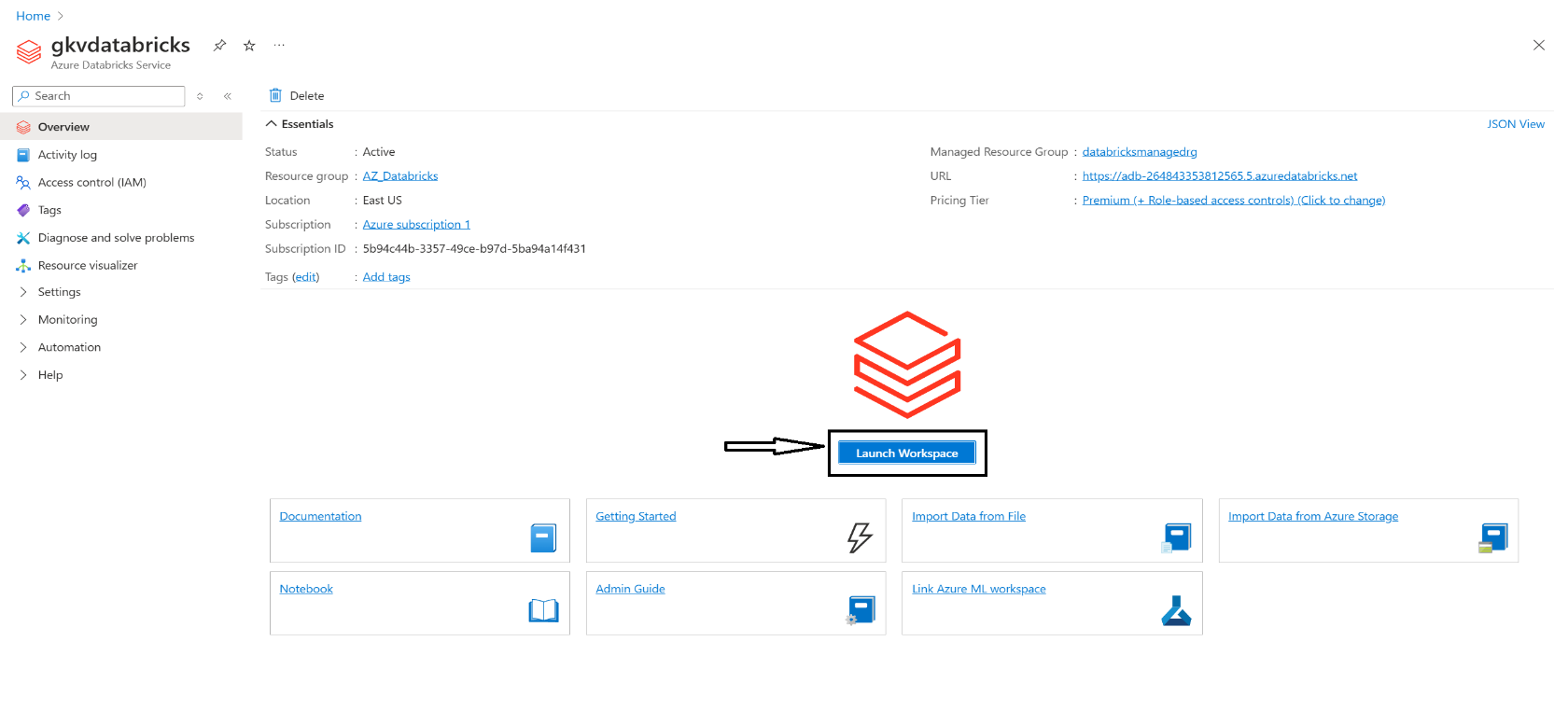
Two containers are to be created one as source and the other as sink.

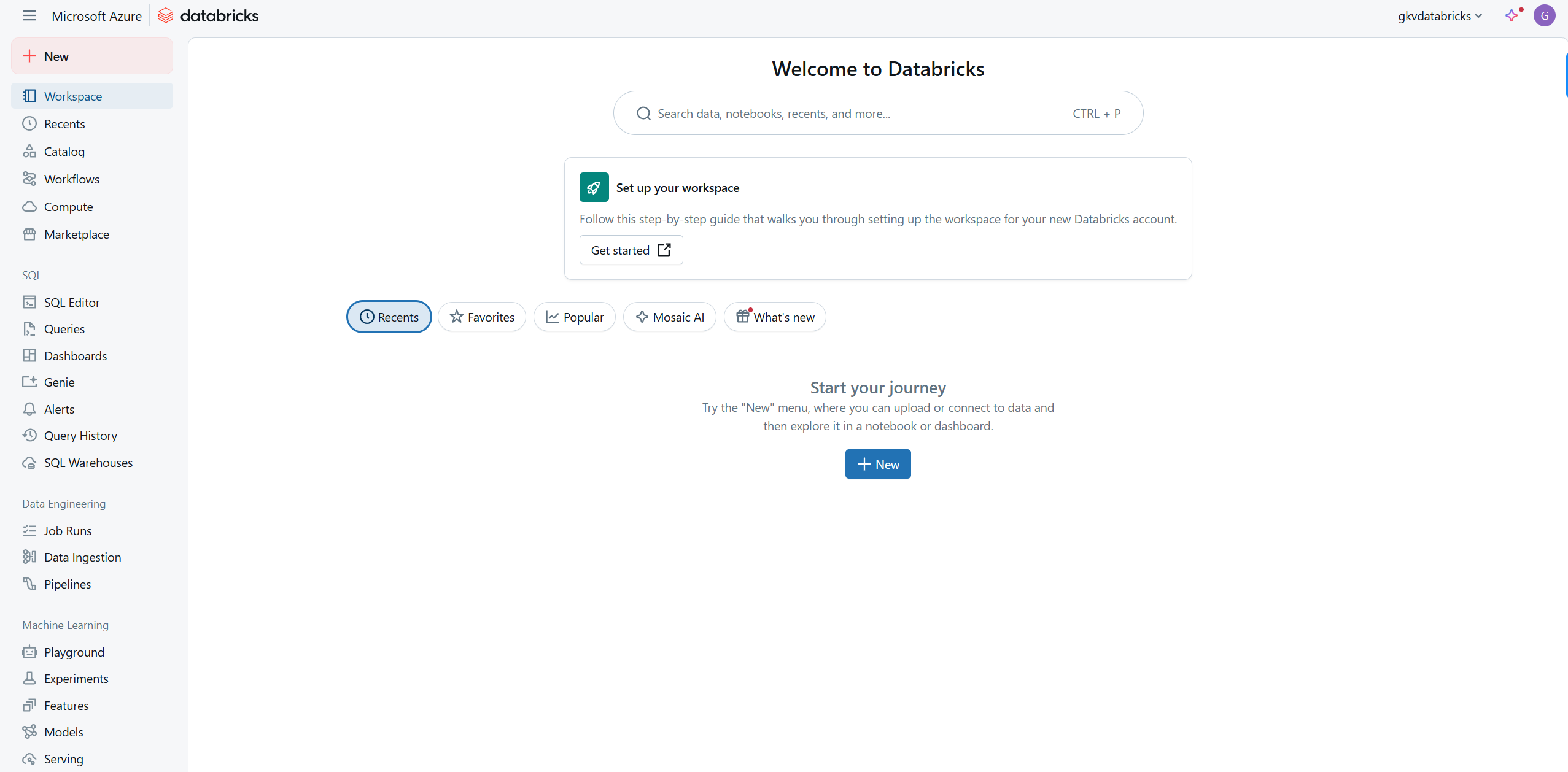


Step 4: Create Databricks workspace.

