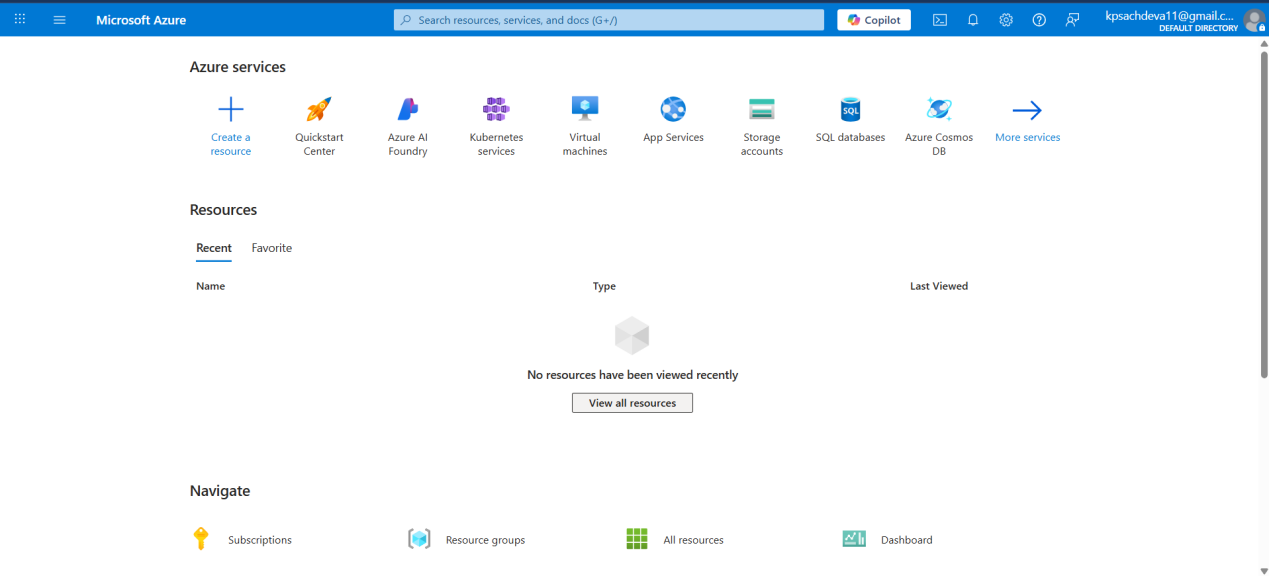
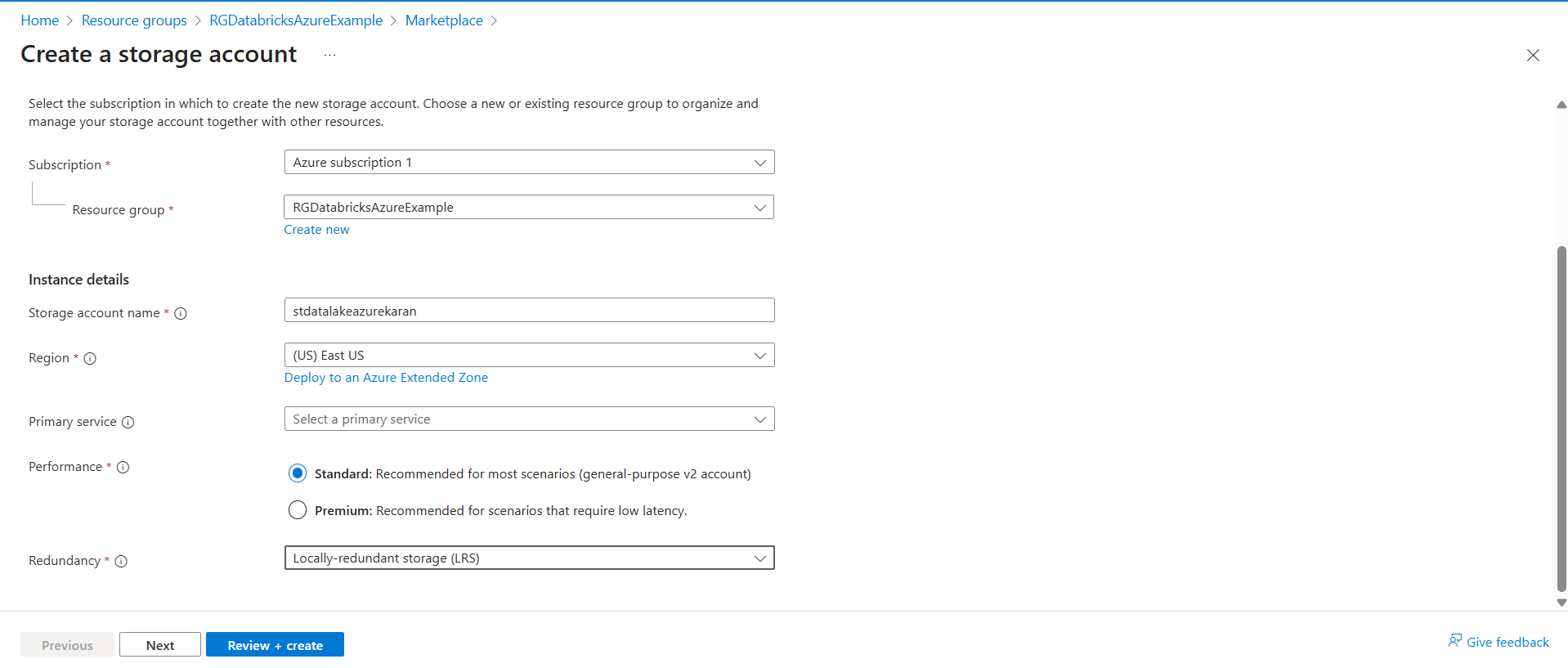
# Databricks using Azure Account

