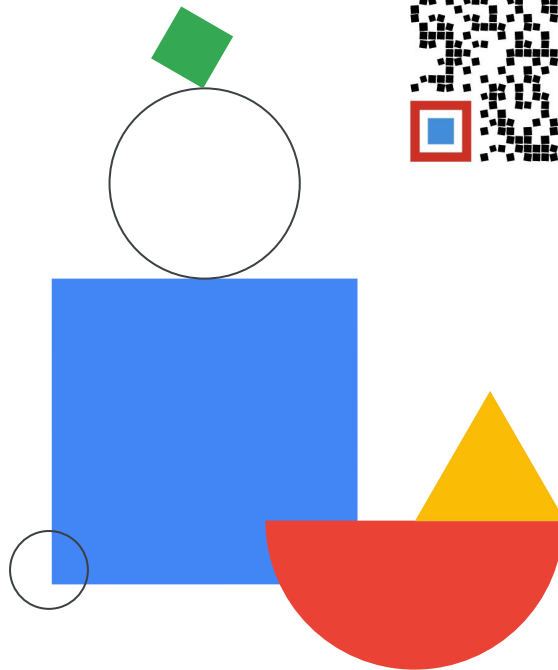


SCAN to REGISTER your Attendance



# BigQuery Demo

Trainer: Paggie Chen



# Agenda

- 01 **How to Create a Table**
  - Native table
  - From public dataset
  - External table (through biglake connection or not)

---
- 02 **Core Concepts**
  - Project, Dataset, Table
  - Partitioned & Clustered table
  - Table & View
  - Authorized view (column-level access control)
  - Row-level access control
  - Policy Tag
  - Data Lineage

---
- 03 **Build Pipelines**
  - Data Canvas
  - Dataform

---
- 04 **Augment with BigQuery ML**
  - Scenario 1: [Prebuilt API] Document AI with BQML
  - Scenario 2: [Vertex AI] Image Data Summarization with BQML
  - Scenario 3: [Native Model] Custom BQML Model

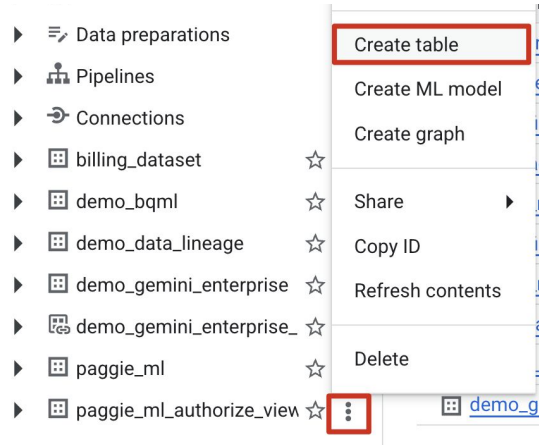
---
- 05 **Data Governance**
  - Data Profile, Data Quality Scan

# How to Create a Table

## Native table

From public dataset

External table (through a connection or not)



## Source

Create table from

Empty table

Google Cloud Storage

Upload

Drive

Google Bigtable

Amazon S3

Azure Blob Storage

Existing table/view

Table type

Native table

# How to Create a Table

Native table

From public dataset

External table (through a connection or not)



The image shows the Google Cloud Marketplace interface. At the top, there is a 'Marketplace' header with a shopping cart icon and a search bar containing 'Search Marketplace'. Below the header, the breadcrumb 'Marketplace > Data' is visible. A filter menu is open, showing 'Filter Type to filter' and a list of categories: AI Agent (1), Analytics (58), Data (84), Machine learning (6), and Generative AI (3). The 'Data' category is selected. To the right, it says '352 results'. Two data cards are displayed. The first card is for 'google-books-ngrams-2020', a BigQuery Public Data set, with a description: 'Google Books word counts found in books from 1500-2012'. The second card is for 'Solid Date', a comprehensive, materialized calendar and date dimension table, with a description: 'solid\_date is a comprehensive, materialized calendar and date dimension table, enumerating every discrete date value over 500 years from...'. Both cards are labeled 'Data' at the bottom.