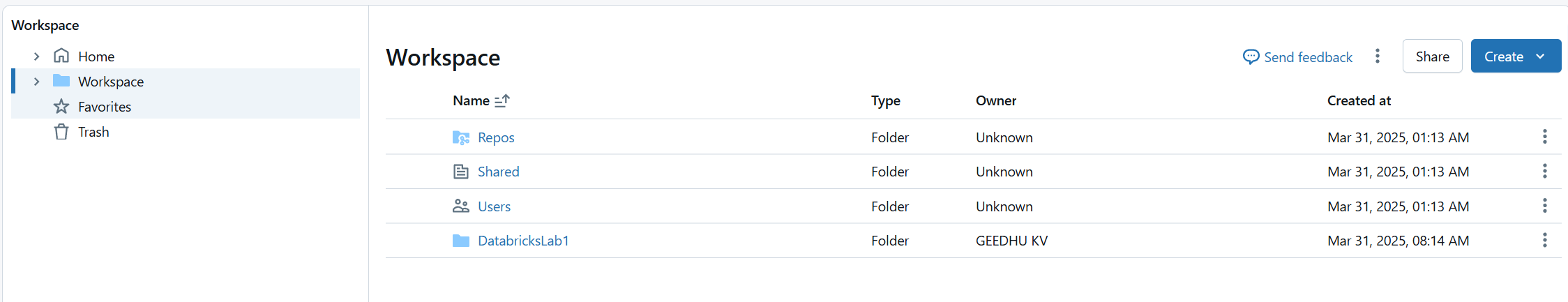


Step 5: Click on Launch Workspace

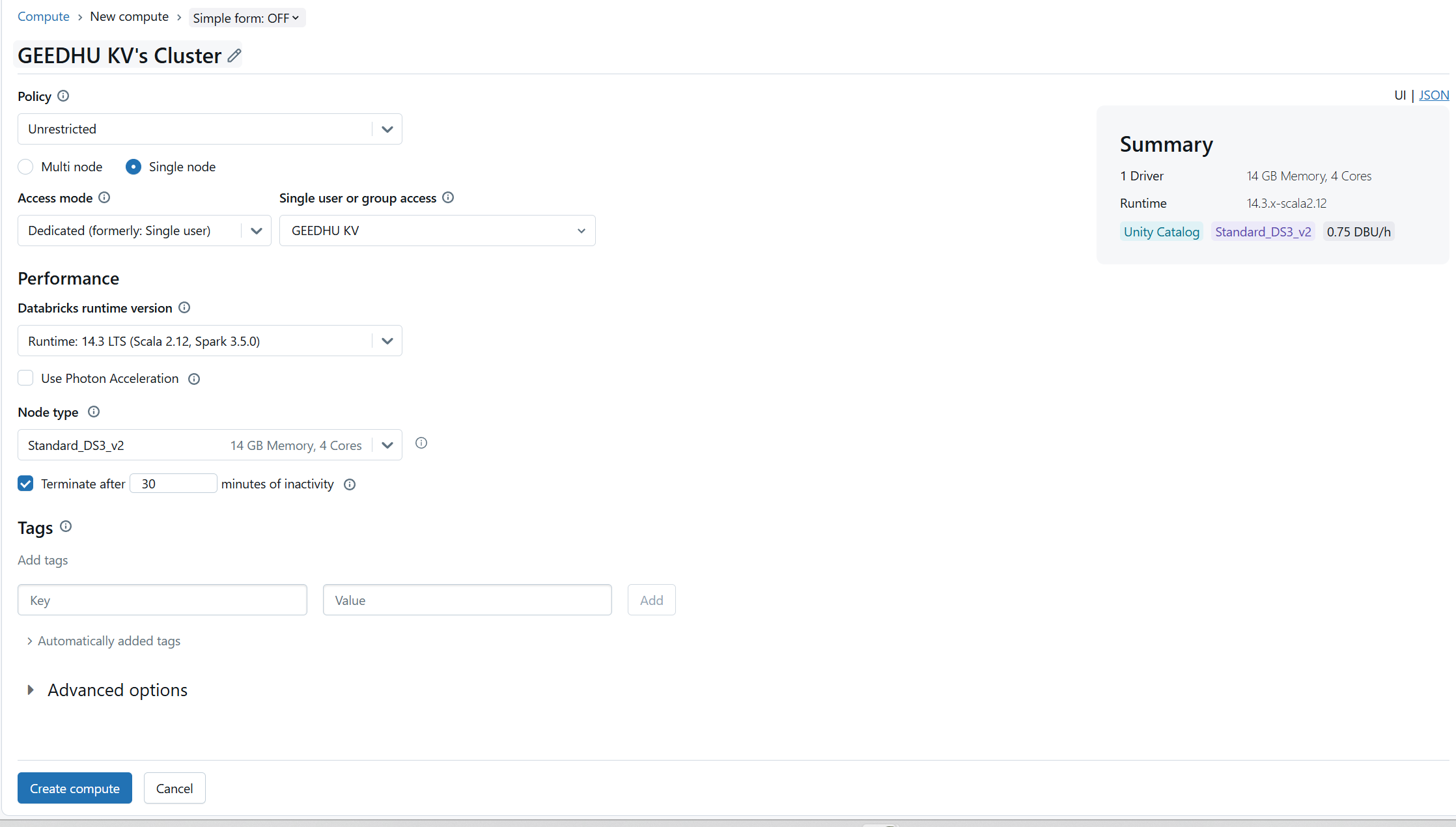




Step 6: Create a folder inside the Workspace.

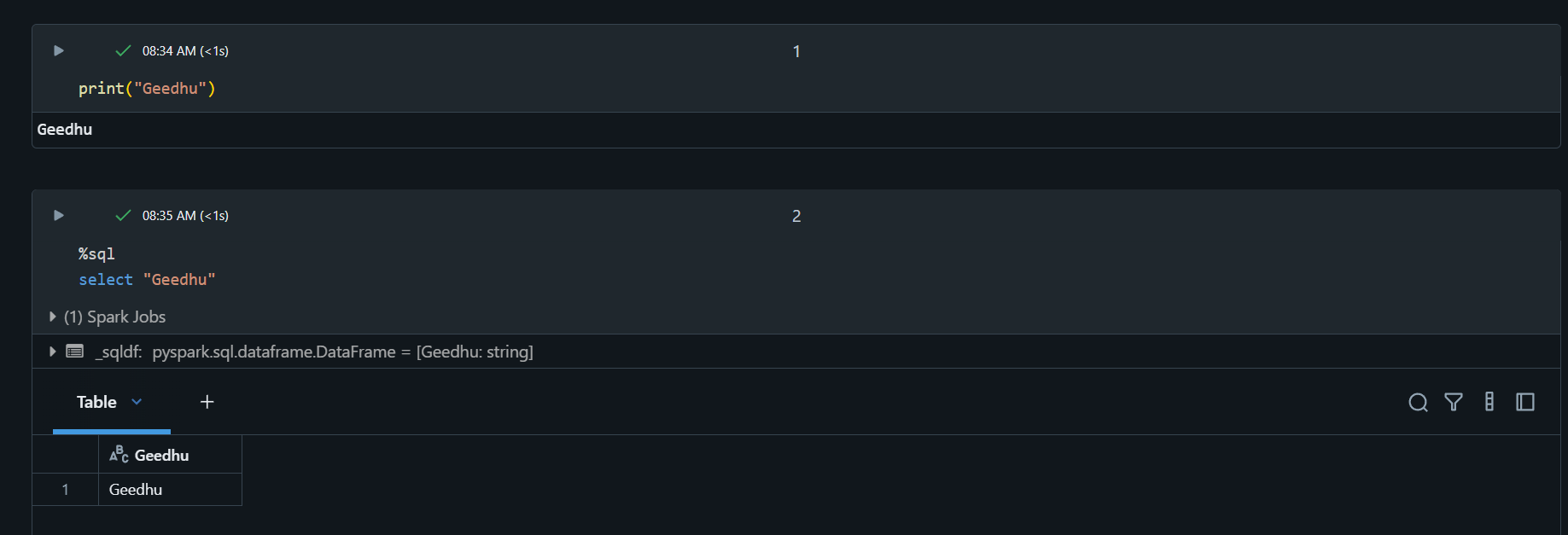


Step 7: Create a cluster in the compute.



Step 8: Create a notebook.

Working with magic commands.



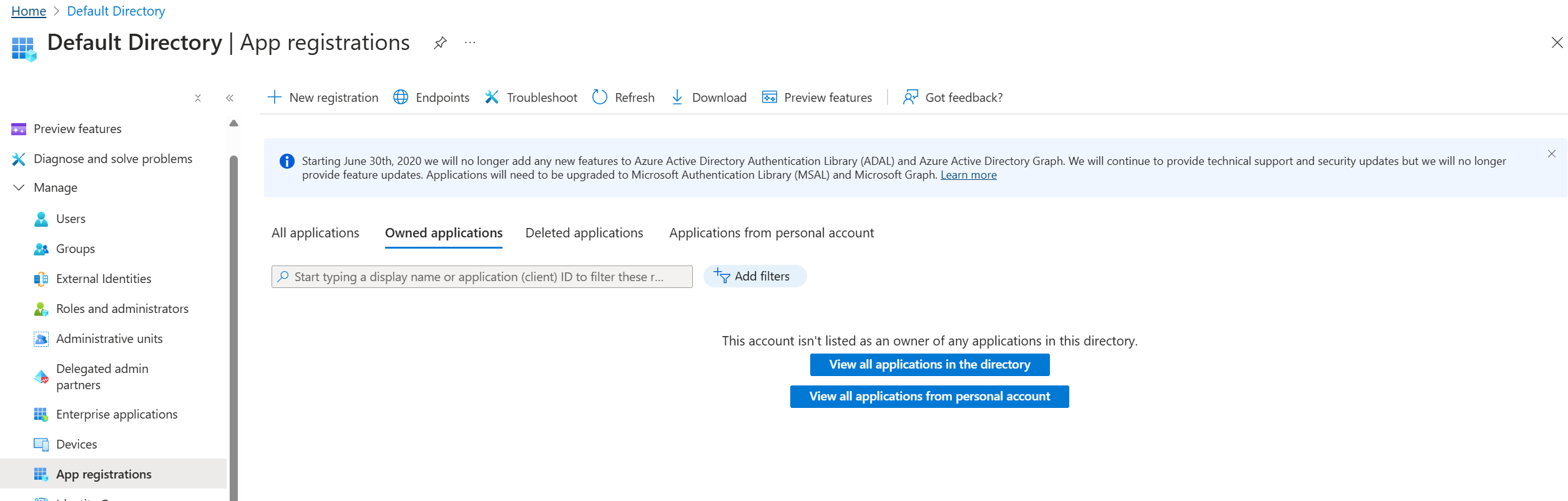
Step 9: Access Azure Datalake storage from Databricks using the **“service principles”** application.

Two Steps involved:

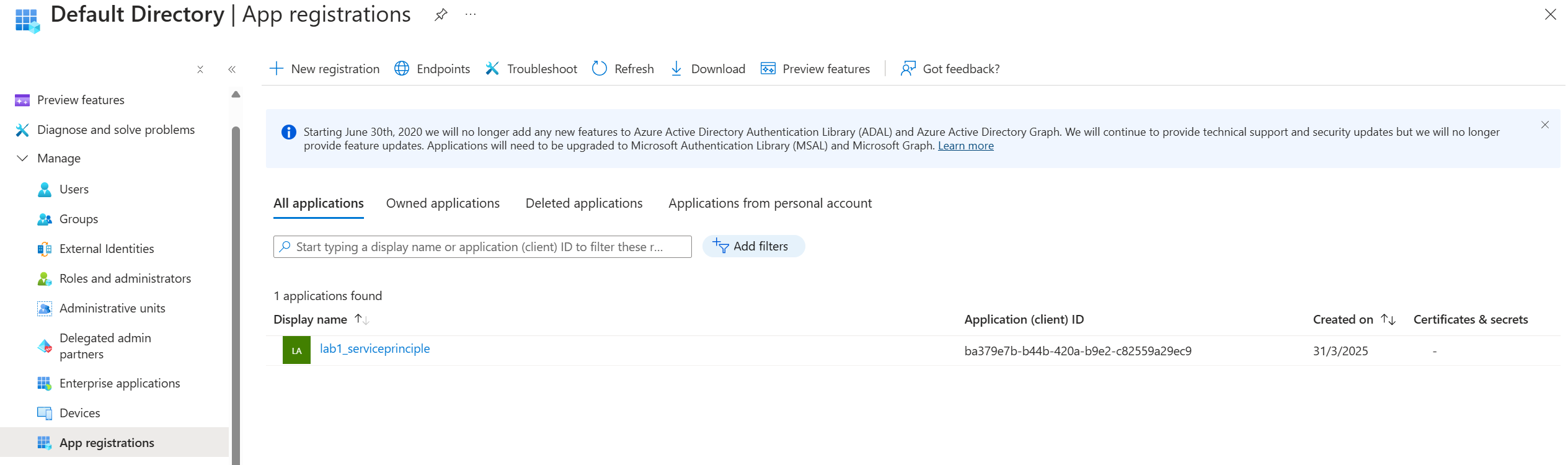
Step a: Create a “Service Principle” application and provide access to read the data from the datalake.

