Lab: Integrate Databricks Unity Catalog with OneLake
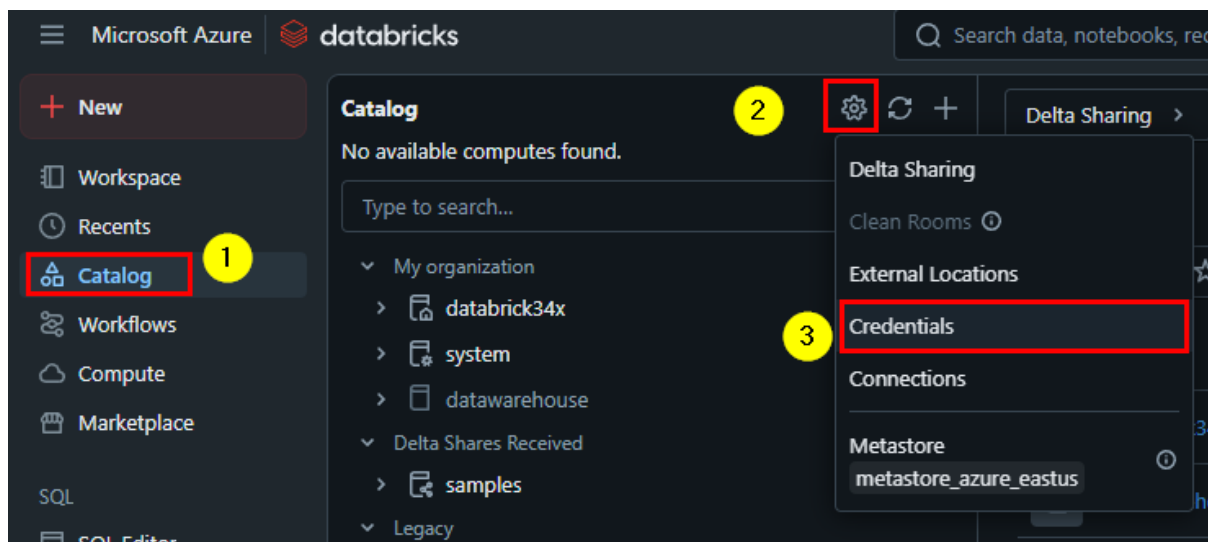
Objective:

Understand the concept of integrating OneLake with the Unity Catalog.

Tasks:

1. Create a unity catalog in Databricks Workspace
2. Create external credentials and location
3. Set up your Cloud storage connection
4. Create a OneLake Shortcut

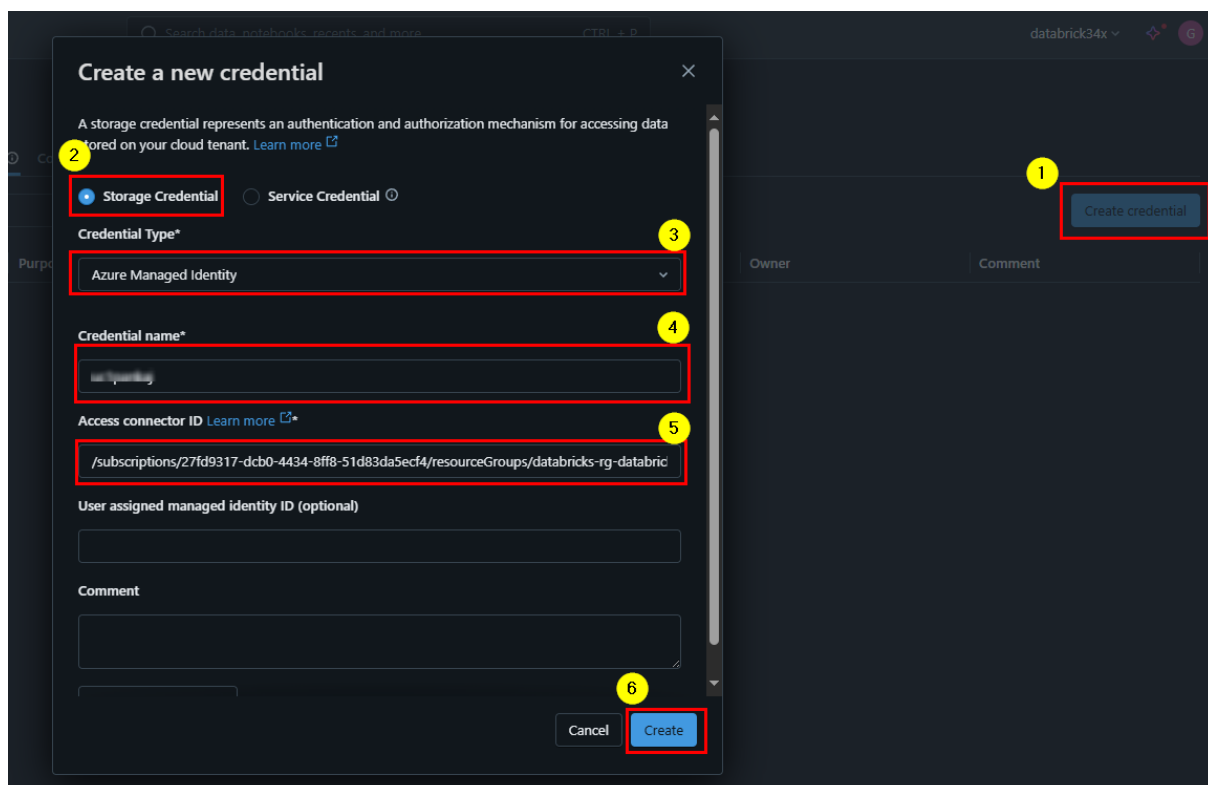## Task 1: Create credentials and external location