## How to Create a Table

Native table

From public dataset

**External table**  
**(through a connection or not)**

Table type

External table

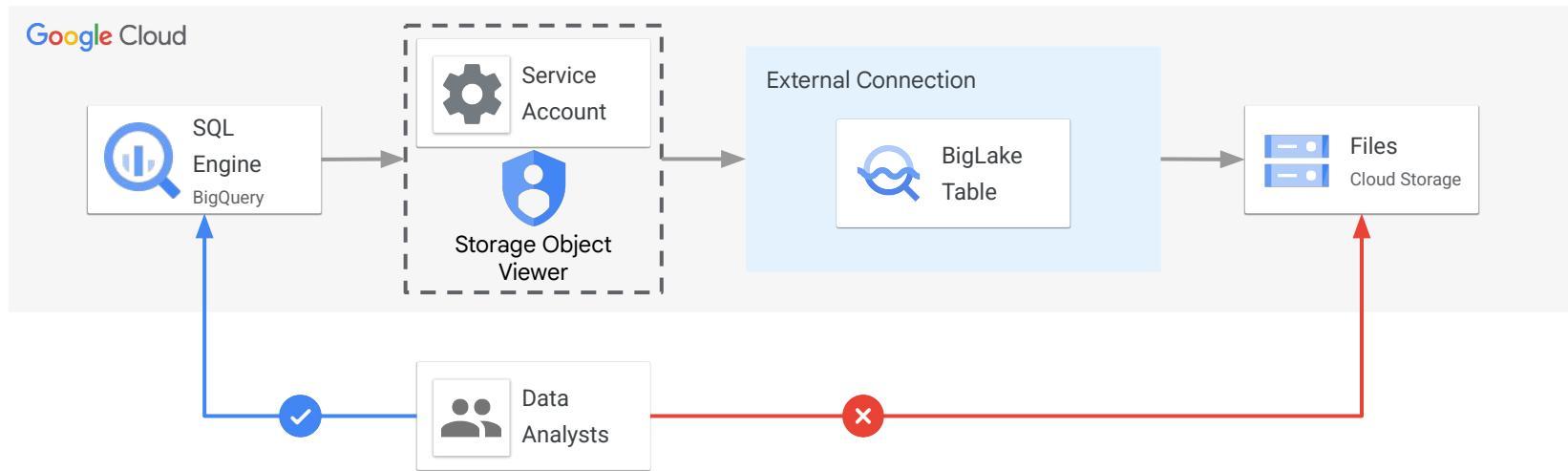


Regional / dual region GCS buckets are recommended for External table.



Create a Lakehouse table using a Cloud resource connection

# Access delegation enforces fine-grained security



## Access delegation

- Decouples access to BigLake table from underlying data store, fine-grained security at table level.
- An external connection associated with a service account is used to connect to the data store.

## Caution

- Users should not have access to the objects directly in Cloud Storage.
- Direct file access lets users bypass governance policies.

# Non-BigLake Table (Creation)

## Create table

### Source

Create table from  
Google Cloud Storage

Select file from GCS bucket or use a [URI pattern](#) \*  
 paggie-bucket/test.csv

File format  
CSV

Source Data Partitioning

### Destination

Project \*  
paggie-svc-project

Dataset \*  
paggie\_ml

Table \*  
demo\_gcs\_external\_table

Maximum name size is 1,024 UTF-8 bytes. Unicode letters, marks, numbers, connectors, dashes, and spaces are allowed.

Table type  
External table

**i** Regional / dual region GCS buckets are recommended for External table.

Create a BigLake table using a Cloud resource connection

### Schema

Auto detect

# BigLake Table (Creation)

## Create table

### Source

Create table from  
Google Cloud Storage

Select file from GCS bucket or use a [URI pattern](#) \*  
 paggie-bucket/test.csv

File format  
CSV

Source Data Partitioning

### Destination

Project \*  
paggie-svc-project

Dataset \*  
paggie\_ml

Table \*  
demo\_gcs\_biglake\_table

Maximum name size is 1,024 UTF-8 bytes. Unicode letters, marks, numbers, connectors, dashes, and spaces are allowed.