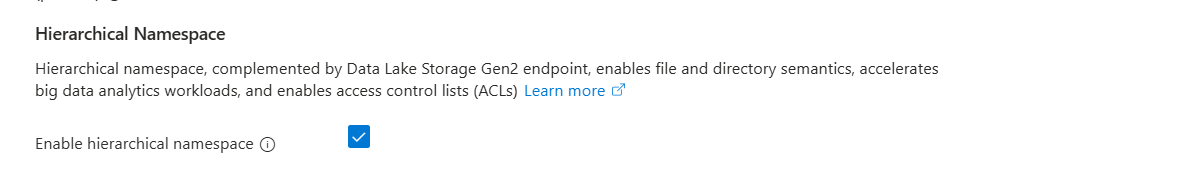
1. Microsoft Azure account homepage:



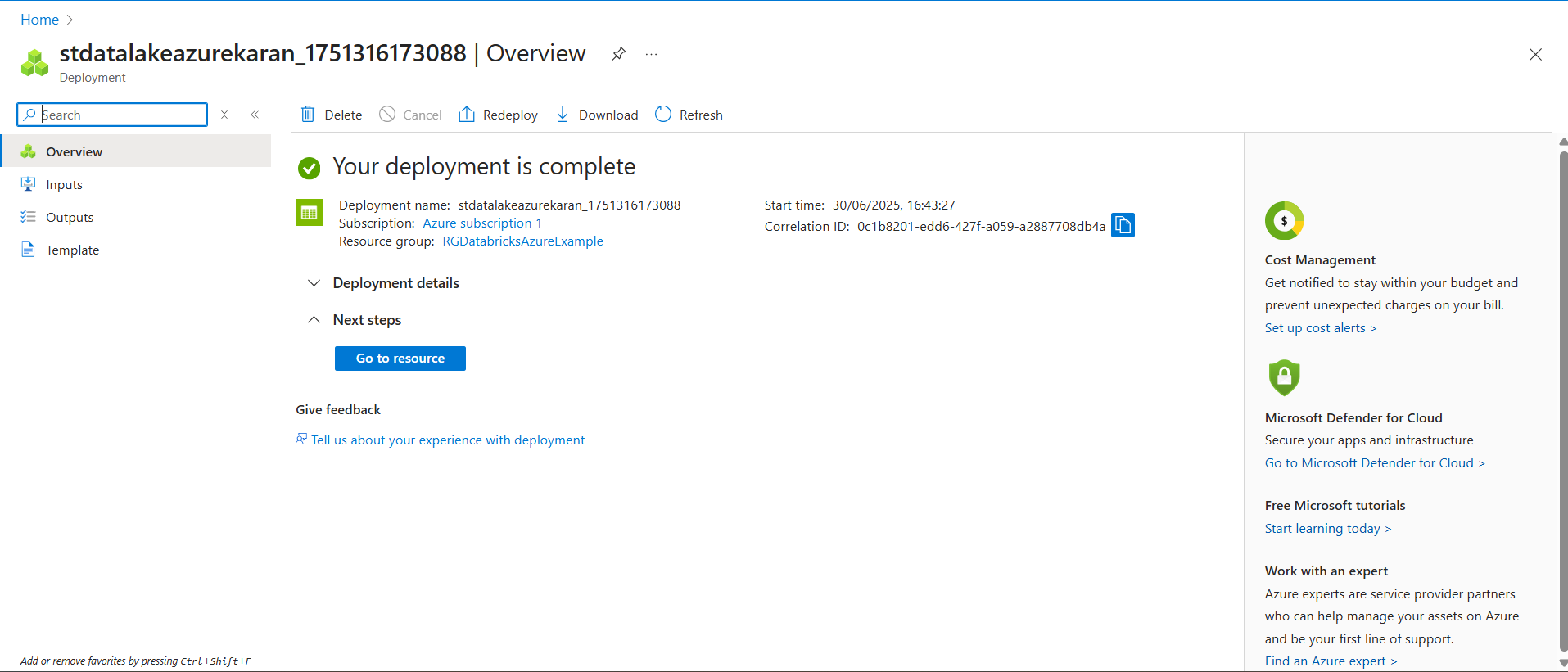
1. Creating a microsoft storage account in azure:



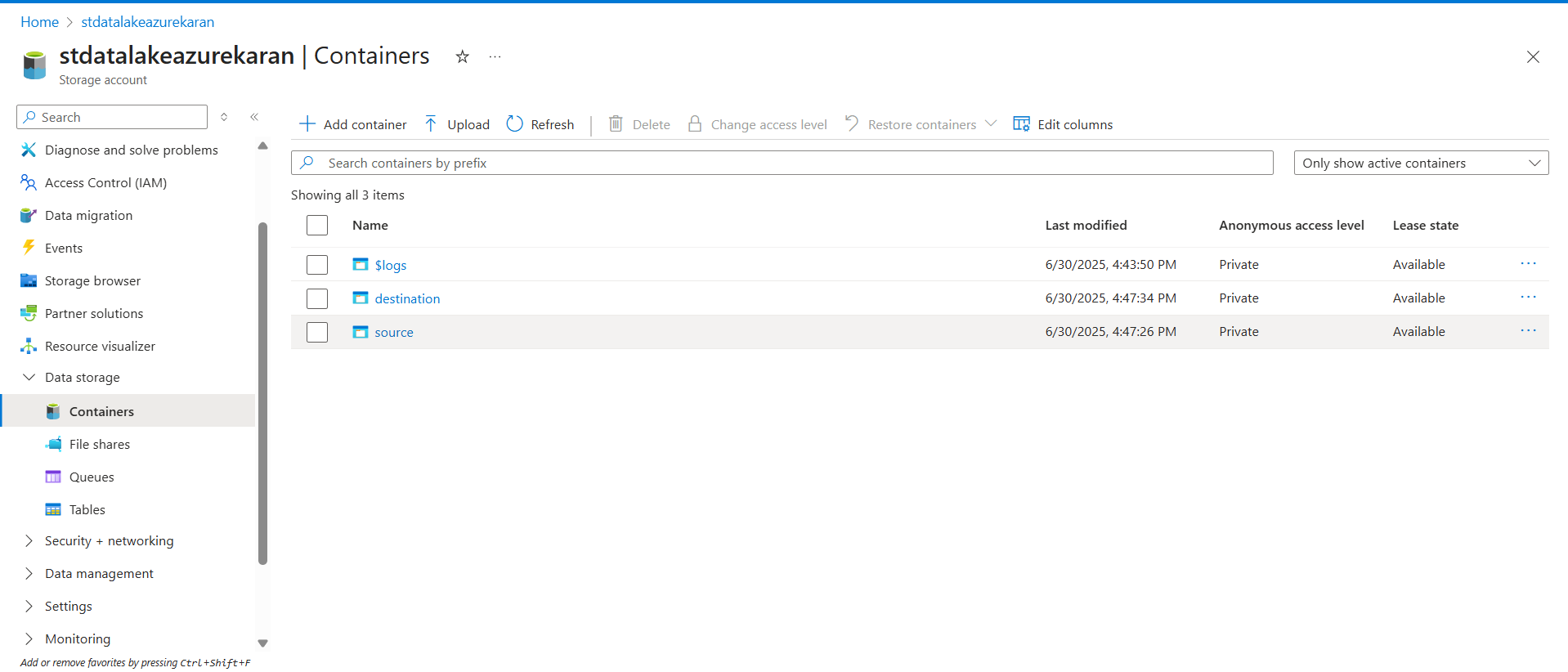
1. Creating a hierarchical namespace for our datalake



