🧩 Get Started with Dataplex: Challenge Lab (ARC117)

Overview

This Google Cloud Skill Badge demonstrates proficiency in managing and organizing data using **Dataplex Universal Catalog**, a unified data fabric for metadata management, governance, and data discovery at scale. The challenge lab validates hands-on experience with creating lakes, zones, and assets, along with implementing metadata aspects to manage data efficiently.

Objectives

The lab involved completing the following tasks within a real Google Cloud environment:

1. **Create a Dataplex Lake with a Raw Zone**

   - Configured a lake named **Customer Engagements**.

   - Created a **Raw Event Data** zone to store unprocessed event information.

2. **Attach a Cloud Storage Asset**

   - Created a **Cloud Storage bucket** in the same region.

   - Attached it as a **regional asset** named *Raw Event Files* to the raw zone.

3. **Implement Metadata Governance with Aspect Types**

   - Created an **aspect type** named *Protected Raw Data Aspect* containing an enumerated field *Protected Raw Data Flag* with values `Y` and `N`.

   - Associated the aspect with the *Raw Event Data* zone to classify and govern sensitive data

Task 1 — Create a lake with a raw zone

1. Open Dataplex Universal Catalog

1. In Cloud Console click **Navigation menu → View all products**.

2. Under **Analytics**, click **Dataplex Universal Catalog**.

3. If a "Welcome to the new Dataplex Universal Catalog experience" pop-up appears, close it.

2. Create the lake

1. Under **Manage lakes**, click **Manage**.

2. Click **+ Create lake**.

3. Fill fields:

   - **Display Name:** `Customer Engagements`
   - **ID:** leave default (or match display name)
   - **Region:** `<REGION>`
   - Leave other options default.

4. Click **Create**. Wait (can take a few minutes) until the lake status becomes **Active**.

3. Add a raw zone to the lake

1. In the Dataplex UI, open the **Customer Engagements** lake (click its name).

2. Click **+ Add zone**.

3. Fill fields:

- **Display Name:** `Raw Event Data`
- **ID:** leave default
- **Type: Raw zone**
- **Data locations: Regional**
- **Region:** `<REGION>` (ensure it matches lake region)
- Leave other defaults (Discovery enabled is fine).

4. Click **Create**. Wait until the zone shows **Active**.

Task 2 — Create and attach a Cloud Storage bucket to the zone

1. Create a Cloud Storage bucket (via Dataplex UI)

1. In the Dataplex UI, go to **Zones** tab of the `Customer Engagements` lake, then click the `Raw Event Data` zone.
2. Open the **Assets** tab and click **+ ADD ASSET → + ADD AN ASSET**.
3. In **Type** choose **Storage bucket**.
4. **Display name:** `Raw Event Files` (or similar)
5. For **Bucket**, click **Browse**.
6. Click **+ Create new bucket (shopping basket icon)** to create a new bucket without leaving the UI.
7. For the new bucket:
   - **Bucket name:** use your **Project ID** (the lab suggested using Project ID) or another globally unique name.
   - **Location type: Region**
   - **Region:** `<REGION>`
   - Keep default retention / access options (or confirm "Public access will be prevented" then click Confirm).
8. Click **Create**, then **Select** the bucket you created.
9. For **Discovery settings**, choose **Inherit** to inherit zone discovery behavior.
10. Click **Continue**, then **Submit** to attach the bucket as the asset.
11. Wait until asset shows as attached (Active/Registered).

Alternative: Create bucket via Cloud Storage console

- Open **Storage → Browser → Create bucket**, enter name (Project ID), choose **Region** = `<REGION>`, create. Then attach it in Dataplex as above using Browse → Select.

Task 3 — Create an aspect type and add the aspect to the zone

1. Create the aspect type

1. In Dataplex UI, open the **Settings / Catalog** area (or look for **Aspect types** in the left menu). If not visible, open the **Universal Catalog** sections—aspect types are usually under Governance / Metadata / Aspect types.
2. Click **+ Create aspect type** (or **Create**).
3. Fill the aspect type details:
   - **Name:** `Protected Raw Data Aspect`
   - **Location / Region:** `<REGION>` (ensure region is same as lake/zone)
4. Add a field:
   - **Field name:** `Protected Raw Data Flag`
   - **Field type: Enumerated** (enum)

- **Values:** add two enumerations: `Y` and `N`
- (Optional) Add descriptions for the aspect and field.

5. Save the aspect type. Confirm creation completed successfully.

(If UI labels differ slightly: look for "Create Aspect Type", "New Aspect", or similar; the key is name, location, and a single enumerated field with `Y` and `N` .)

2. Add the aspect to the zone (or to the asset)

1. Navigate to the `Raw Event Data` zone in the Dataplex UI.

2. You can add aspects either at the zone level or to individual assets. Lab asked to **add the aspect to the zone**:
   - Open the zone details page and look for **Aspects** or **Metadata** tab.
   - Click **Add aspect**, choose the `Protected Raw Data Aspect` you created.
   - For the `Protected Raw Data Flag` value select either `Y` or `N` (choose accordingly; for lab you may set `N` unless instructed otherwise).
   - Submit/save.

3. Confirm the aspect appears in the zone's aspects list.