Table type  
External table

**i** Regional / dual region GCS buckets are recommended for External table.

Create a BigLake table using a Cloud resource connection

Connection ID \*  
asia-east1.demo\_biglake\_connection

# BigLake Table (Creation)

Create a BigLake table using a Cloud resource

Connection ID \*  
asia-east1.demo\_biglake\_connection

## Schema

Create table

Cancel

You are missing permissions and may need to talk to your administrator. Original error message: Failed to create table:  
Access Denied: BigQuery BigQuery: Permission denied while globbing file pattern. bqcx-304589458397-ba4e@gcp-sa-bigquery-condel.iam.gserviceaccount.com does not have storage.objects.get access to the Google Cloud Storage object. Permission 'storage.objects.get' denied on resource (or it may not exist). Please make sure gs://paggie-bucket/test.csv is accessible via appropriate IAM roles, e.g. Storage Object Viewer or Storage Object Creator.

Grant BQ Service Account  
access in GCS

Grant access to "paggie-bucket"

Grant principals access to this resource and add roles to specify what actions the principals can take. Optionally, add conditions to grant access to principals only when a specific criteria is met. [Learn more about IAM conditions](#)

Resource

 paggie-bucket

Add principals

Principals are users, groups, domains, or service accounts. [Learn more about principals in IAM](#)

New principals \*

bqcx-304589458397-ba4e@gcp-sa-bigquery-condel.iam.gserviceaccount.com

Assign roles

Roles are composed of sets of permissions and determine what the principal can do with this resource. [Learn more](#)

Role \*  
Storage Object Viewer

IAM condition (optional) ⓘ  
IAM conditions disabled

Grants access to view objects and their metadata, excluding ACLs. Can also list the objects in a bucket.

+ Add another role

Save

Cancel

# Non-BigLake Table

paggie-svc-project / Datasets / paggie\_ml / Tables / demo\_gcs\_external\_table

demo\_gcs\_external\_table [Query](#) [Open in](#) [Share](#) [Delete](#)

[Schema](#) [Details](#) [Insights](#) [Lineage](#) [Data Profile](#) [Data Quality](#)

Generate table and column descriptions with Gemini. Ensure a profile scan to ground the

Filter Enter property name or value

<input type="checkbox"/>	Field name	Type	Mode	Description	Key	Collation	Default
<input type="checkbox"/>	first_name	STRING	NULLABLE	-	-	-	-
<input type="checkbox"/>	last_name	STRING	NULLABLE	-	-	-	-
<input type="checkbox"/>	gender	STRING	NULLABLE	-	-	-	-
<input type="checkbox"/>	age	INTEGER	NULLABLE	-	-	-	-

Edit schema

# BigLake Table

paggie-svc-project / Datasets / paggie\_ml / Tables / demo\_gcs\_biglake\_table

demo\_gcs\_biglake\_table BigLake [Query](#) [Open in](#) [+](#)

[Schema](#) [Details](#) [Insights](#) [Lineage](#) [Data Profile](#) [Data Quality](#)

Generate table and column descriptions with Gemini. Ensure a profile scan

Filter Enter property name or value

<input type="checkbox"/>	Field name	Type	Mode	Description	Key
<input type="checkbox"/>	first_name	STRING	NULLABLE	-	-
<input type="checkbox"/>	last_name	STRING	NULLABLE	-	-
<input type="checkbox"/>	gender	STRING	NULLABLE	-	-
<input type="checkbox"/>	age	INTEGER	NULLABLE	-	-

Edit schema

[View row access policies](#)