1. Data storage is created in azure

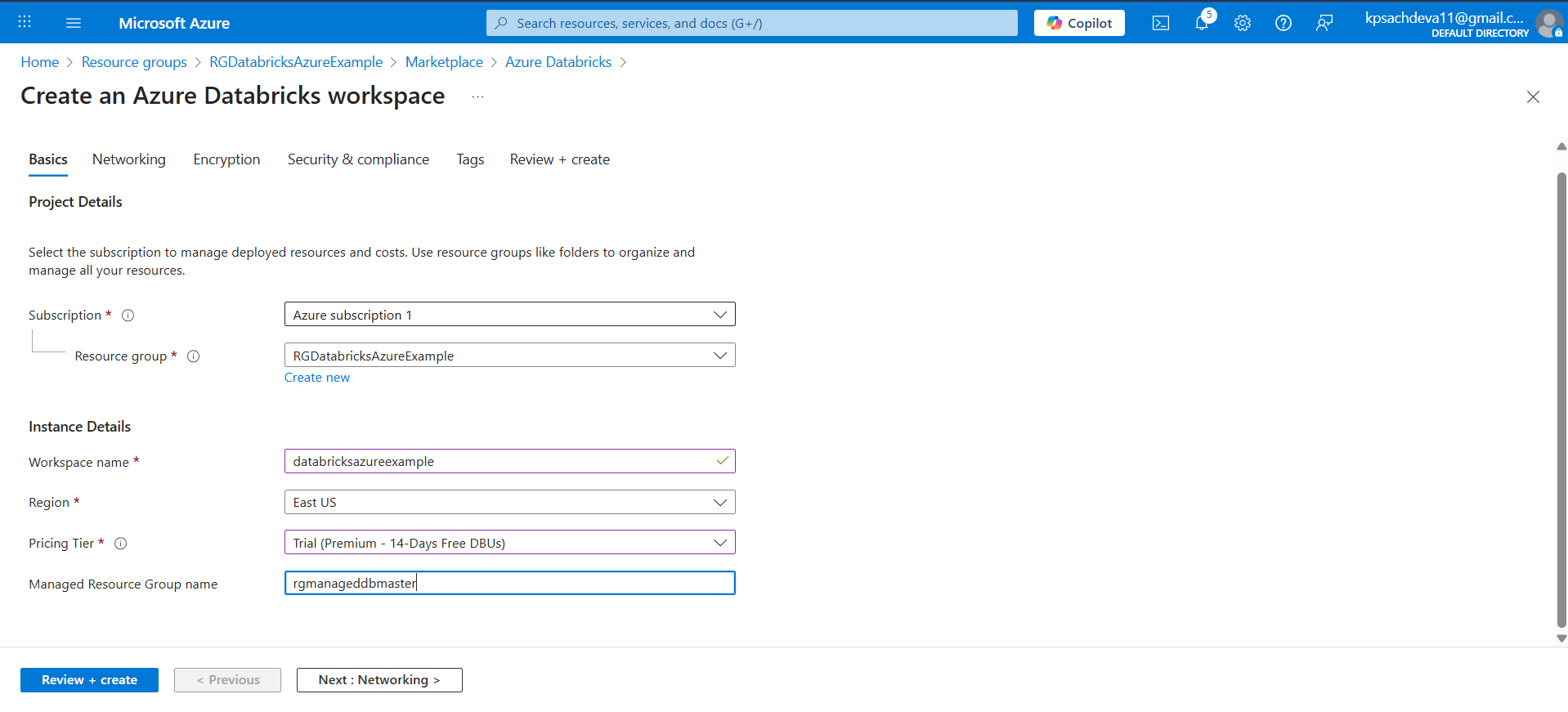


1. Create two containers source and destination