1. Open the Databricks workspace (https://adb-3426738885164031.11.azuredatabricks.net/) in a browser window.

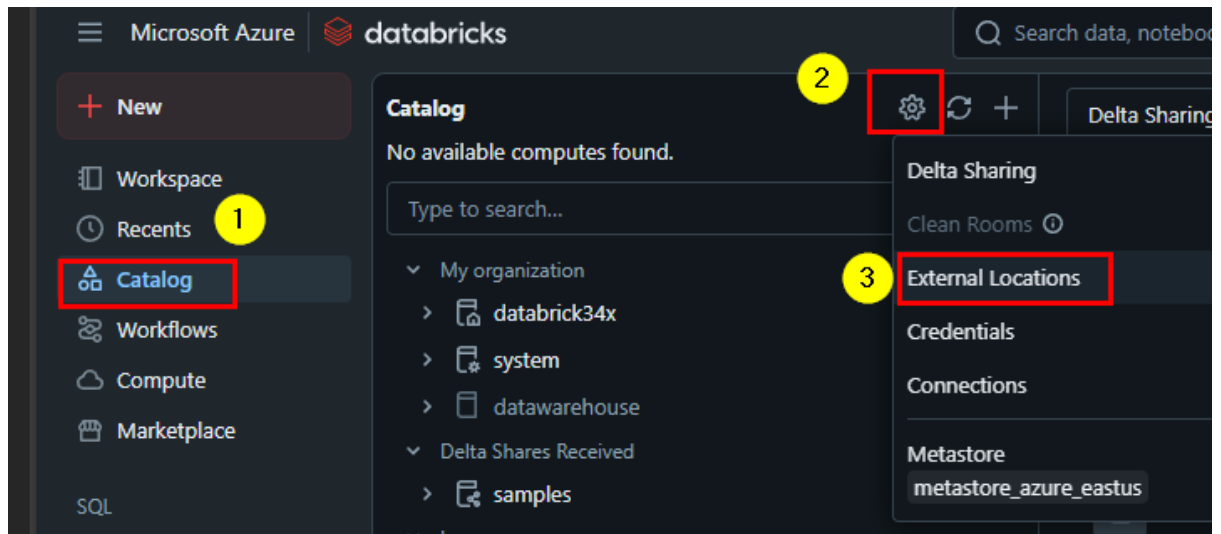2. Navigate to catalog click on the gear icon and select Credentials

3. Click **Create credential** and provide the following details as shown:

   Access connector ID:

   /subscriptions/27fd9317-dcb0-4434-8ff8-51d83da5ecf4/resourceGroups/RG-Genpact/providers/Microsoft.Databricks/accessConnectors/databricks-access-connector