## Core Concepts

### Project, Dataset, Table

Partitioned & Clustered table

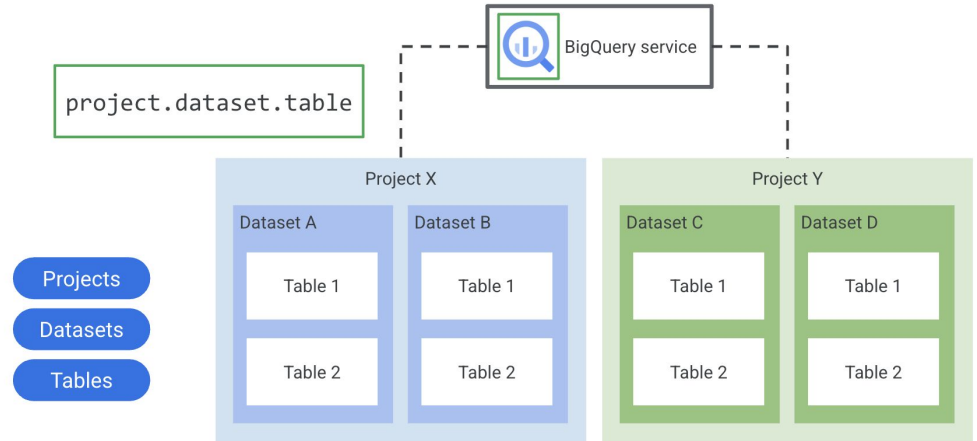
Table & View

Authorized view (column-level access control)

Row-level access control

Policy Tag

Data Lineage



# Core Concepts

Project, Dataset, Table

Partitioned & Clustered table

Table & View

Authorized view (column-level access control)

Row-level access control

Policy Tag

Data Lineage

## Clustered and Partitioned Tables

Orders table  
Not Clustered; Not partitioned

Order_Date	Country	Status
2022-08-02	US	Shipped
2022-08-04	JP	Shipped
2022-08-05	UK	Canceled
2022-08-06	KE	Shipped
2022-08-02	KE	Canceled
2022-08-05	US	Processing
2022-08-04	JP	Processing
2022-08-04	KE	Shipped
2022-08-06	UK	Canceled
2022-08-02	UK	Processing
2022-08-05	JP	Canceled
2022-08-06	UK	Processing
2022-08-05	US	Shipped
2022-08-06	JP	Processing
2022-08-02	KE	Shipped
2022-08-04	US	Shipped

Orders table  
Clustered by Country; Not partitioned

Order_Date	Country	Status
2022-08-04	JP	Shipped
2022-08-04	JP	Processing
2022-08-05	JP	Canceled
2022-08-06	JP	Processing
2022-08-06	KE	Shipped
2022-08-02	KE	Canceled
2022-08-04	KE	Shipped
2022-08-02	KE	Shipped
2022-08-05	UK	Processing
2022-08-06	UK	Canceled
2022-08-02	UK	Canceled
2022-08-06	UK	Processing
2022-08-02	US	Shipped
2022-08-05	US	Processing
2022-08-05	US	Shipped
2022-08-04	US	Shipped

Orders table  
Clustered by Country; Partitioned by Order\_Date (Daily)

	Order_Date	Country	Status
Partition: 2022-08-02	2022-08-02	KE	Shipped
	2022-08-02	KE	Canceled
Clusters: Country	2022-08-02	UK	Processing
	2022-08-02	US	Shipped

	Order_Date	Country	Status
Partition: 2022-08-04	2022-08-04	JP	Shipped
	2022-08-04	JP	Processing
Cluster: Country	2022-08-04	KE	Shipped
	2022-08-04	US	Shipped

	Order_Date	Country	Status
Partition: 2022-08-05	2022-08-05	JP	Canceled
	2022-08-05	UK	Canceled
Cluster: Country	2022-08-05	US	Shipped
	2022-08-05	US	Processing

	Order_Date	Country	Status
Partition: 2022-08-06	2022-08-06	JP	Processing
	2022-08-06	KE	Shipped
Cluster: Country	2022-08-06	UK	Canceled
	2022-08-06	UK	Processing

# Core Concepts

Project, Dataset, Table

Partitioned & Clustered table

Table & View

Authorized view (column-level access control)