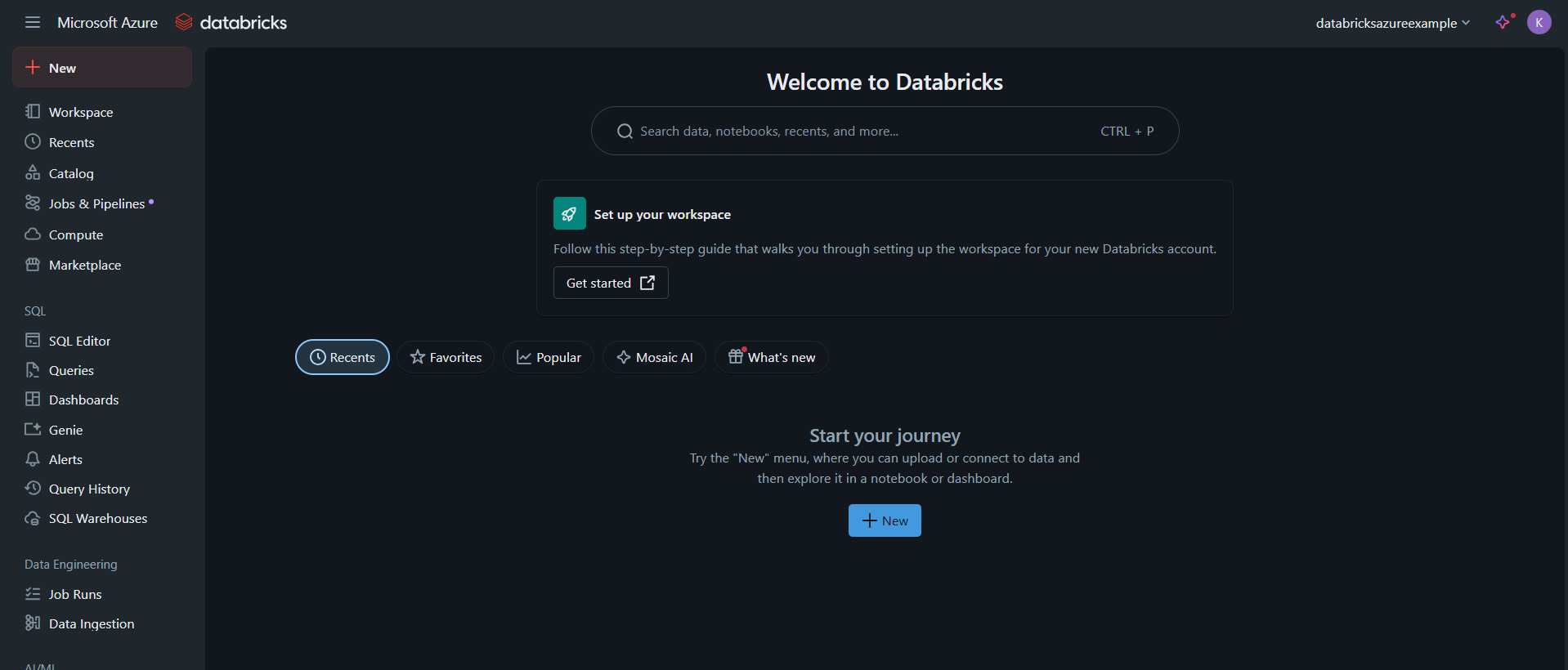


## Creating an Azure Databricks resource

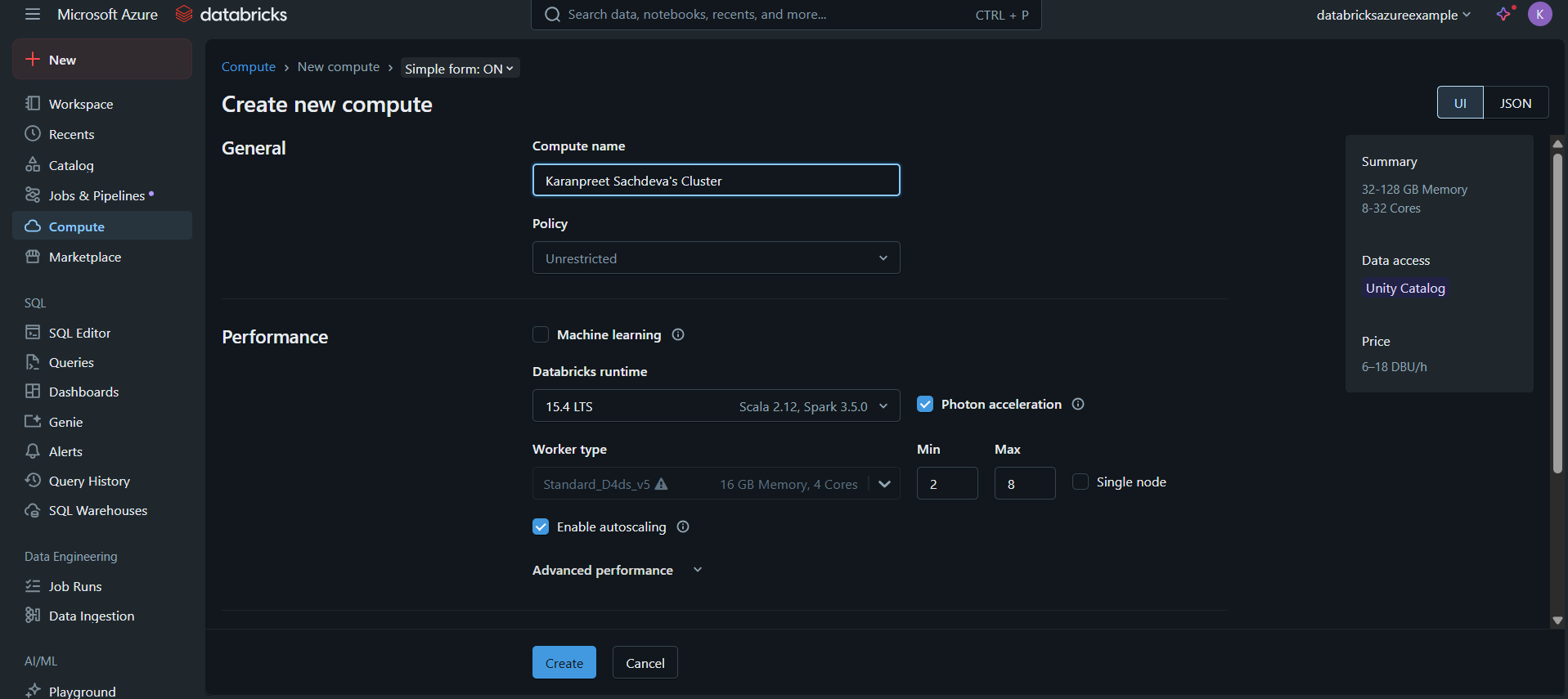
1. Creating an azure databricks resource