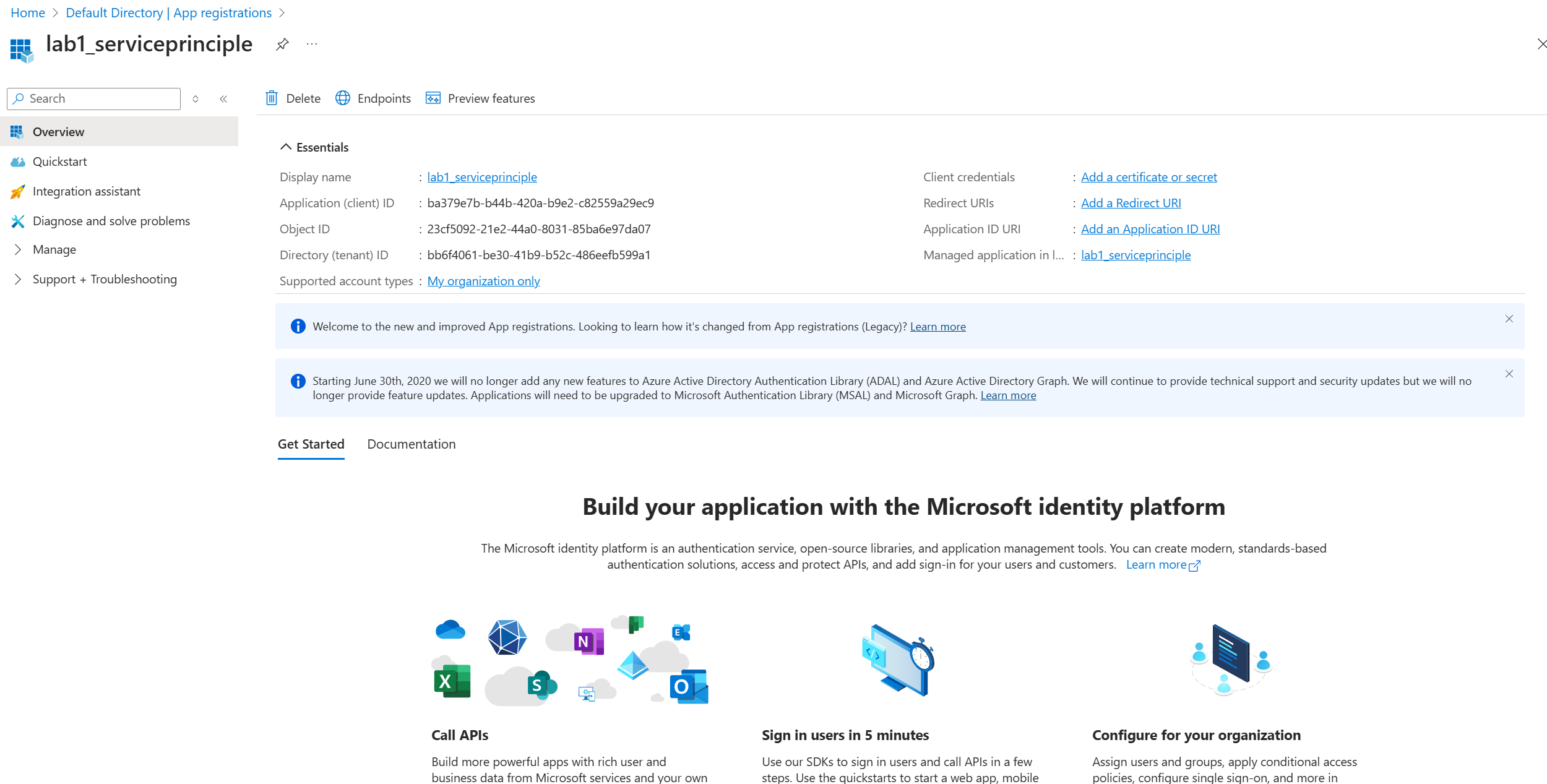
Step b: Import the credentials to the Databricks workspace.

Step 9a: Navigate to the “Microsoft EntraID” service and then to the “App Registrations”.



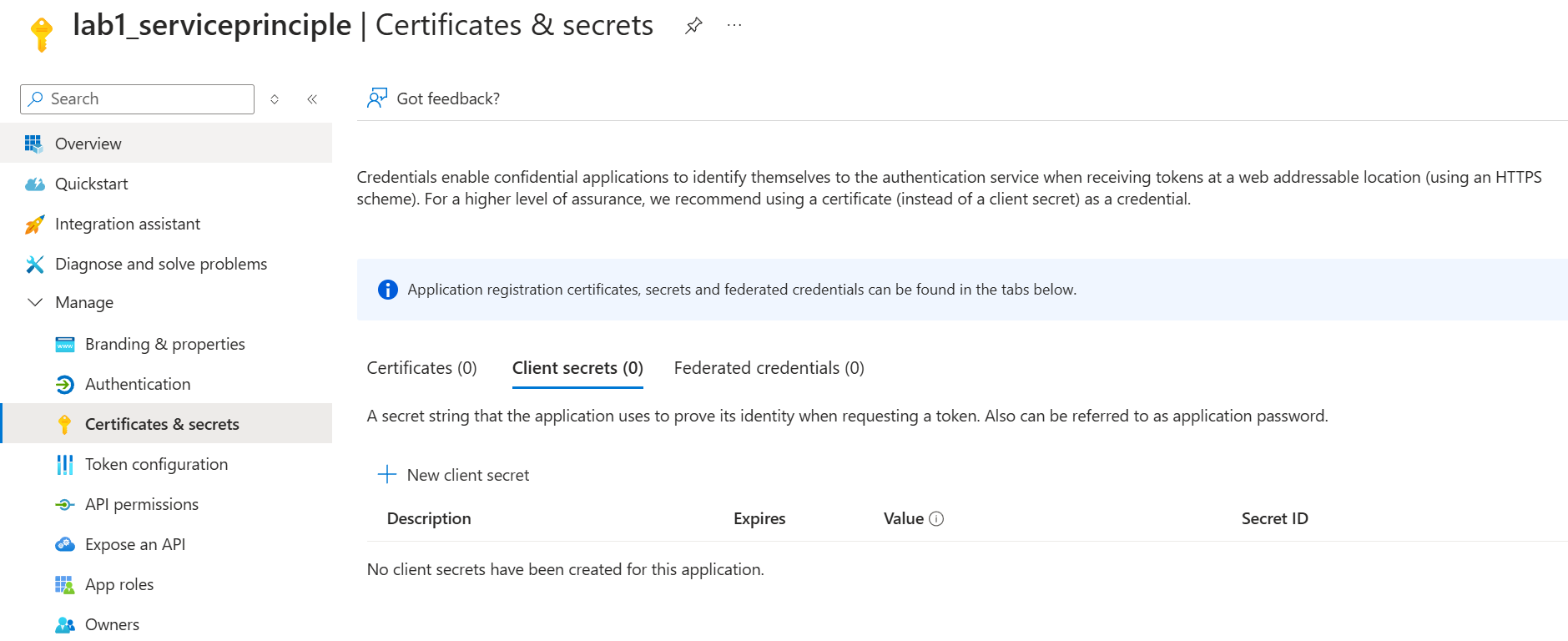
Step 9b: Click on “New registration”.





Copy the Application (client) ID and Directory(tenant) ID from the page.

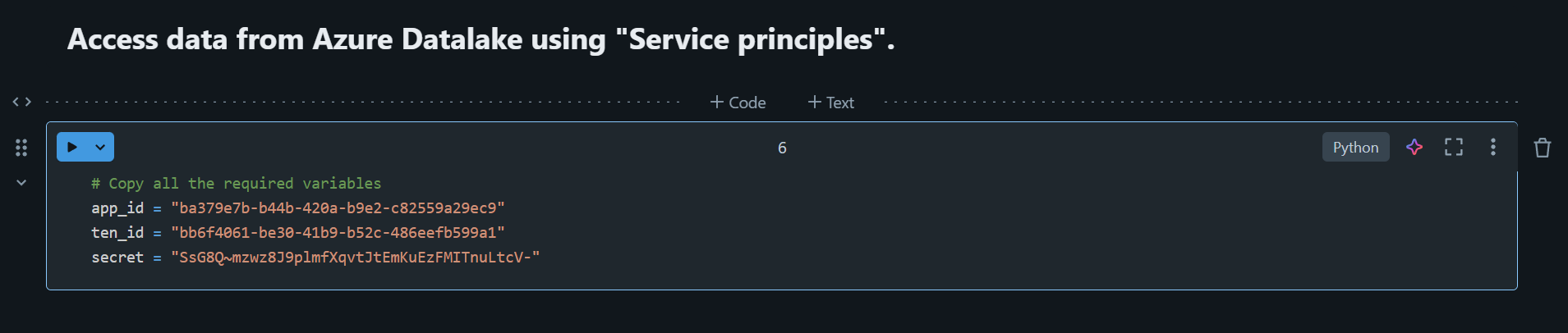
Step 9c: Navigate to “Certificates & Secrets” and create a “New client secret”. Make sure to copy the “value” because it appears only once.



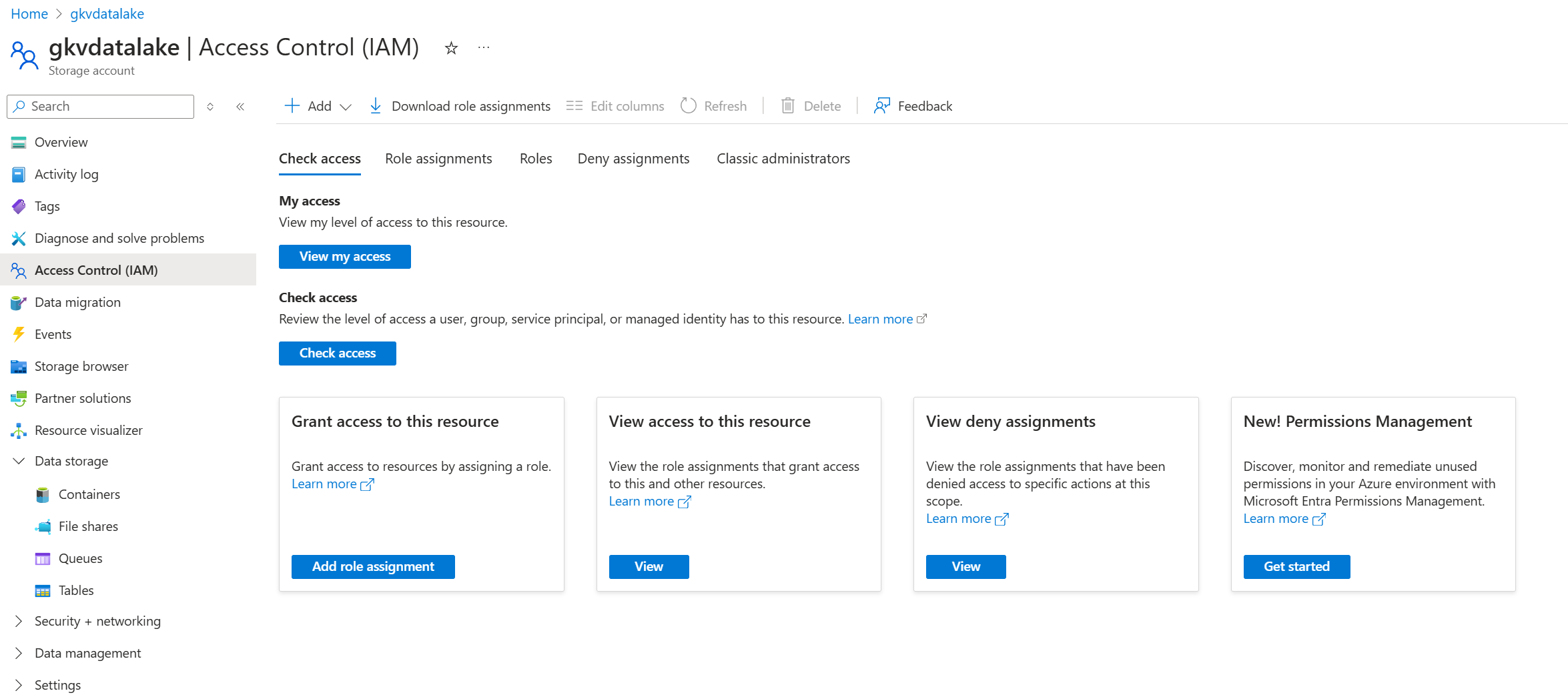


Step 9d: Now navigate to the Databricks Workspace and create a notebook.