Row-level access control

Policy Tag

Data Lineage

## Clustered Tables

**Table 1**  
Not clustered

Order_Date	Country	Status
2022-08-02	US	Shipped
2022-08-04	JP	Shipped
2022-08-05	UK	Canceled
2022-08-06	KE	Shipped
2022-08-02	KE	Canceled
2022-08-05	US	Processing
2022-08-04	JP	Processing
2022-08-04	KE	Shipped
2022-08-06	UK	Canceled
2022-08-02	UK	Processing
2022-08-05	JP	Canceled
2022-08-06	UK	Processing
2022-08-05	US	Shipped
2022-08-06	JP	Processing
2022-08-02	KE	Shipped
2022-08-04	US	Shipped

**Table 2**  
Clustered by country

Order_Date	Country	Status
2022-08-04	JP	Shipped
2022-08-04	JP	Processing
2022-08-05	JP	Canceled
2022-08-06	JP	Processing
2022-08-06	KE	Shipped
2022-08-02	KE	Canceled
2022-08-04	KE	Shipped
2022-08-02	KE	Shipped
2022-08-05	UK	Canceled
2022-08-06	UK	Canceled
2022-08-02	UK	Processing
2022-08-06	UK	Processing
2022-08-02	US	Shipped
2022-08-05	US	Processing
2022-08-05	US	Shipped
2022-08-04	US	Shipped

**Table 3**  
Clustered by country and status

Order_Date	Country	Status
2022-08-05	JP	Canceled
2022-08-04	JP	Processing
2022-08-06	JP	Processing
2022-08-04	JP	Shipped
2022-08-02	KE	Canceled
2022-08-06	KE	Shipped
2022-08-04	KE	Shipped
2022-08-02	KE	Shipped
2022-08-05	UK	Canceled
2022-08-06	UK	Canceled
2022-08-02	UK	Processing
2022-08-06	UK	Processing
2022-08-05	US	Processing
2022-08-02	US	Shipped
2022-08-05	US	Shipped
2022-08-04	US	Shipped

## Optimizing Queries on Clustered Tables

**Orders table**  
Clustered by Order\_Date, Country and Status

Order_Date	Country	Status
2022-08-02	KE	Canceled
2022-08-02	KE	Shipped
2022-08-02	UK	Processing
2022-08-02	US	Shipped
2022-08-04	JP	Processing
2022-08-04	JP	Shipped
2022-08-04	KE	Shipped
2022-08-04	US	Shipped
2022-08-05	JP	Canceled
2022-08-05	UK	Canceled
2022-08-05	US	Processing
2022-08-05	US	Shipped
2022-08-06	JP	Processing
2022-08-06	KE	Shipped
2022-08-06	UK	Canceled
2022-08-06	UK	Processing

**Example 1: Query filters on Order\_Date and Country**  
✓ - Query optimized for clustering

Order_Date	Country	Status

**Example 2: Query filters on Country and Status**  
✗ - Not optimized for clustering

Order_Date	Country	Status

## Core Concepts

Project, Dataset, Table

Partitioned & Clustered table

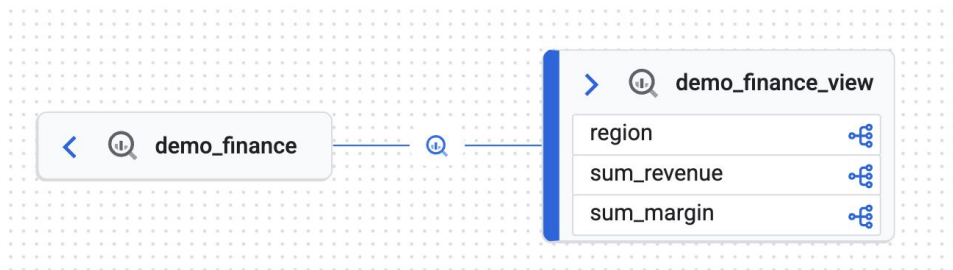
### Table & View

Authorized view (column-level access control)

Row-level access control

Policy Tag

Data Lineage



paggie-svc-project / Datasets / paggie\_ml\_authorize\_view / Tables / demo\_finance\_view

demo\_finance\_view Query Chat Open in Share Copy Delete Refresh

Schema Details Table Explorer Preview Insights Lineage Data Profile Data Quality

### View info

Edit Details

View ID	paggie-svc-project.paggie_ml_authorize_view.demo_finance_view
Created	Mar 31, 2026, 12:51:51 AM UTC+3
Last modified	Apr 19, 2026, 10:47:17 PM UTC+3
View expiration	NEVER
Use Legacy SQL	false
Description	
Labels	
Primary key(s)	
Tags	

