Tagging Dataplex Assets

Tools Used:

- Google Cloud Console
- Cloud Dataplex
- Data Catalog
- BigQuery

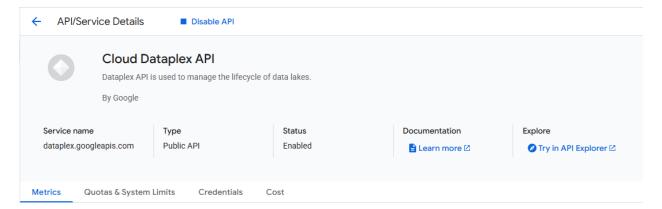
Objective

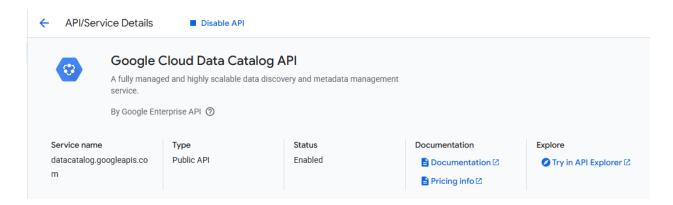
The objective of this lab was to demonstrate how to use Google Cloud's **Dataplex** and **Data Catalog** to:

- 1. Create a lake, zone, and asset.
- 2. Build and apply a **tag template** to a dataset.
- 3. Search for and identify assets using the applied tags.

Enable Dataplex and Data Catalog APIs

- 1. In the Google Cloud Console, enter **Cloud Dataplex API** in the top search bar.
- 2. Click on the result for **Cloud Dataplex API** under Marketplace.
- 3. If not already enabled, click Enable.
- 4. Repeat steps 1-3 for Google Cloud Data Catalog API.





Task 1. Create a lake, zone, and asset

In this task, you create a new Dataplex lake to store customer order information, add a curated zone to the lake, and then attach a pre-created BigQuery dataset as a new asset in the zone.

Create a lake

1. In the Google Cloud Console, in the Navigation menu (=), click View all products. Under Analytics, click Dataplex Universal Catalog.

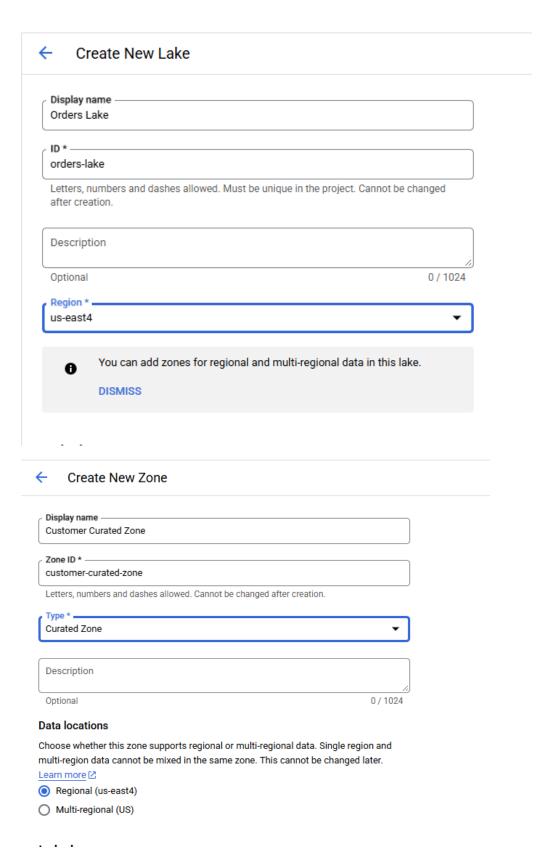
If prompted Welcome to the new Dataplex experience, click **Close**.

- 2. Under Manage lakes, click Manage.
- 3. Click Create.
- 4. Enter the required information to create a new lake:

Property	Value
Display Name	Orders Lake
ID	Leave the default value.
Region	

Leave the other default values.

