

Create a Secure Data Lake on Cloud Storage: Challenge Lab

✓ Task 1: Create a lake in Dataplex and add a zone

1 Create a Lake

1. Sign in as **User 2** in the lab.
2. Navigate to:

Navigation Menu > Dataplex

Click:

- Create Lake
- Fill:
- **Display Name:** Customer-Lake
- **ID:** Leave default (customer-lake-xxxxx)
- **Region:** Select the *region from the Lab Details panel.*
- Click:

Create

Cloud Storage

Create a bucket

Get Started

Pick a globally unique, permanent name. [Naming guidelines](#)

qwiklabs-gcp-03-249e0632f6b0-bucket

Tip: Don't include any sensitive information

Labels (optional)

Continue

Good to know

Location pricing

Item	Cost
us (multiple regions in United States)	\$0.026 per GB-month
With default replication	\$0.020 per GB written

Current configuration: Multi-region / Standard

Estimate your monthly cost

Cloud Storage

Create a bucket

Only locations allowed by the organization's resource location policy can be selected.

Location type

- Multi-region Highest availability across largest area
- Dual-region High availability and low latency across 2 regions
- Region** Lowest latency within a single region

us-central1 (Iowa)

Add cross-bucket replication via Storage Transfer Service
As data is added or changed, replicate it to another bucket, enabling you to store a copy that follows different bucket settings - e.g., region, storage class, etc. [Learn more](#)

Continue

Choose how to store your data

Cloud Storage

Buckets

Bucket details

qwiklabs-gcp-03-249e0632f6b0-bucket

Location	Storage class	Public access	Protection
us-central1 (Iowa)	Standard	Not public	Soft Delete

Objects Configuration Permissions Protection Lifecycle Observability New

Folder browser Buckets > qwiklabs-gcp-03-249e0632f6b0-bucket

Create folder Upload ...

Filter by name prefix only Filter Filter object

Name	Size	Type	Created	Storage
No rows to display				

Created bucket qwiklabs-gcp-03-249e0632f6b0-bucket X

Get started with Cloud Storage

- Getting bucket information
- Uploading objects
- Downloading objects
- Use cases for Cloud Storage
- Terraform samples

2 Add a Zone

1. Inside **Customer-Lake**, go to **Zones**.

2. Click:

3. Add Zone

- Fill:
- **Display Name:** Public-Zone
- **ID:** Leave default.
- **Zone Type:** Raw
- **Storage Type:** Regional

- Enable **Metadata Discovery** under **Discovery Settings**.
- Under **Labels**:
- Key: domain_type
- Value: source_data
- Click:

Create

Create New Lake

Display name: Customer-Lake

ID: customer-lake

Description (Optional):

Region: us-central1

You can add zones for regional and multi-regional data in this lake.

Lakes	CREATE	LEARN				
CUSTOMER-LAKE	0 Assets	Active Status	0 Resources requiring action	us-central1 (Iowa) Region	July 1, 2025 Last modified	None Labels
PUBLIC-LAKE	0 Assets	Active Status	0 Resources requiring action	us-central1 (Iowa) Region	July 1, 2025 Last modified	None Labels

Creation of lake (projects/qwiklabs-gcp-03-249e0632f6b0/locations/us-central1/lakes/customer-lake) finished.

✓ Task 2: Create an entry group

1. Go to:

Navigation Menu > Data Catalog > Entry groups

Click:

- Create Entry Group
- Fill:
 - **Name:** Custom entry group
 - **Location:** Lab region
- Click:

The screenshot shows two views of the Google Cloud Platform Data Catalog interface.

Top View: Create New Zone

- Left Sidebar:** Discover, Search, Manage Metadata (Catalog, Glossaries), Manage Lakes (Secure, Process, Manage, Profile, Data Quality), Govern, Settings.
- Form Fields:**
 - Display name: Public-Zone
 - Zone ID *: public-zone
 - Type *: Raw Zone
 - Description: (Optional) [Placeholder: 0 / 1024]
- Right Panel:**
 - Zones in a lake**
 - Raw**: Store data in any format from external sources in Cloud Storage. Useful for data requiring further processing before it is ready for consumption.
 - Curated**: Structured data in Cloud Storage must conform to certain file formats and organized in a hive-compatible directory layout. Useful for data ready for consumption and analytics workloads. [Learn more about zone design recommendations](#)
 - Decide how tables are named in Dataproc Metastore and BigQuery**: When publishing discovered tables and filesets in Dataproc Metastore and

Bottom View: Labels

- Left Sidebar:** Discover, Search, Manage Metadata (Catalog, Glossaries), Manage Lakes (Secure, Process, Manage, Profile, Data Quality), Govern, Settings.
- Form Fields:**
 - Choose whether this zone supports regional or multi-regional data. Single region and multi-region data cannot be mixed in the same zone. This cannot be changed later. [Learn more](#)
 - Regional (us-central1) (radio button selected)
 - Multi-regional (US) (radio button)
- Labels Section:**
 - Labels are keyvalue pairs that allow you to group related objects together or with other Cloud Platform resources.
 - Key 1: domain_type, Value 1: source_data
 - + ADD LABEL
- Discovery settings:**
 - All assets will inherit these settings, unless configuration overrides exist at the asset level.
 - Scan files in this zone to detect and register table metadata.
 - Enable metadata discovery