**A view is a virtual table defined by a SQL query. Query results contain only the data from the tables and fields specified in the query that defines the view.**

### Storage info

Logical views are virtual and provide a reusable reference to a set of data, but don't physically store any data. [Learn more](#)

### Query

Edit Query

```
1 SELECT region, sum(revenue) AS sum_revenue, sum(margin) AS sum_margin
2 FROM `paggie-svc-project.paggie_ml.demo_finance`
3 GROUP BY region
```

## Core Concepts

Project, Dataset, Table

Partitioned & Clustered table

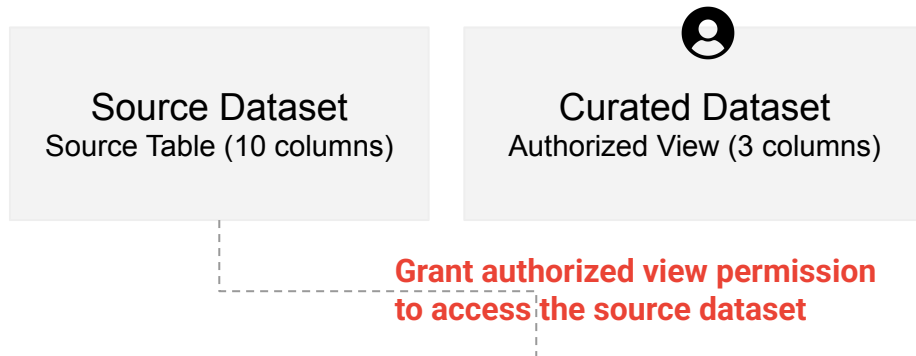
Table & View

**Authorized view (column-level access control)**

Row-level access control

Policy Tag

Data Lineage



### Authorized views

Authorize a view to access the data in this dataset. An authorized view lets you restrict the visibility of the data in this dataset, while still allowing users and groups to process the data by using the authorized view.

### Currently Authorized Views

Project ID	Dataset ID	Table ID	
paggie-svc-project	paggie_ml_authorize_view	demo_finance_view	

### Authorize view

Authorized View \*

Add authorization

## Core Concepts

Project, Dataset, Table

Partitioned & Clustered table

Table & View

Authorized view (column-level access control)

**Row-level access control**

Policy Tag

Data Lineage

### Row access policies

Limit access for a group of users to a subset of table rows. To add, edit, or remove row access policies, use DDL statements. [Learn more](#)

 Filter Filter row access policies


Name ↑	Filter expression	Created	Last modified	
service_filter	workload_type = 'Container'	Mar 31, 2026, 1:07:19 AM	Mar 31, 2026, 1:07:19 AM	<a href="#">View</a>

### Permissions for "service\_filter"

 Filter Enter property name or value

Role / Principal ↑

▼ BigQuery Filtered Data Viewer (1)

 paggiechen8866@gmail.com

## Core Concepts

Project, Dataset, Table

Partitioned & Clustered table

Table & View

Authorized view (column-level access control)

**Row-level access control**

Policy Tag

Data Lineage

paggie-svc-project / Datasets / paggie\_ml / Tables / workload

☆ workload 🔍 Query 🗨 Chat 📄 Open in ▾ 👤 Share ▾

Schema Details Preview Table Explorer Preview Insights

❗ Table preview is not supported for tables using row-level security.

```
1 SELECT * FROM `paggie-svc-project.paggie_ml.workload` LIMIT 1000
```

✅ Query completed

**Only workload\_type='Container' is shown in the query result**

Query results 🗨 Chat 📄

Job information Results Visualization JSON Execution details

Row	workload_id	workload_type	app_name
1	4	Container	Spotify
2	1	Container	UberEats

