Describe how Azure Databricks manages Delta Lake

Note: *In this reading you can see the steps involved in the process of using Azure Databricks to manage Data Lake.*

Unit notebook

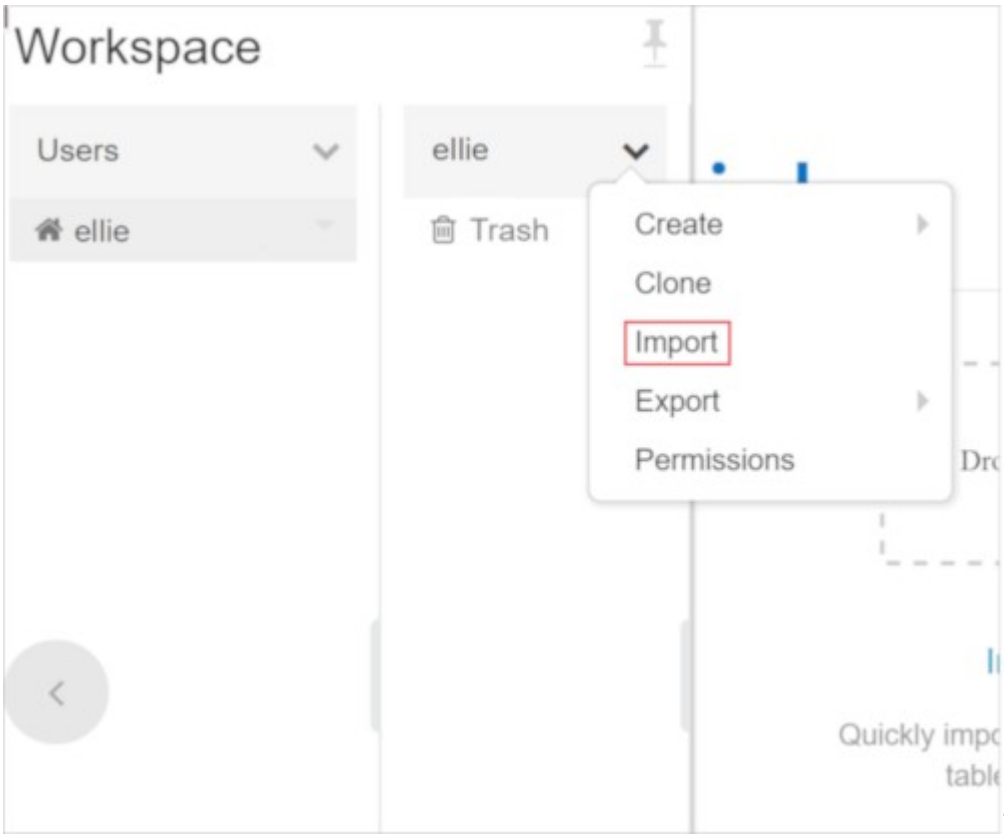
In this unit, you need to complete the exercises within two Databricks Notebooks. To begin, you must first import the notebooks by cloning a Databricks archive.

Clone the Databricks archive

1. If you do not currently have your Azure Databricks workspace open then access the Azure portal, navigate to your deployed Azure Databricks workspace and select **Launch Workspace**.

2. In the left pane, select **Workspace** > **Users**, and then select your username (the entry with the house icon).

3. In the pane that appears, select the arrow next to your name, and select **Import**.



The menu option to import the archive.

4. In the **Import Notebooks** dialog box, select the URL bar and paste in the following address:

1

https://github.com/solliancenet/microsoft-learning-paths-databricks-notebooks/blob/master/data-engineering/DBC/08-Azure-Databricks-Security-Data-Protection.dbc?raw=true

5. Select **Import**.

6. Select the **09-Building-And-Querying-A-Delta-Lake** folder that appears.

Complete the following Notebook:

In your Azure Databricks workspace, open the **09-Building-And-Querying-A-Delta-Lake** folder that you imported within your user folder.

Open the **3.Managed-Delta-Lake** notebook. Make sure you attach your cluster to the notebook before following the instructions and running the cells within.

Within the notebook, you will discover Delta Lake's key features that allow for query optimization and garbage collection, resulting in improved performance.

After you've completed the notebook, return to this screen, and continue to the next step.