4. Create an external location

5. Create a container in Azure Data Lake storage

Storage url: abfss://<replace-with-your-containerName>@storage34x.dfs.core.windows.net/
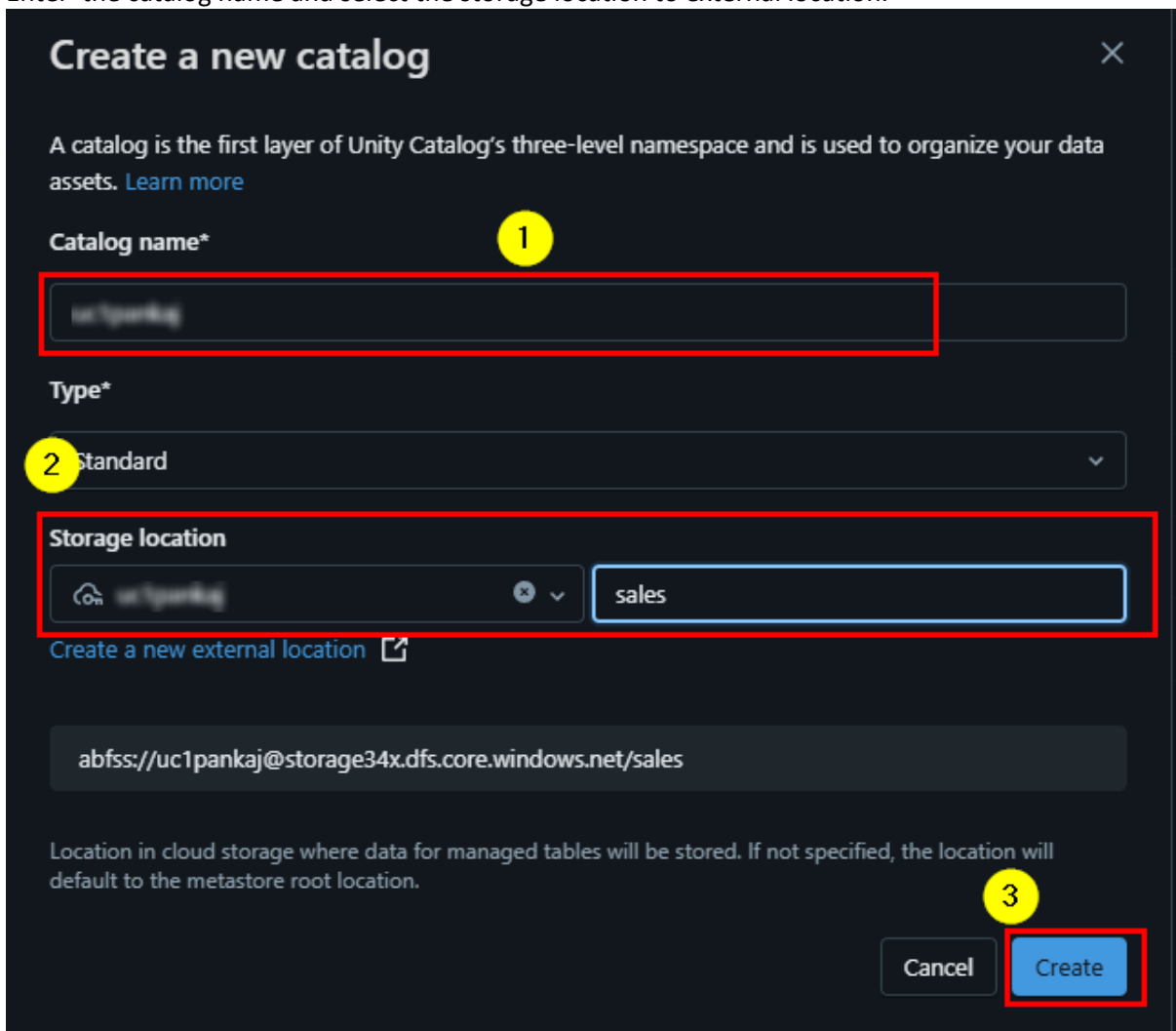
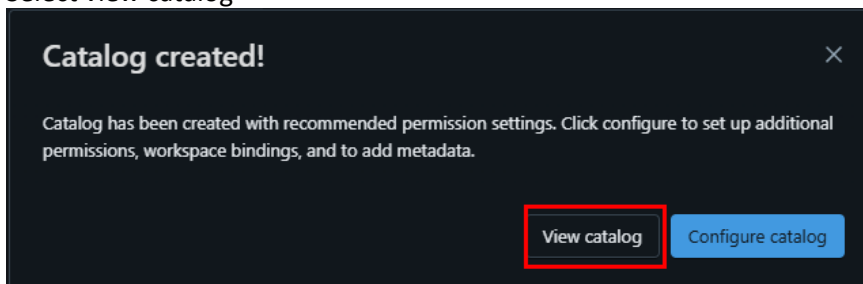## Task 2: Create a unity catalog in Databricks Workspace