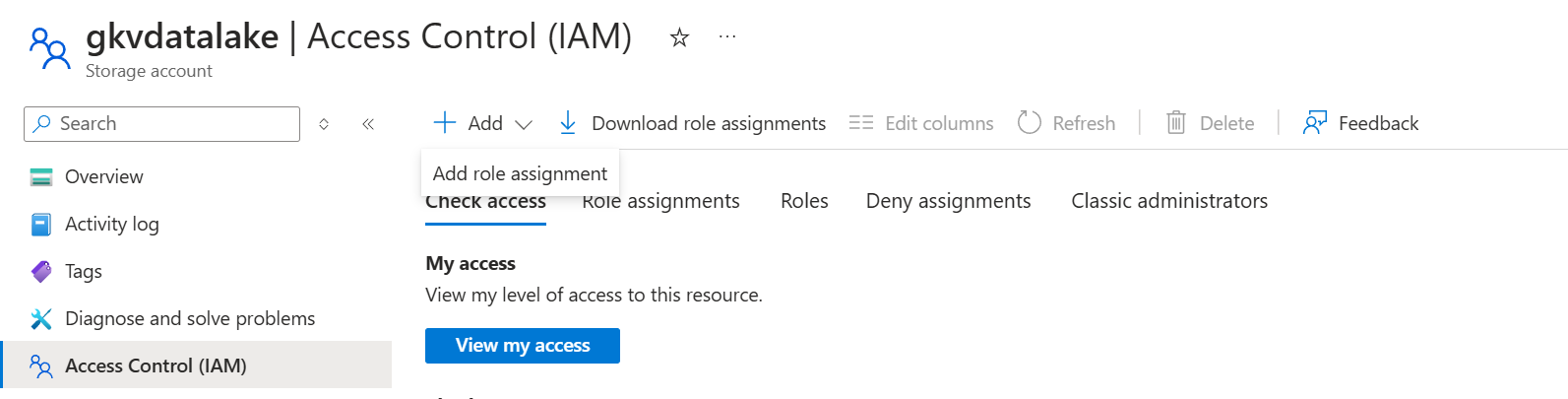
Include all the relevant details required from Azure

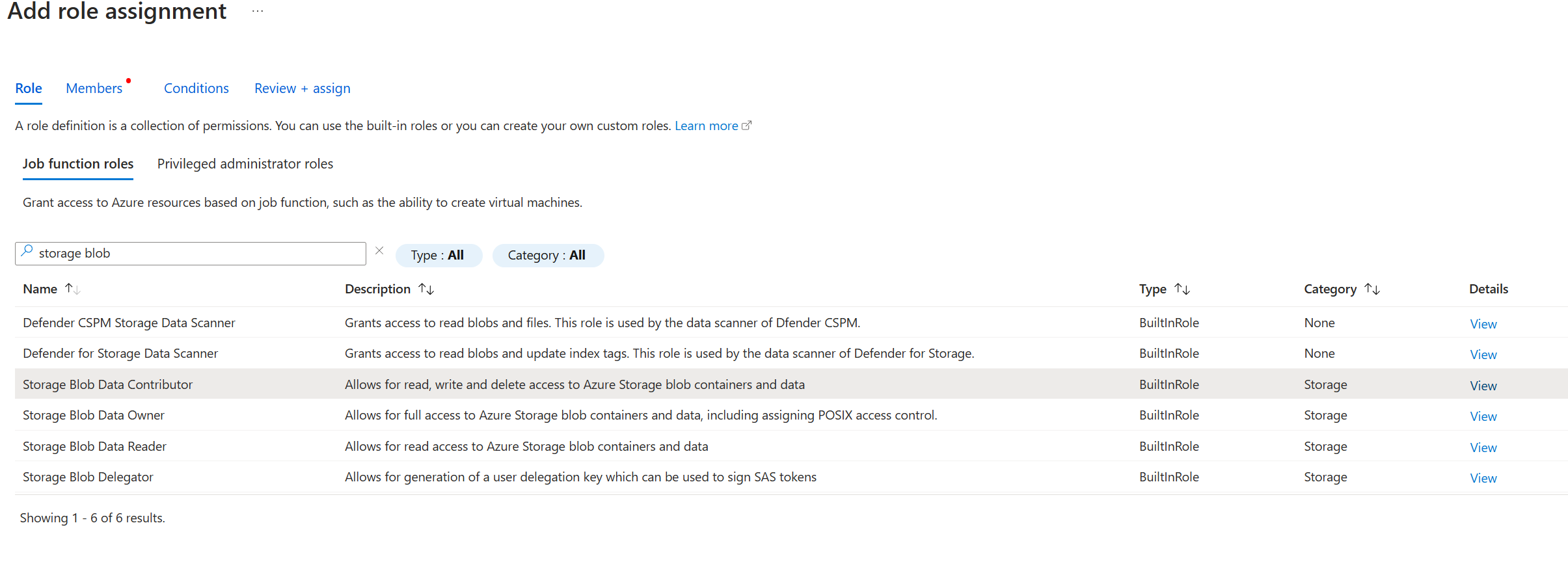


Step 9e: Again navigate to Azure and then navigate to the storage area where the data is available and then navigate to Access Control (IAM)

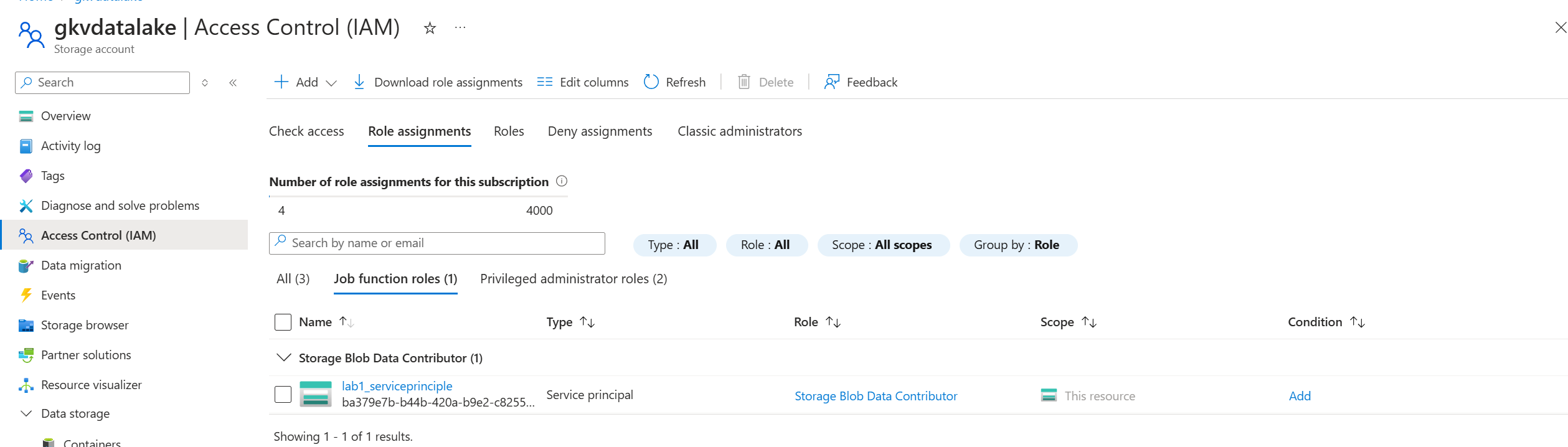


Step 9f: Create a new role assignment.