Display name	Type	Status	Assets requiring action	Assets	Data locations	Last modified	Labels
PUBLIC-ZONE	Raw Zone	Active	0	0	Regional (us-central1)	July 1, 2025	domain_type:source_data

✓ Task 3: Attach an existing Cloud Storage bucket to the zone

1 Attach the Bucket

1. Stay in **Dataplex > Customer-Lake > Public-Zone > Assets.**
2. Click:
3. Add Asset

Fill:

- **Asset Type:** Storage Bucket
- **Display Name:** Customer Raw Data
- **ID:** Leave default.
- **Storage bucket:** Select:

[PROJECT-ID]-customer-bucket

Under **Advanced Settings:**

- Enable **Metadata Discovery** (ignore if already enabled).
- Enable **CSV Parsing Options:**
 - First row as header: Enable
 - Delimiter: ,
 - Encoding: UTF-8

Click:

Save

The screenshot shows the Google Cloud Data Catalog interface. On the left, there is a navigation sidebar with sections like Discover, Manage Metadata, Manage Lakes, and Govern. Under Manage Lakes, 'Manage' is selected. The main area shows a step-by-step process for adding assets to a zone:

- 1 Add assets: Add either Storage buckets or BigQuery datasets to this zone.
- 2 Advanced settings (optional): Change deletion policy and discovery settings for all assets.
- 3 Review assets.

The 'Add Assets' form is open, showing a 'New asset' entry. The 'Type' is set to 'Storage bucket', 'Display name' is 'Customer Raw Data', and 'ID' is 'customer-raw-data'. A note indicates that GCS buckets are required for access by Dataplex. The 'Description' field is optional, with a character limit of 1024.

Below this, the 'Public-Zone' page is shown. It displays the total assets (0) and assets requiring action (0). The zone details are: Zone ID: public-zone, Display name: Public-Zone, Type: Raw Zone, Status: Active. The 'ASSETS' tab is selected, showing a table with one row for 'Customer Raw Data' (Storage bucket, Status: Creating). A message at the bottom says 'Asset creation started'.

✓ Task 4: Create a tag template & attach to the asset

1 Create a Tag Template

1. Go to:

Navigation Menu > Data Catalog > Tag Templates

Click:

- Create Tag Template
- Fill:

- **Name:** Customer Data Tag Template
- **ID:** customer_data_tag_template
- **Location:** Default region.
- Under **Fields:**
- **Field 1:**
 - Name: Data Owner
 - Type: String
- **Field 2:**
 - Name: PII Data
 - Type: Enumerated
 - Add values:
 - Yes
 - No
- Click:

Create

Create template

Template display name Customer Data Tag Template

Template ID * customer_data_tag_template

Location * us-central1 (Iowa)

Visibility

Public (recommended) Any user who has the required view permissions for a data entry can view all the public tags associated with it. Public tags support using simple search.

Private Users can search or view the tags and the data entries associated with the tags only if they are granted the required view permissions on both the tag template and the data entries. Private tags don't support simple search.

Private tags are not migrated to DataPlex Internal Format. Consider selecting

Create template

public visibility.

Fields

Create one or more fields in your template, such as Business Owner (string), Project Number (double), or Contains PII (boolean)

New field

Field display name: Data Owner

Field ID *: data_owner

Field description:

Make this field required

Type *: String

DONE

Create template

Field description:

Make this field required

Type *: Enumerated

Enumerated Values

Values 1 *: Yes

Values 2 *: No

+ ADD VALUE

DONE

Entry group de...

DATA CATALOG DATAPLEX UNIVERSAL CATALOG NEW HIDE INFO PANEL

Custom entry group

ENTRIES DETAILS

List of entries in the entry group

+ CREATE FILESET EDIT FILESET

- DELETE FILESET

<input type="checkbox"/>	Name	System	Type	Description	Created
No entries to display. To manually add a GCS fileset, click create button above.					

Custom entry group

Edit or delete roles below, or select "Add principal" to grant new access.

Show inherited roles in table

Display roles inherited from the parent resources in the table below

Filter Enter property name or value

Role / Principal Inheritance

- Cloud Dataplex Service Agent (1)
- DataCatalog EntryGroup Creator (1)
- DataCatalog EntryGroup Owner (1)
- Editor (2)
- Owner (3)
- Viewer (2)

2 Attach the Tag Template to the Asset

1. Go to:

Navigation Menu > Data Catalog > Search

Search:

[PROJECT-ID]-customer-bucket

- (or filter by CLOUD STORAGE to find your bucket asset.)
- Click the bucket asset entry.
- Click:
- Add Tag

Choose:

Customer Data Tag Template

- Fill:
- **Data Owner:** (Enter your name here)
- **PII Data:** Yes
- Click:Save

 How to Search for Assets Using a Tag Template: Customer Data Tag Template

Go to: Navigation Menu > Data Catalog > Search

In the **Filter**:

- Go to **Tag Template**, select:

Customer Data Tag Template

You may also filter by **Tag Field Values** if needed.

The search will now return **all assets tagged using your Customer Data Tag Template**, allowing you to view all tagged entries efficiently.