1. Select Catalog from the left vertical menu, select Catalogs and click Create Catalog.
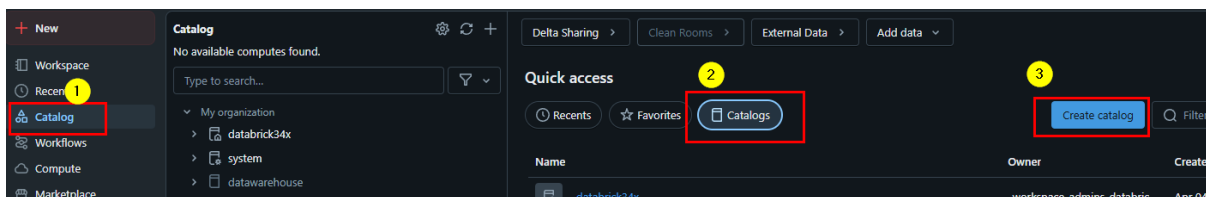
2. Enter the catalog name and select the storage location to external location.

3. Select view catalog

4. Create a new notebook in Databricks inside the workspace
5. In the first cell enter the following code and create an empty cell and run it, if no cluster is created then create a new one.

```
use catalog uc1pankaj;
create schema if not exists sales;
use schema sales;
create table if not exists department
(
  deptcode int,
  deptname string,
  location string
);

INSERT INTO department VALUES
  (10, 'FINANCE', 'EDINBURGH'),
  (20, 'SOFTWARE', 'PADDINGTON');

select * from department;
```

6. Create a new code cell and run the following code to create an external table. Replace your container name:

```
create external table trips_external
using delta
location 'abfss://<your-container-name>@storage34x.dfs.core.windows.net/trip'
as select * from samples.nyctaxi.trips;
```

7. Terminate the compute resource.

Task 3: Set up your Cloud storage connection in Fabric

1. Switch back to the Microsoft Fabric portal and from setting open connections.
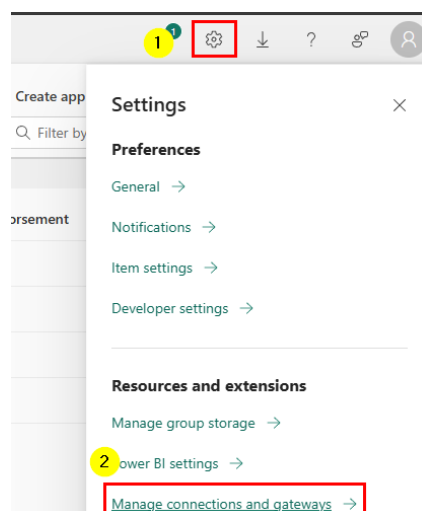
2. Enter the Connection name, type, storage URL container name, authentication type: key, and Account key.

3. Open the connection and copy the connection ID:

Task 4: Create a OneLake Shortcut

1. **Import the sync notebook** to your Fabric workspace. Download [This notebook](#)

2. **Configure the parameters** in the first cell of the notebook to integrate Unity Catalog tables. The Databricks API, authenticated through PAT token, is utilized for exporting Unity Catalog tables. The following snippet is used to configure the source (Unity Catalog) and destination (OneLake) parameters. Ensure to replace them with your own values.

   ```
   # Databricks workspace

   dbx_workspace = "<databricks_workspace_url>"

   dbx_token = "<pat_token>"

   # Unity Catalog

   dbx_uc_catalog = "catalog1"

   dbx_uc_schemas = '["sales"]'


   # Fabric

   fab_workspace_id = "<workspace_id>"

   fab_lakehouse_id = "<lakehouse_id>"

   fab_shortcut_connection_id = "<connection_id>"

   # If True, UC table renames and deletes will be considered

   fab_consider_dbx_uc_table_changes = True
   ```

3. Run all cells of the notebook to start synchronizing Unity Catalog Delta tables to OneLake using shortcuts. Once notebook is completed, shortcuts to Unity Catalog Delta tables are available in the lakehouse, SQL analytics endpoint, and semantic model.