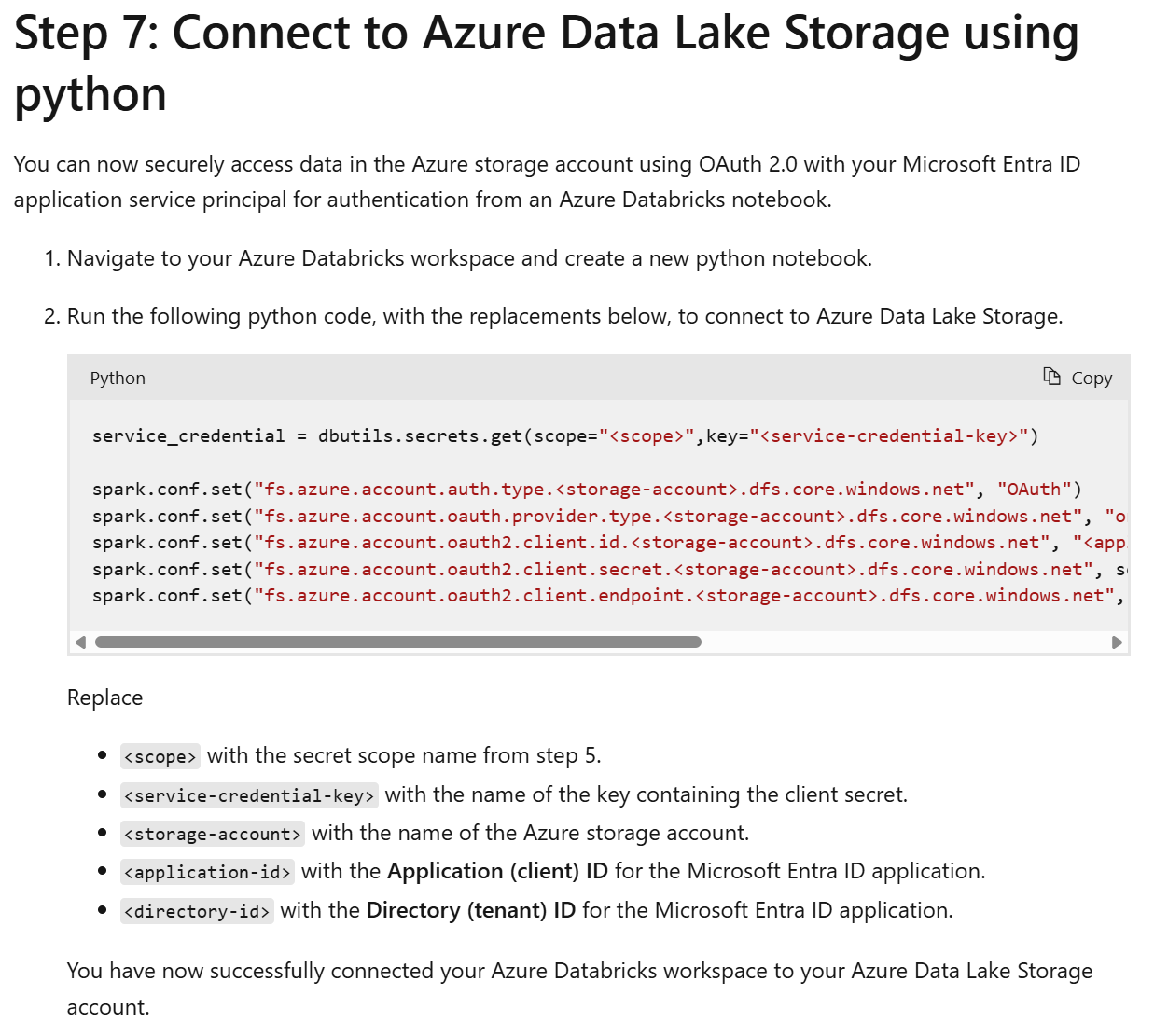


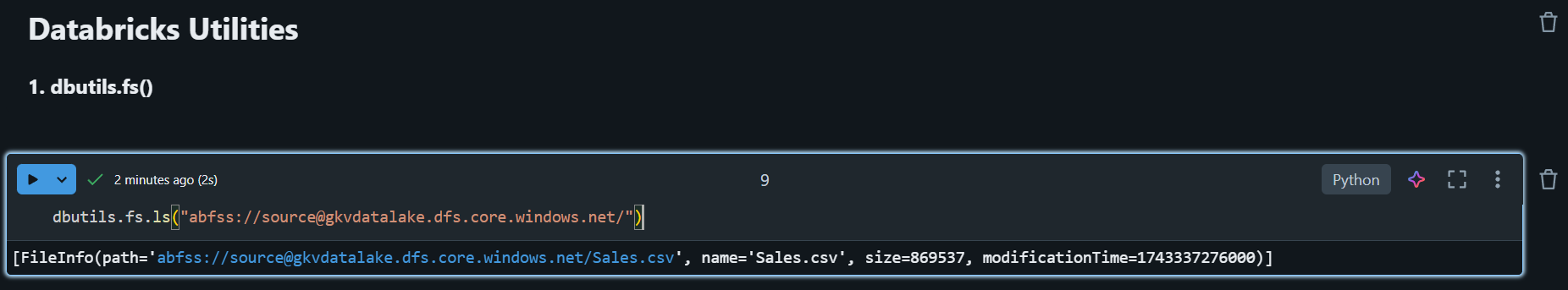


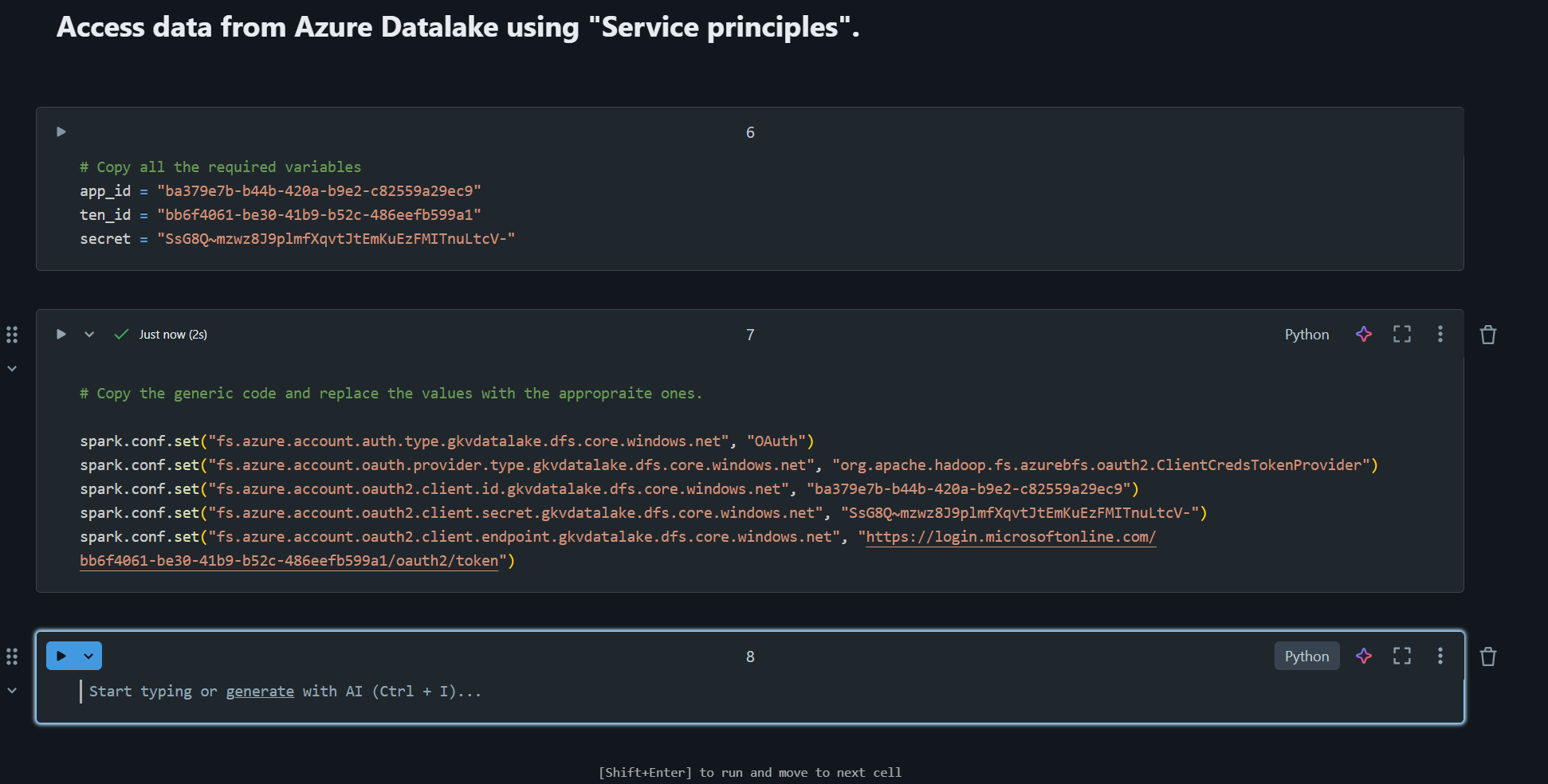




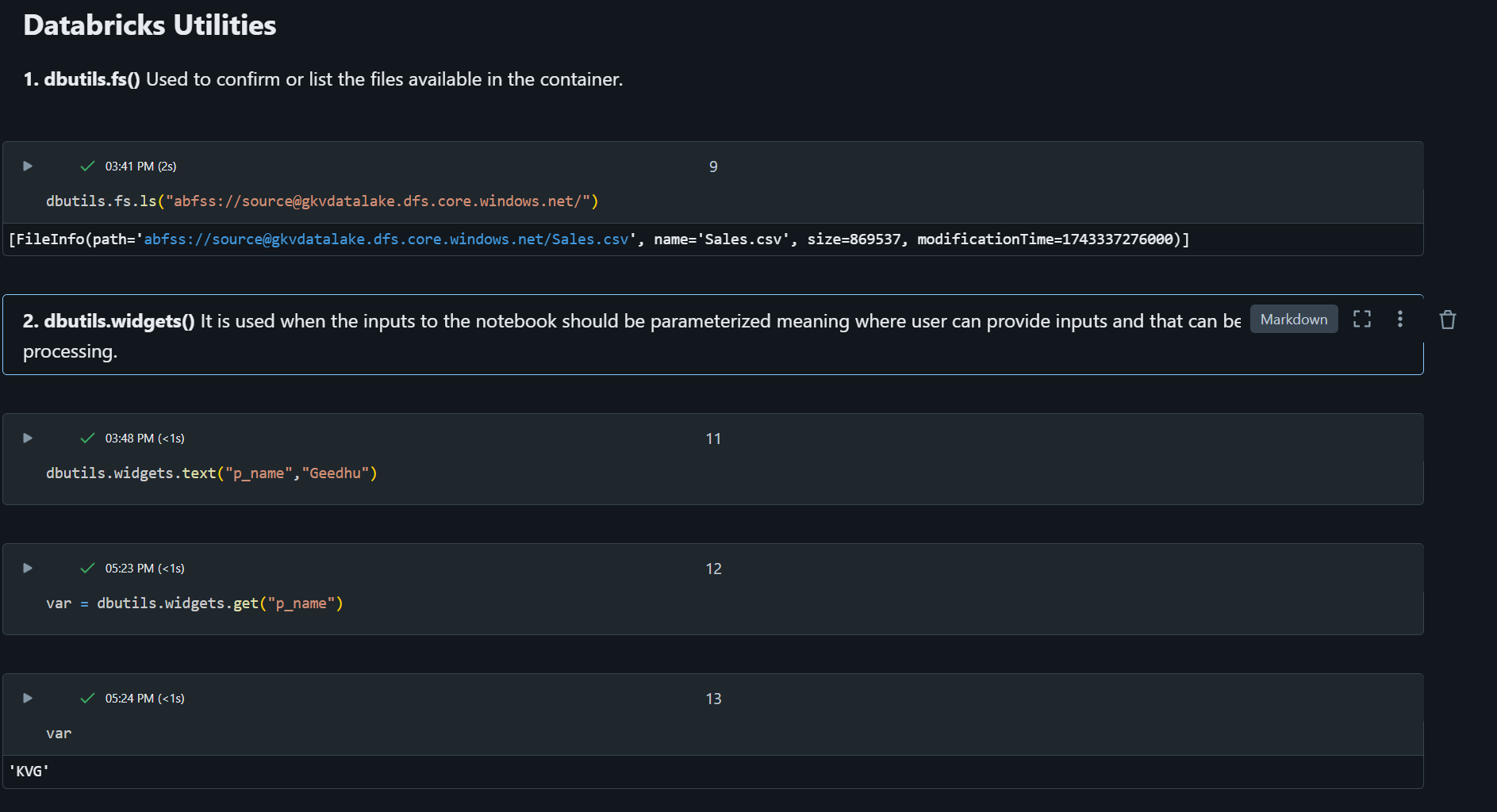
Step 9g: Copy the code from already existing Microsoft doc to the Databricks notebook.