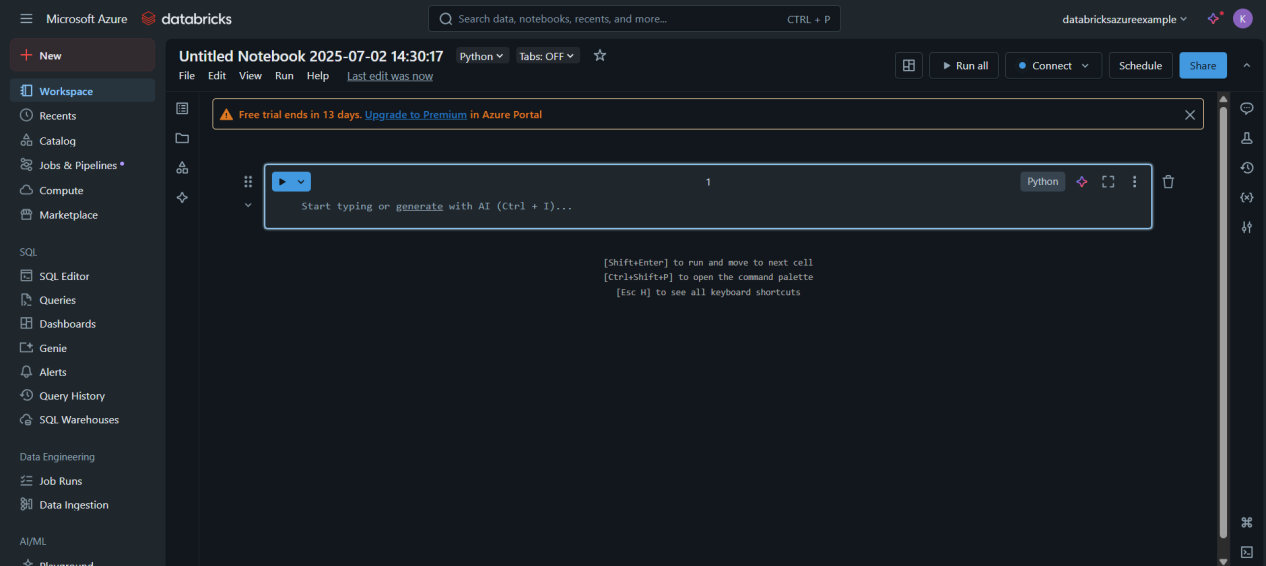
1. Launched Azure Databricks workspace



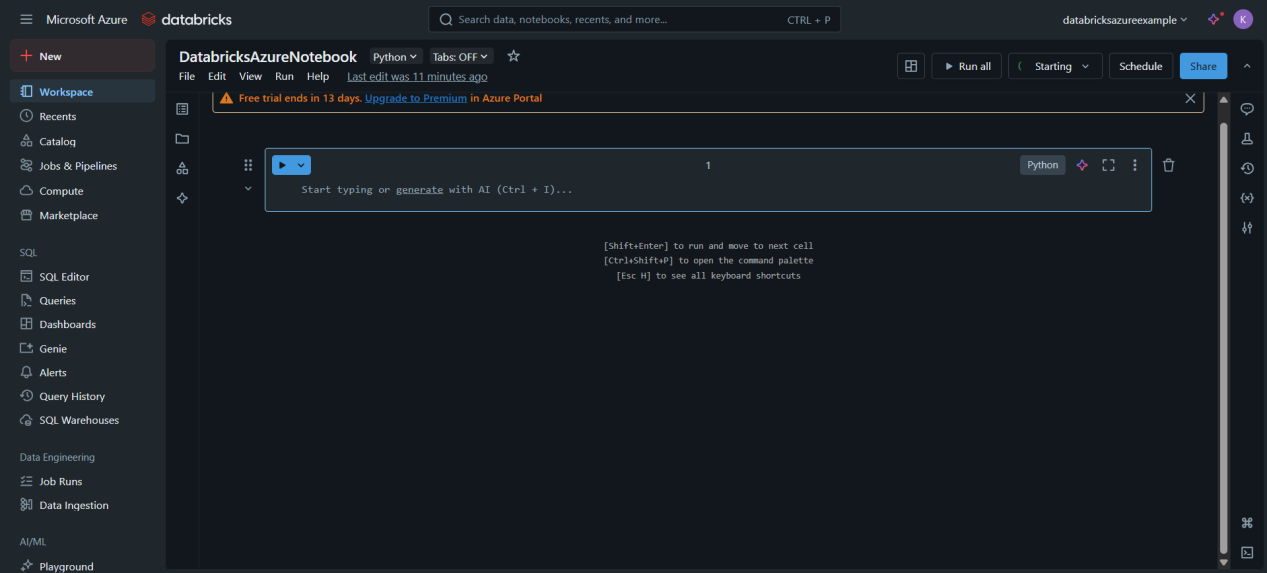
1. Creating the cluster



1. Creating the notebook