5. Click Create.



Add a zone to the lake

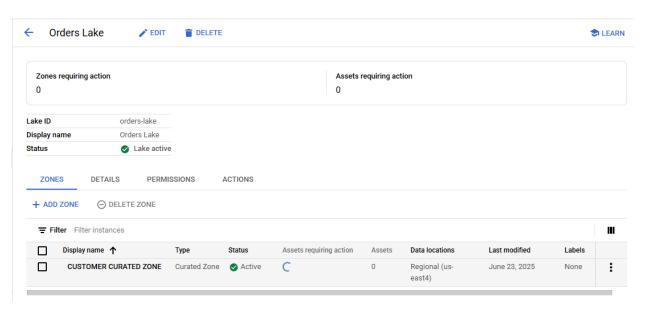
- 1. On the Manage tab, click on the name of your lake.
- 2. Click Add zone.
- 3. Enter the required information to create a new zone:

Property	Value
Display Name	Customer Curated Zone
ID	Leave the default value.
Туре	Curated zone
Data locations	Regional

Leave the other default values.

For example, the option for **Enable metadata discovery** under **Discovery settings** is enabled by default and allows authorized users to discover the data in the zone.

4. Click Create.



Attach an asset to a zone

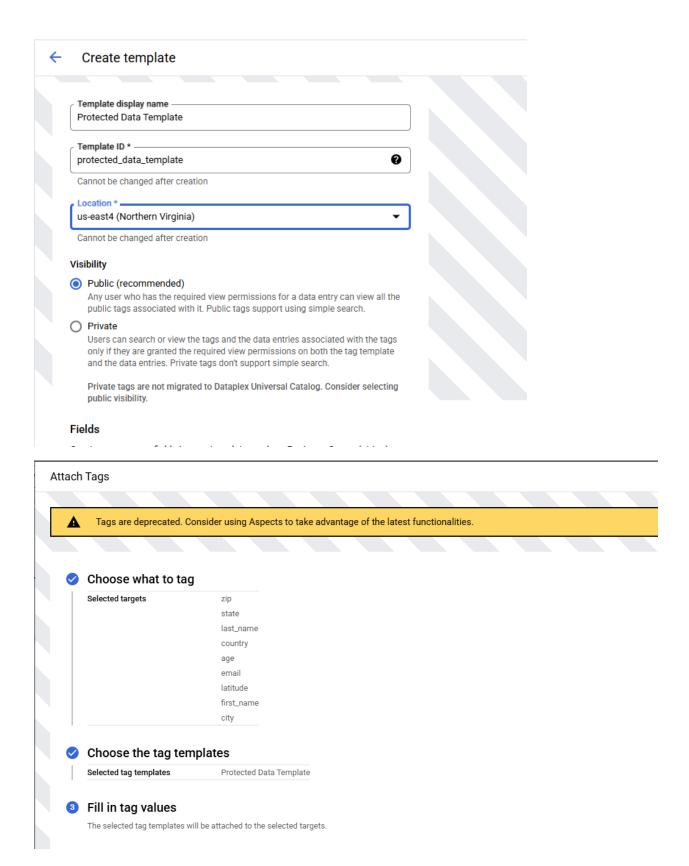
- 1. On the **Zones** tab, click on the name of your zone.
- 2. On the Assets tab, click Add assets.

- 3. Click Add an asset.
- 4. Enter the required information to attach a new asset:

Property	Value
Туре	BigQuery dataset
Display Name	Customer Details Dataset
ID	Leave the default value.
Dataset	customers

Leave the other default values.

- 5. Click Continue.
- 6. For **Discovery settings**, select **Inherit** to inherit the Discovery settings from the zone level, and then click **Continue**.
- 7. Click **Submit**.



Task 2. Create a tag template

To start tagging data, you first need to create one or more tag templates. A tag template can be a public or private tag template. When you create a new tag template, the option to create a public tag template is the default and recommended option. Users who have the required view permissions for a data asset can view all the public tags associated with it. This supports simple search for discovery while also adhering to data access controls already implemented on the underlying data.

In this task, you create a public tag template to label BigQuery table columns with a protected status. With a public tag template, users who have access to the underlying BigQuery table columns will be able to see the tags applied to the columns.