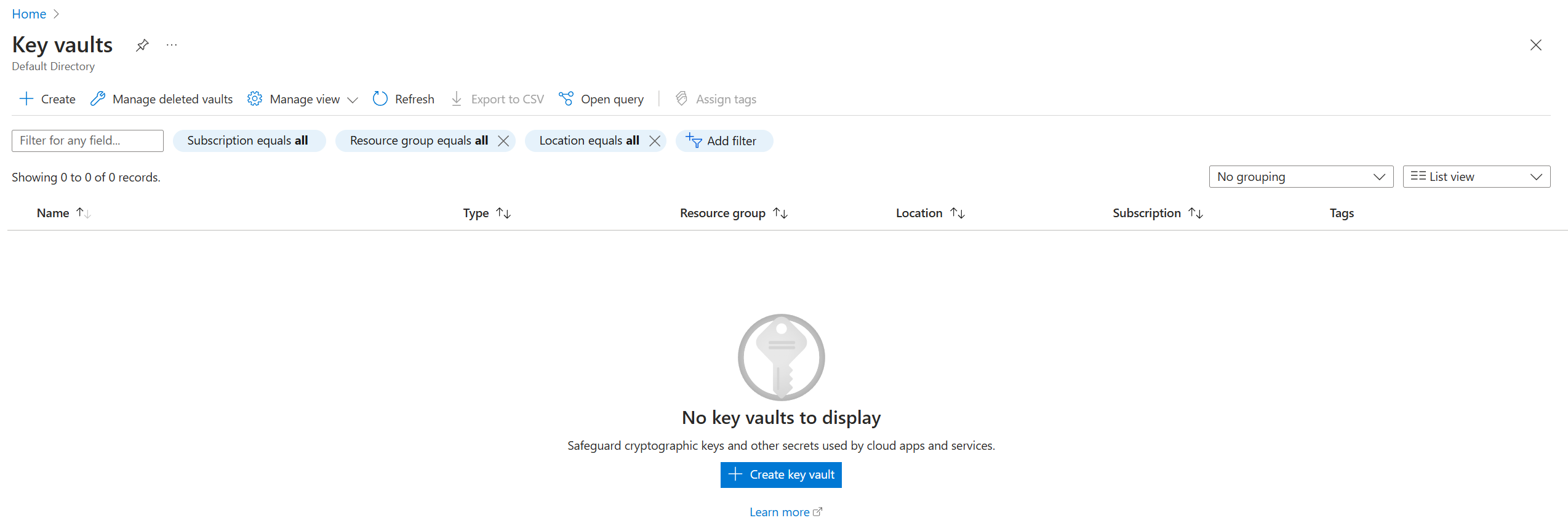


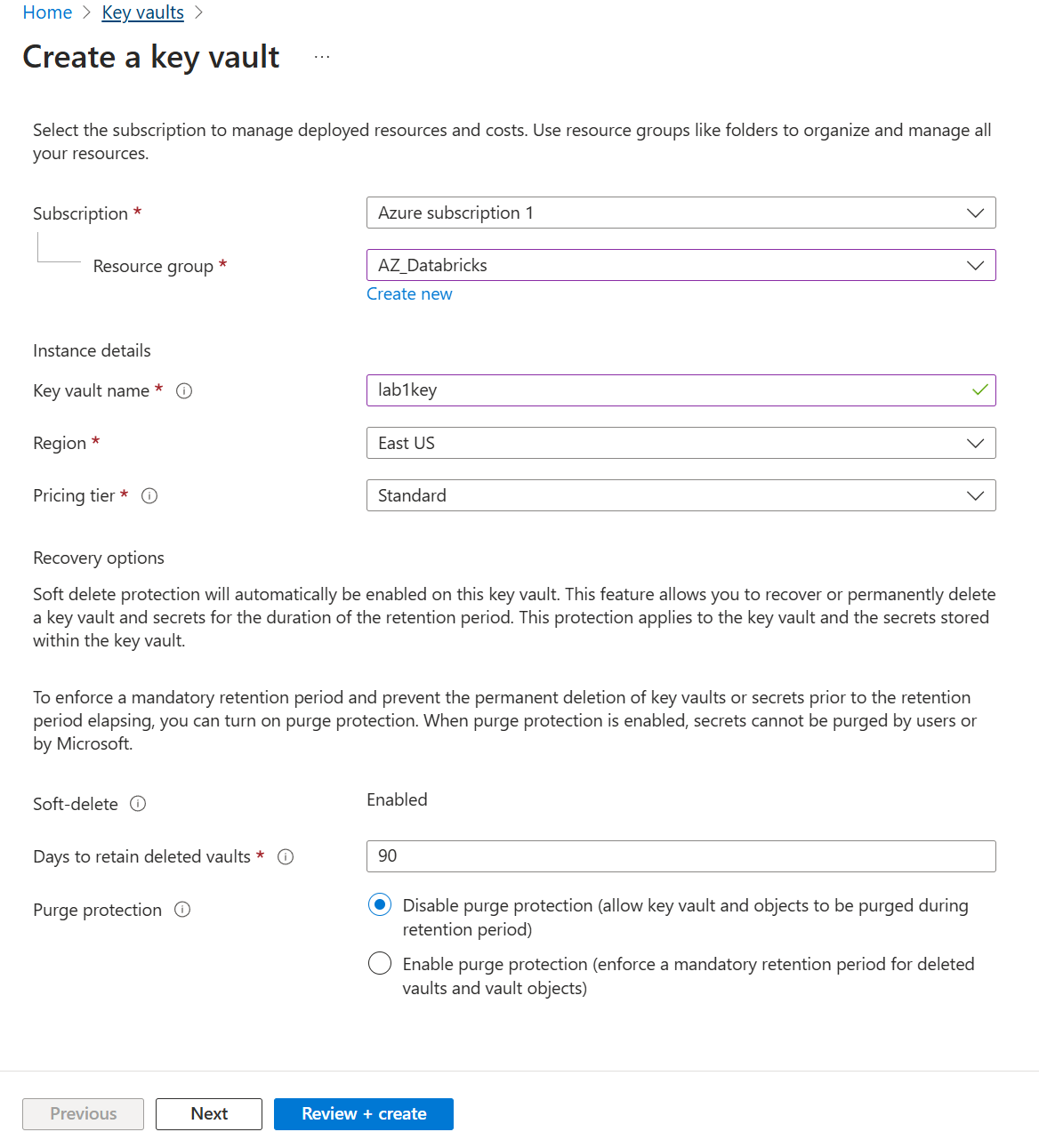
Step 10: Understand Databricks Utilities.





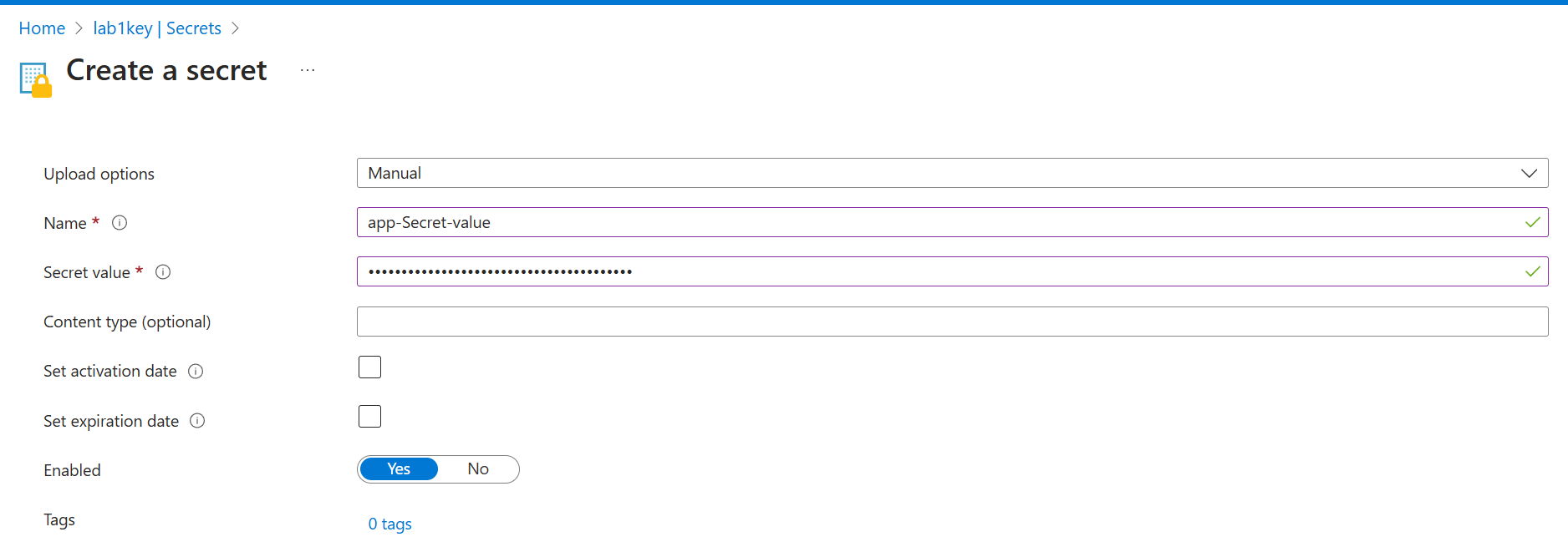
In normal homepage, search “Key vaults” as shown below:







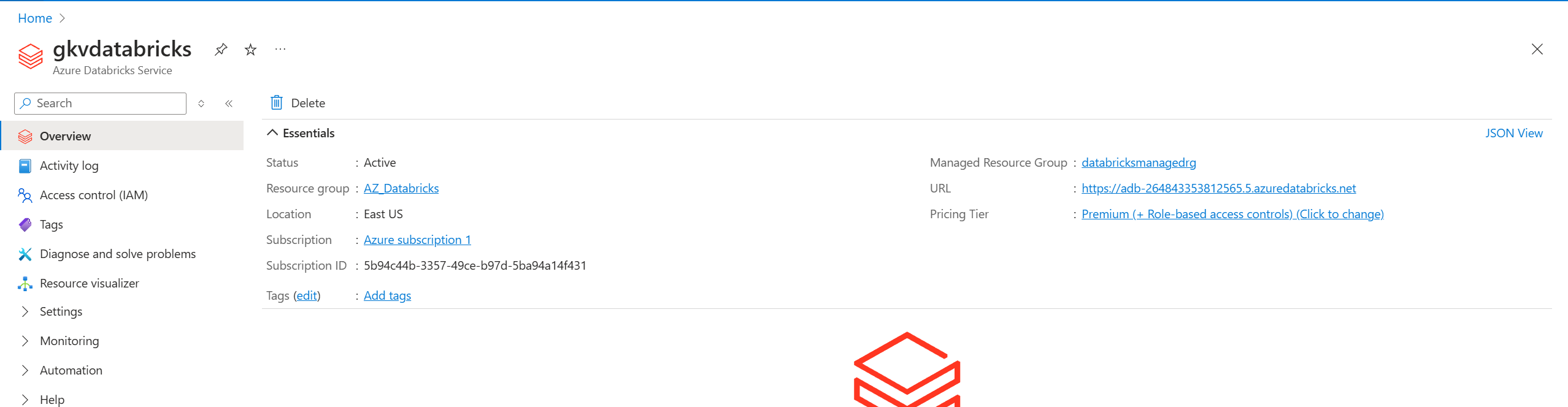
Create a secret object in the key vaults.