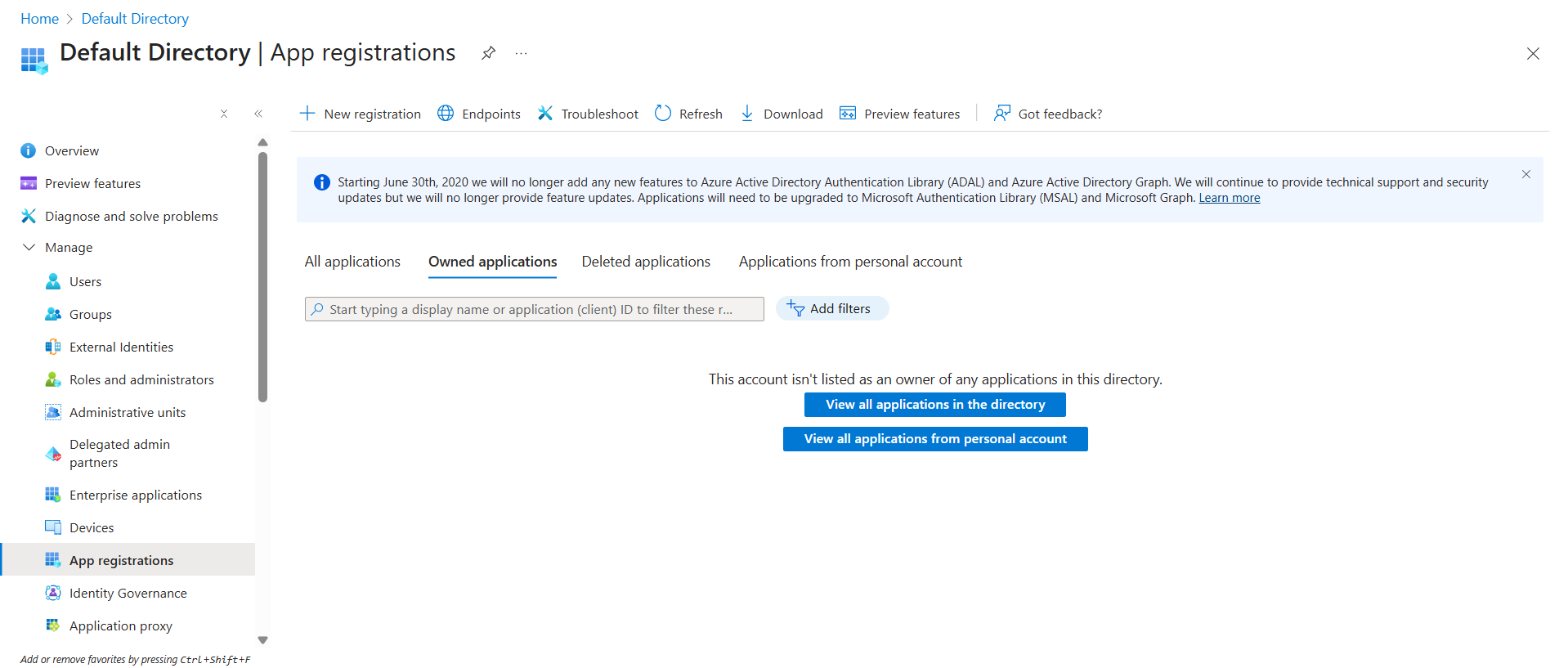


1. Connecting the notebook to the cluster we created

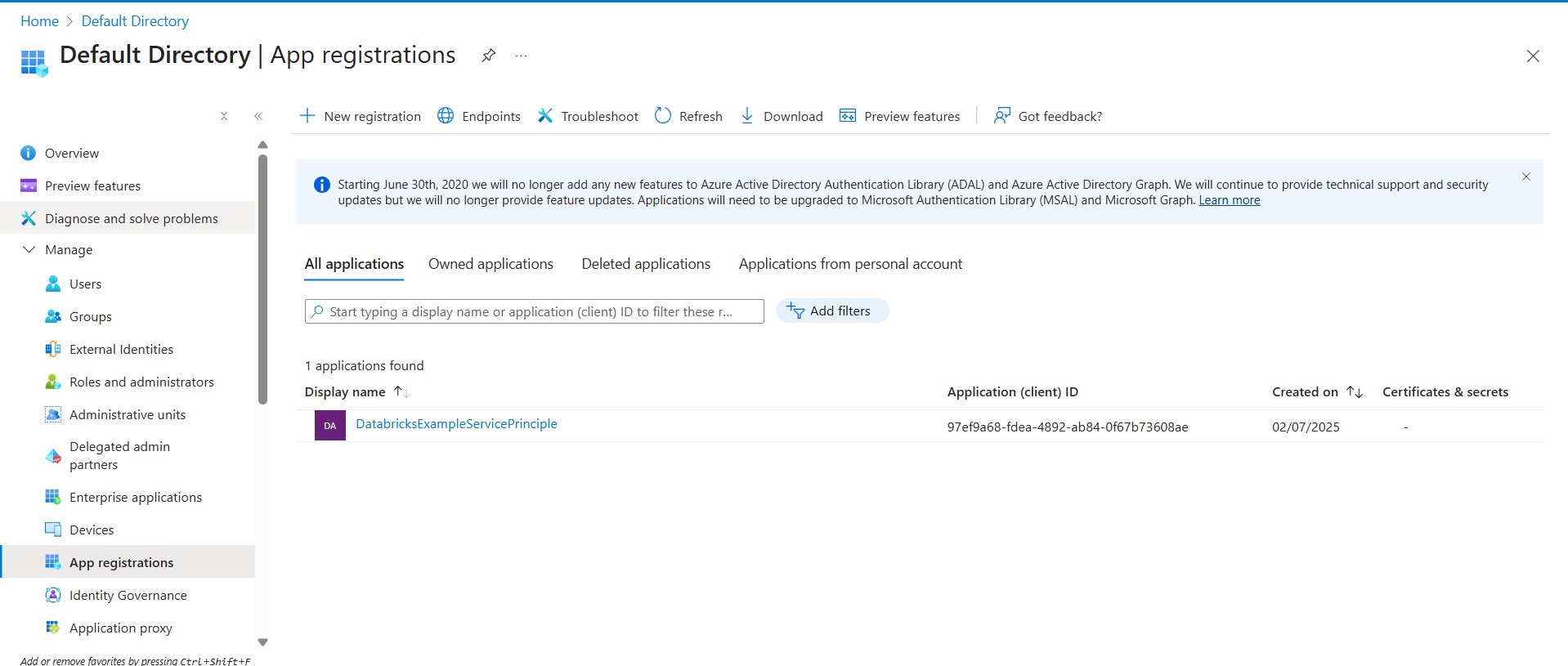


## Accessing the data from datalake