- 1. On the left menu, under Manage Metadata, click Catalog.
- 2. Click Create tag template (Deprecated).
- 3. When prompted Are you sure you want to continue?, click Proceed.

Note: Tag templates are being replaced by aspect types but are still functional at this time. You can proceed with the next steps to create a tag template.

This lab is scheduled to be updated to use aspect types, and these changes will be reflected soon.

4. Enter the required information to define the tag template:

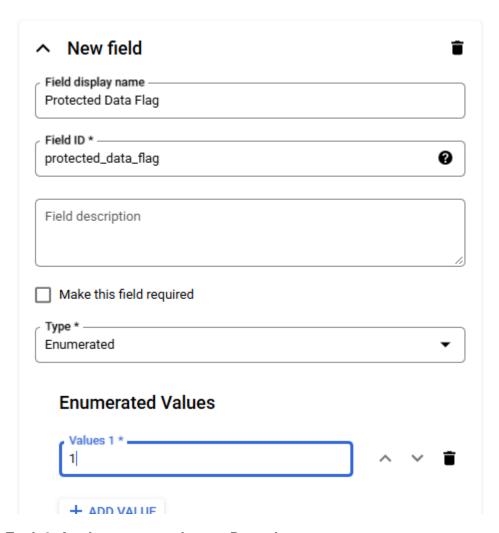
Property	Value
Template Display Name	Protected Data Template
Template ID	Leave the default value.
Location	
Visibility	Public

5. Click **Add field**, and enter the required information to add a new field to the template:

Property	Value

Field Display Name	Protected Data Flag
Field ID	Leave the default value.
Туре	Enumerated

- 6. For **Enumerated values > Values 1**, enter YES.
- 7. Click **Add value**, and for **Values 2**, enter NO.
- 8. Click Done.
- 9. Click Create.



Task 3. Apply a tag template to Dataplex assets

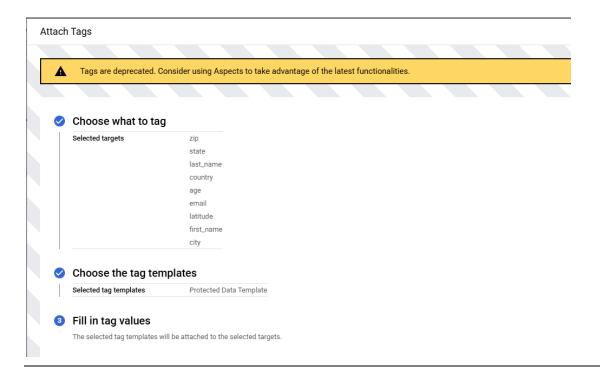
After you create a tag template, you can use it to attach tags to any number of desired data assets to which you have access.

In this task, you apply your previously created tag template to specific columns in the BigQuery table that you want to label with a protected data status.

- 1. On the left menu, under **Discover**, click **Search**.
- 2. For **Filters** > **Systems**, enable the checkbox for **Dataplex**.
- 3. Click on the **customer_details** table.

If you do not see the **customer_details** table, wait a few minutes and then refresh the page to allow the Dataplex asset list to be updated.

- 4. Click Attach tags.
- 5. For **Choose what to tag**, enable the checkboxes for the following columns:
 - a. zip
 - b. state
 - c. last_name
 - d. country
 - e. email
 - f. latitude
 - g. first_name
 - h. city
 - i. longitude
- 6. Click OK.
- 7. For Choose the tag templates, select Protected data template.
- 8. For **Protected data flag**, select **YES**.
- 9. Click Save.



Key Learnings

- **Dataplex** provides a centralized and intelligent way to manage data lakes and warehouses.
- Data Catalog enhances discoverability and governance with metadata tagging.
- Tag Templates enable reusable and consistent metadata annotation across datasets.
- **Public tags** help maintain transparency and easy access within permission boundaries.

Conclusion

This lab successfully demonstrated how Google Cloud's **Dataplex** and **Data Catalog** can be used together to **organize**, **tag**, and **search** data assets efficiently. By applying structured metadata through tag templates, teams can better manage **data governance**, **compliance**, and **discoverability** in large-scale data environments.