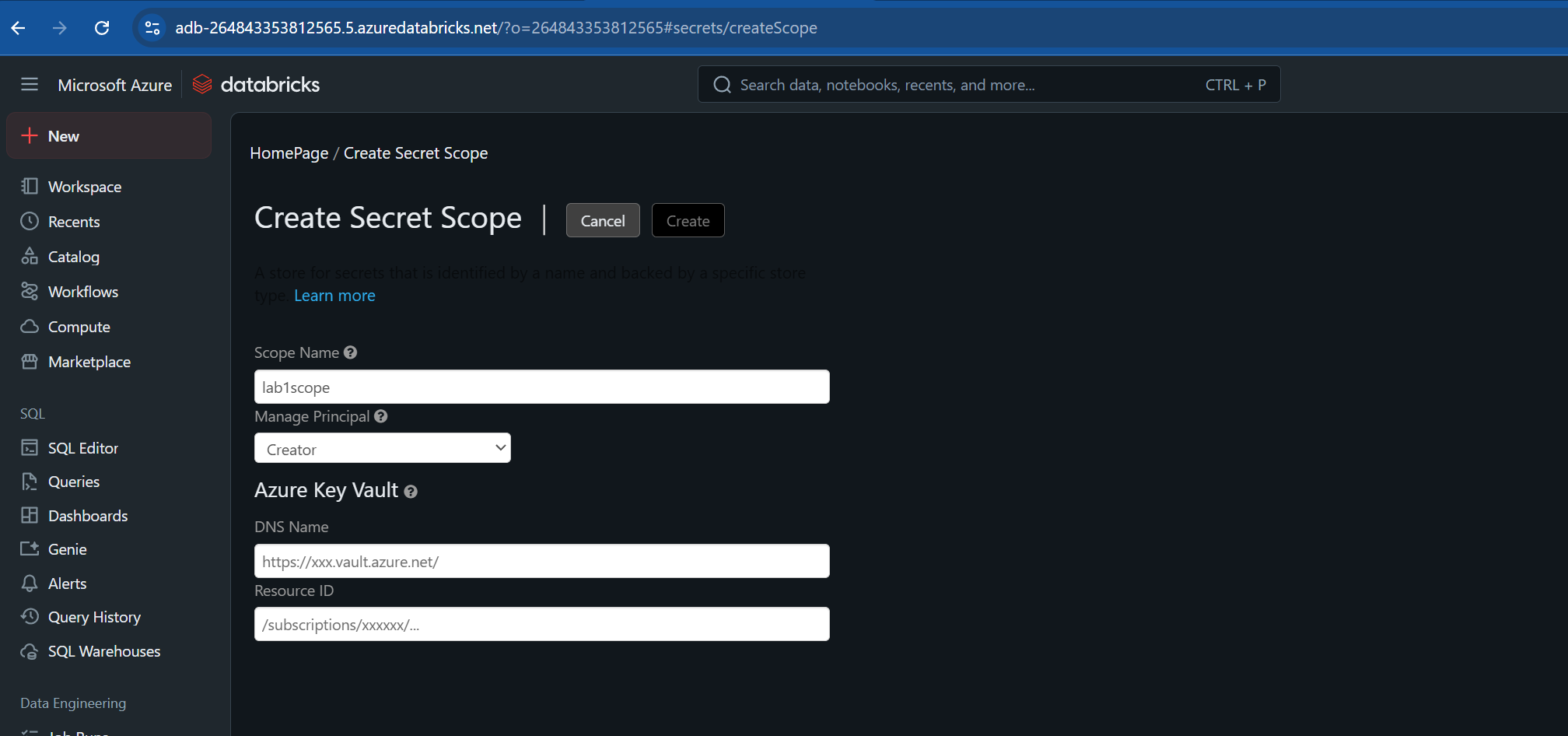


Now secret is created in Azure key vault.

Now connect this secret to the Databricks workspace.

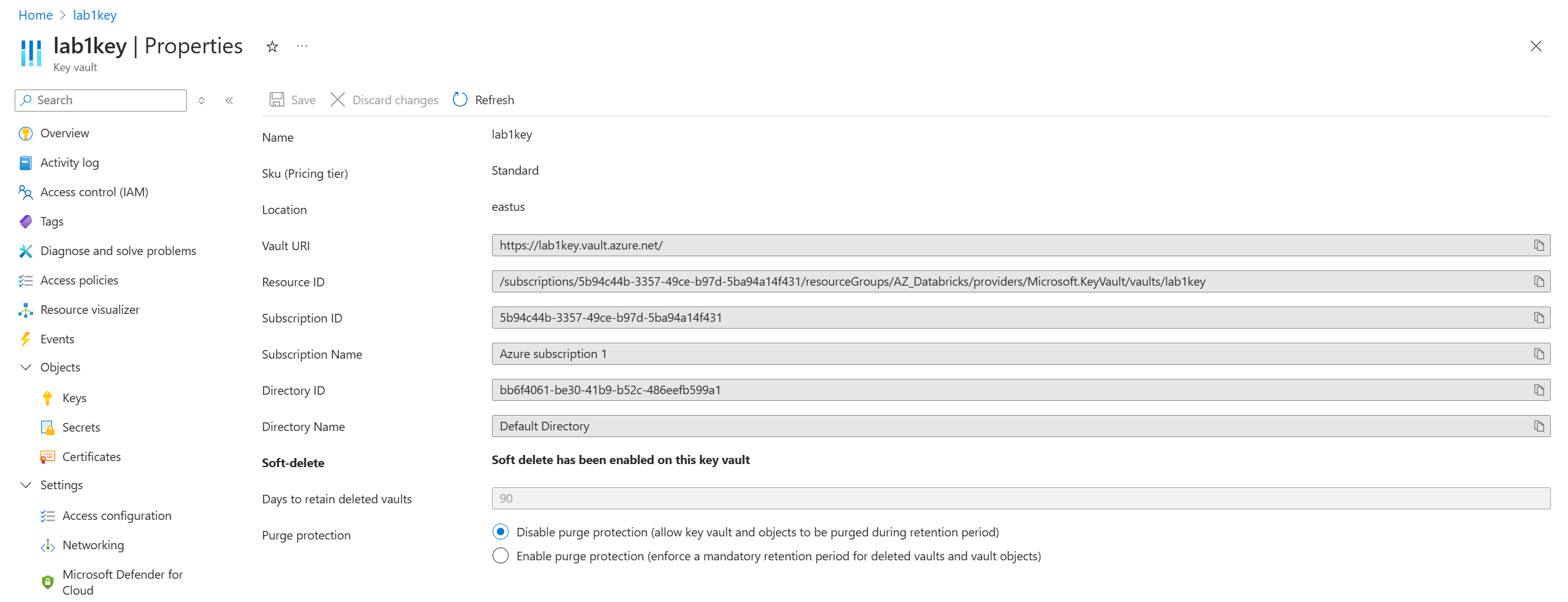
Step 1: go to azure homepage and navigate azure Databricks page created and copy the URL.



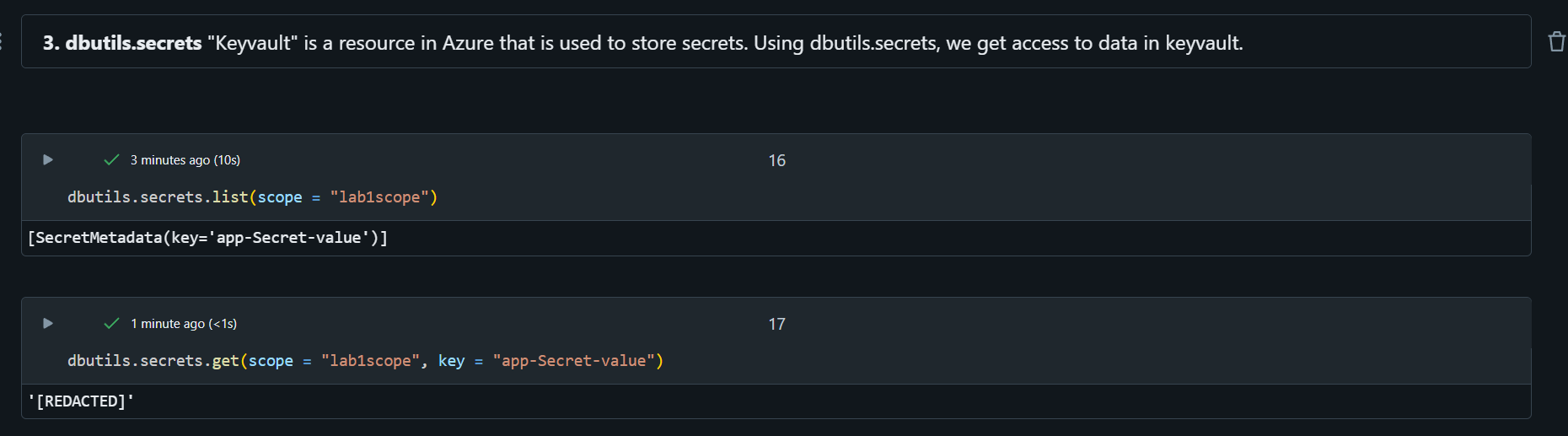
Step 2: Paste the URL in the new tab along with “#secrets/createScope”. 

Give any scope name

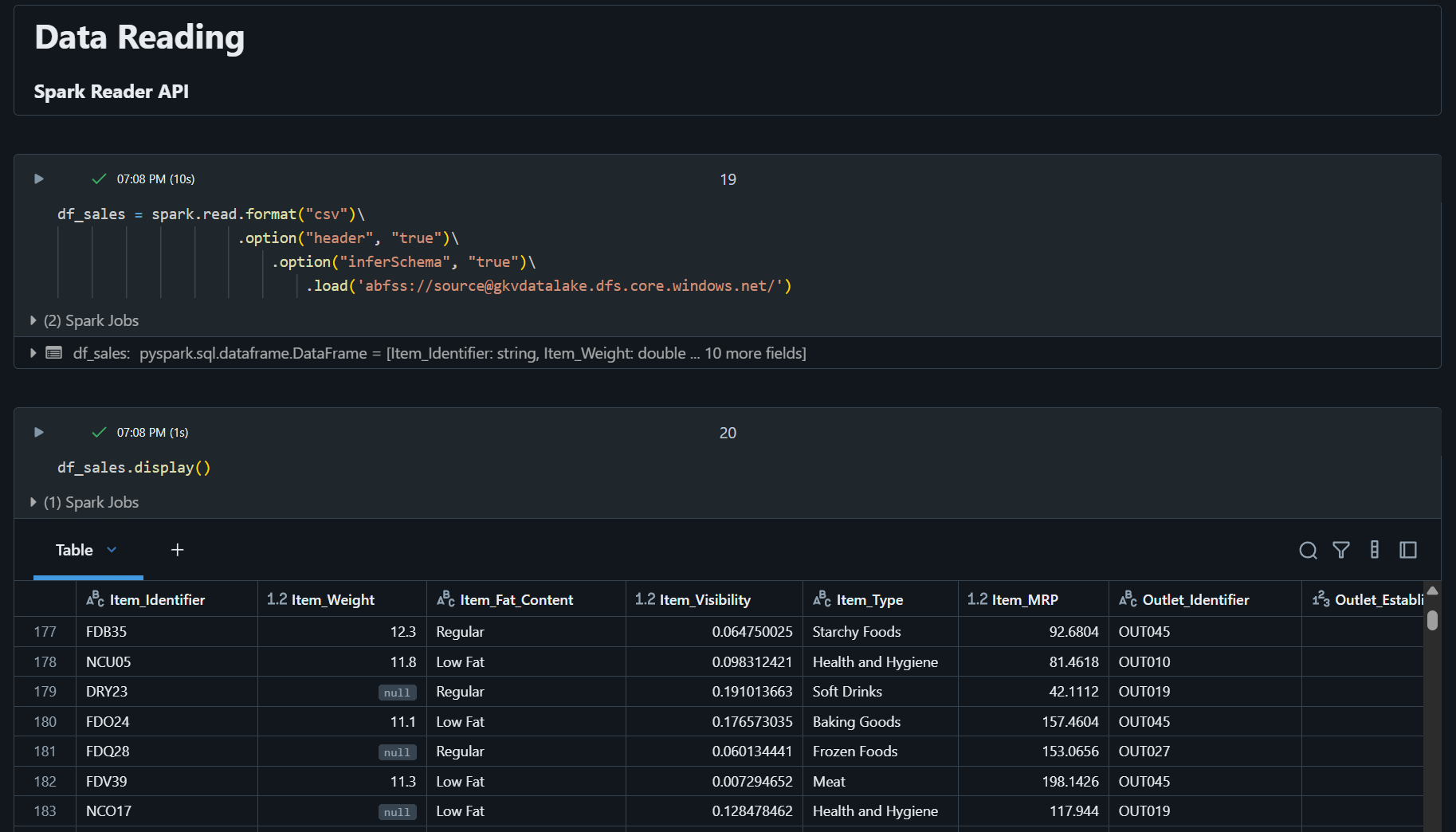
And for Azure key vault values copy data from Key vault created.