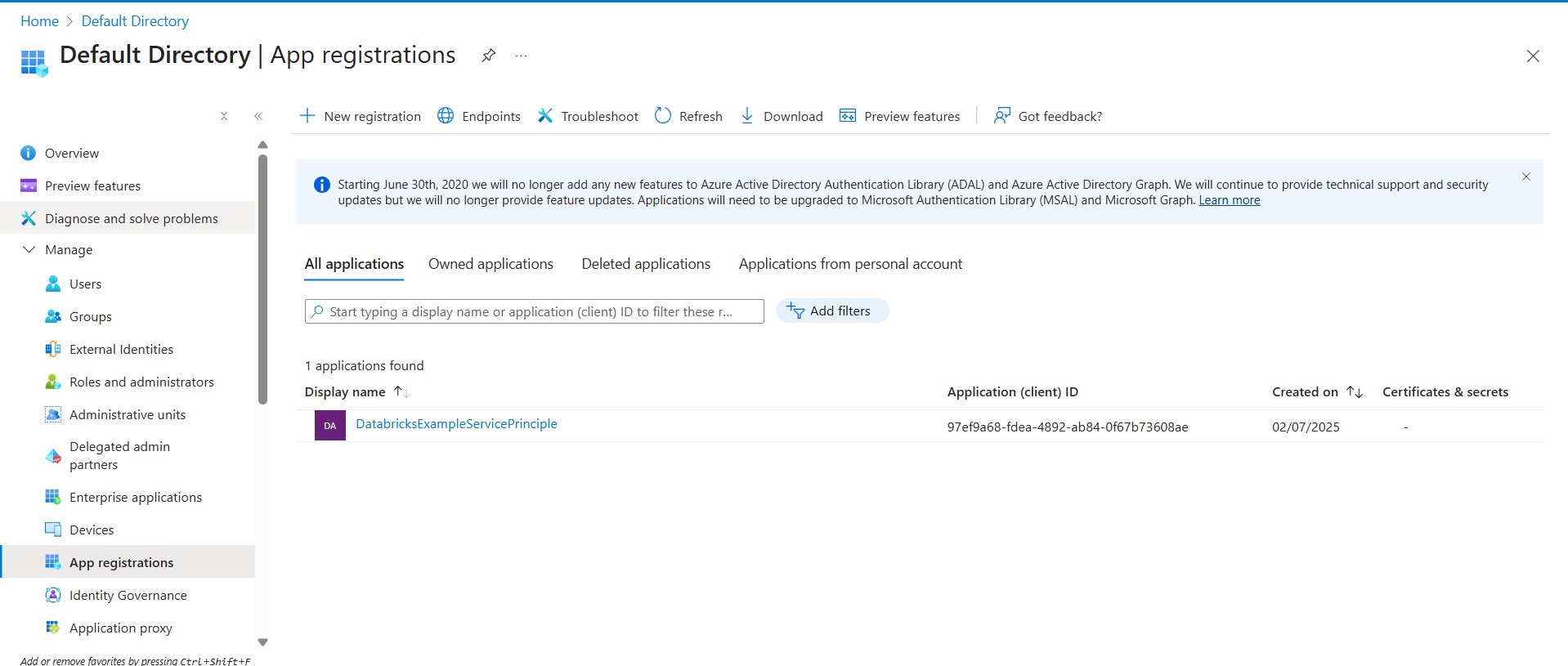
1. Creating an Entra ID (Service Principle layer)



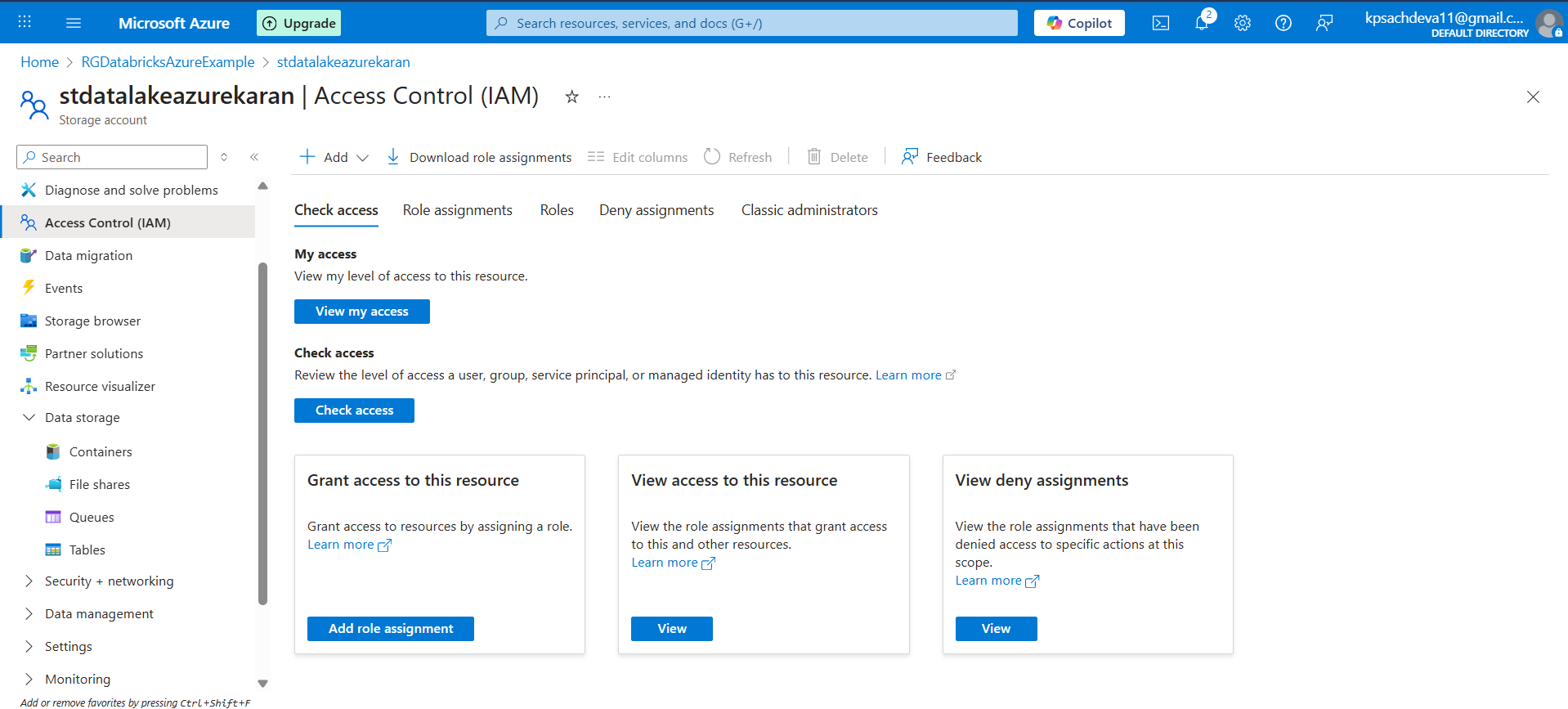
1. Creating a new App Registration in Entra ID