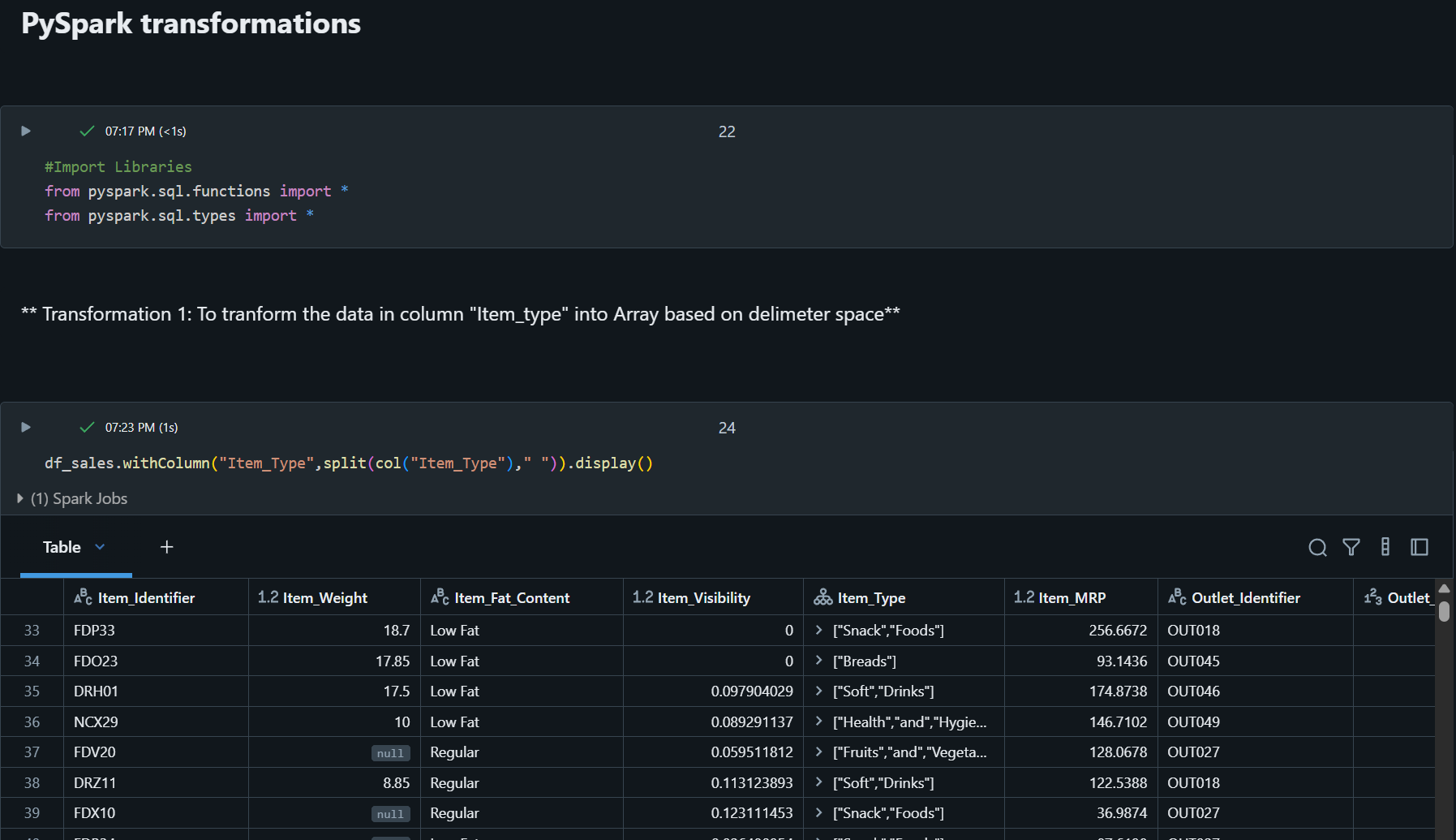
Vault URI is DNS name and Resource Id is resource ID.

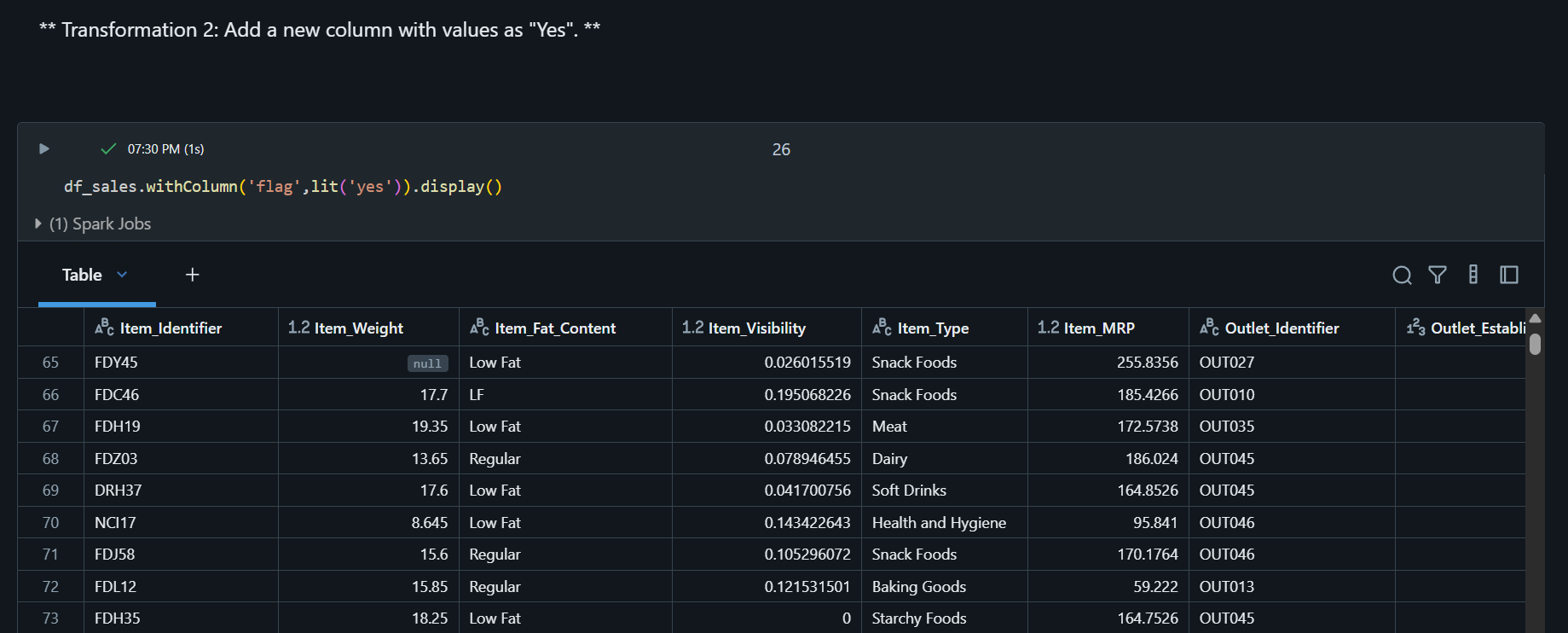


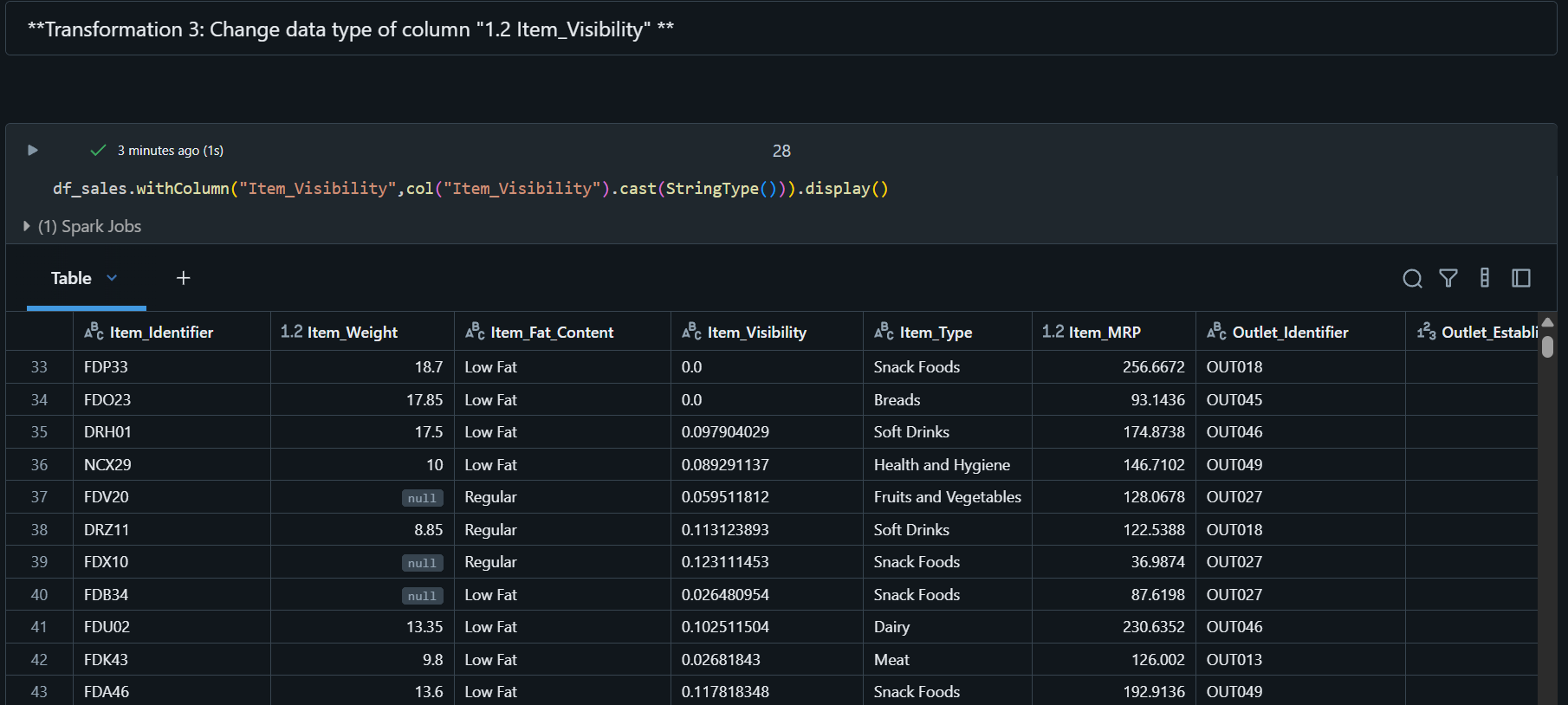
Step 11: Data Reading



Step 12: Performing transformations using pyspark







`