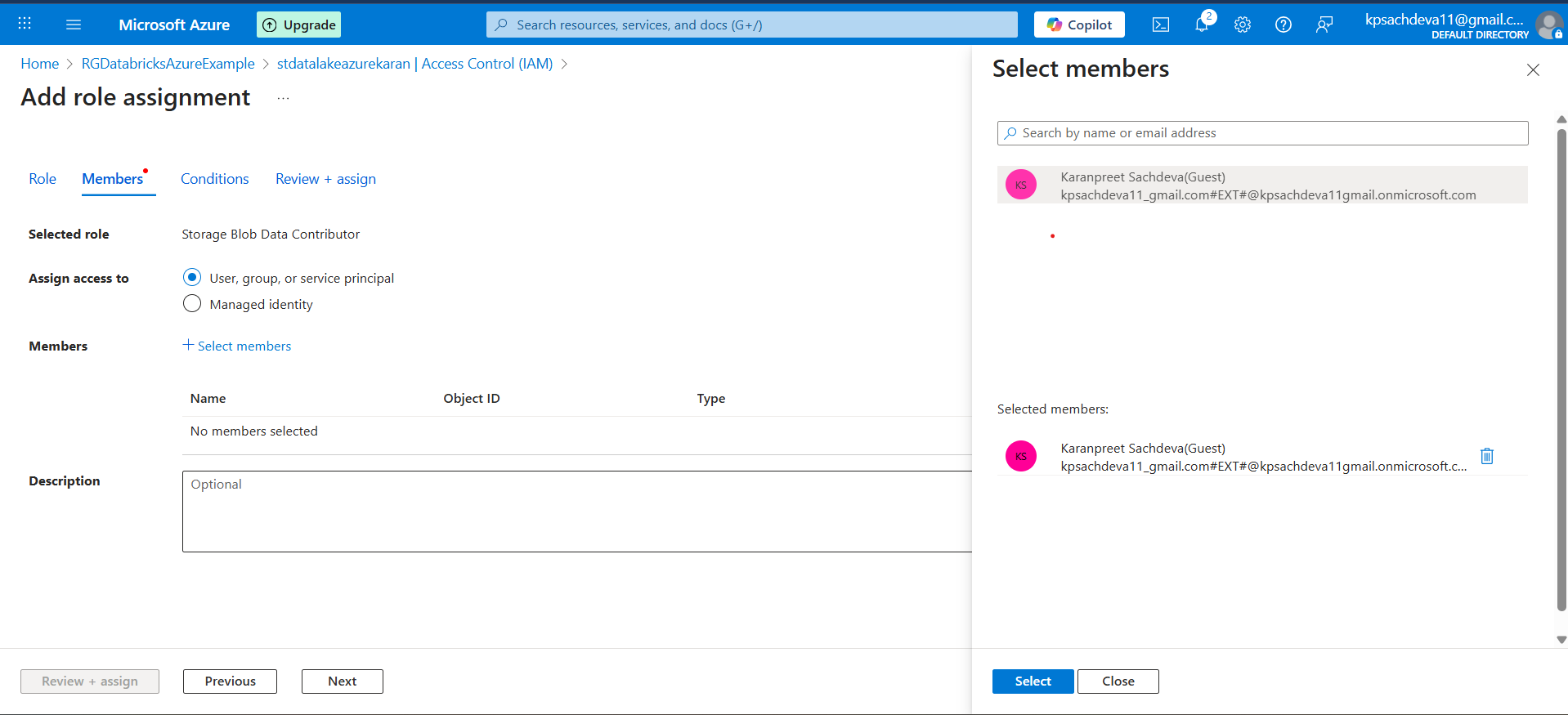
1. Creating a new Client Secret



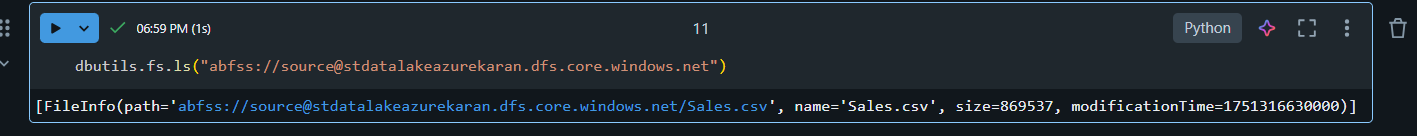
1. Creating an IAM control for Data request coming from Databricks



1. Creating a blob storage contributor role permissions



1. Connection with Datalake successful



Important concept of external delta tables and managed delta tables

1. Managed Delta tables in databricks

MetaStore

Unity metastore, hive metastore. Consists of all the definitions, basically metadata about the databases. In short, the meta data about the database tables will be stored here.

Cloud Storage (Default for Databricks), When we create a databricks resource a default storage account is created which stores the actual data for all the tables and DB

1. External Delta tables in databricks

Cloud Storage (Default from databricks)

Cloud Storage (Our Own)

Here we will store our data. This will be helpful in the case where when we will delete the table, table will be deleted from default cloud storage and meta store but not from our own cloud storage.

MetaStore

This will also consist of all the definitions and meta data about the tables and databases, and will not create the actual data.

## CRUD operations in delta tables

1. Creating the tables:
2. Creating the table

CREATE TABLE salesDB.managed\_table (

id INT,

name STRING,

marks INT

)

using DELTa

1. Inserting data into table

INSERT INTO salesDB.managed\_table

VALUES (1, 'John', 95),

(2, 'Bob', 85),

(3, 'Mary', 90),

(4, 'Mike', 80),

(5, 'Sarah', 75),

(6, 'Jane', 95),

(7, 'David', 85),

(8, 'Emily', 90),

(9, 'Kevin', 80),

(10, 'Olivia', 75),

(11, 'Daniel', 95),

(12, 'Sophia', 85)

\* The data that we are storing in the DB can be found in the metastore, using the catalog in azure databricks.

1. Creating an external table

CREATE TABLE salesDB.externalTable (

id INT,

name STRING,

marks INT

) USING DELTA

LOCATION 'abfss://destination@stdatalakeazurekaran.dfs.core.windows.net/salesDB/external/'