## Core Concepts

Project, Dataset, Table

Partitioned & Clustered table

Table & View

Authorized view

Row-level access control

Policy Tag

Data Lineage

The screenshot shows the BigQuery interface for the 'demo\_finance' table. The breadcrumb path is 'paggie-svc-project / Datasets / paggie\_ml / Tables / demo\_finance'. The table is in 'Preview' mode. The table schema is as follows:

Field name	Type	Mode	Description	Key	Collation	Default Value	Policy Tags
customer	STRING	NULLABLE	-	-	-	-	pii : confidential
sales	STRING	NULLABLE	-	-	-	-	pii : confidential
margin	FLOAT	NULLABLE	-	-	-	-	

A tooltip message states: 'You have restricted access to the data in this column due to attached policy tags.' An arrow points from the 'pii : confidential' tags in the table to the 'Masking Rules' section below.

## Masking Rules

Choose a masking rule, then select principals that the rule would apply to.

The screenshot shows the configuration for a masking rule. The 'Data Policy Name 1' is 'sha256'. The 'Masking Rule 1' is 'Hash (SHA256)'. The 'Principals 1' list includes 'paggiechen8866@gmail.com'. The role 'bigquerydatapolicy.maskedReader' is selected for the principals.

Data Policy Name 1: sha256

Masking Rule 1: Hash (SHA256)

Principals 1: paggiechen8866@gmail.com

Grant the bigquerydatapolicy.maskedReader role to the list of principals.

## Masked Reader Role (Grant on Policy Tag)

paggie-svc-project / Datasets / paggie\_ml / Tables / demo\_finance

demo\_finance

Query

Chat

Open in

+

Schema

Details

Preview

Table Explorer

Preview

Insights

Row	region	revenue	margin
1	AMER	808585.22	173593.12
2	APAC	151675.12	23147.33
3	EMEA	527471.23	113096.77
4	APAC	916885.92	202511.48
5	AMER	153416.53	26817.16
6	APAC	968805.84	110177.68
7	APAC	143594.66	39670.66

```
1 SELECT * FROM 'paggie-svc-project.paggie_ml.demo_finance' LIMIT 1000
```

Query completed

Using on-demand processing quota

Query results

Job information	Results	Visualization	JSON	Execution details	Execution graph
Row	customer	sales	region	revenue	margin
1	jk6UeODBJ0dgy/WMBt0r0Agw...	7MsrZplVRYIw+vrAPIQI/17wq...	AMER	808585.22	173593.12
2	jk6UeODBJ0dgy/WMBt0r0Agw...	7MsrZplVRYIw+vrAPIQI/17wq...	APAC	151675.12	23147.33
3	jk6UeODBJ0dgy/WMBt0r0Agw...	7MsrZplVRYIw+vrAPIQI/17wq...	EMEA	527471.23	113096.77
4	HlsWyzR0Xl8V8Pxbz2wyPk...	7MsrZplVRYIw+vrAPIQI/17wq...	APAC	916885.92	202511.48
5	HlsWyzR0Xl8V8Pxbz2wyPk...	7MsrZplVRYIw+vrAPIQI/17wq...	AMER	153416.53	26817.16
6	HlsWyzR0Xl8V8Pxbz2wyPk...	7MsrZplVRYIw+vrAPIQI/17wq...	APAC	968805.84	110177.68
7	HlsWyzR0Xl8V8Pxbz2wyPk...	7MsrZplVRYIw+vrAPIQI/17wq...	APAC	143594.66	39670.66
8	HlsWyzR0Xl8V8Pxbz2wyPk...	7MsrZplVRYIw+vrAPIQI/17wq...	APAC	328163.43	54515.41
9	q7mUmQl29x9XGliehBcfuMBI+	7MsrZplVRYIw+vrAPIQI/17wq...	AMER	801037.4	128467.01
10	q7mUmQl29x9XGliehBcfuMBI+	7MsrZplVRYIw+vrAPIQI/17wq...	AMER	463728.05	136980.64
11	q7mUmQl29x9XGliehBcfuMBI+	7MsrZplVRYIw+vrAPIQI/17wq...	EMEA	619900.52	150356.65
12	A+9Wg9+byWnuF08b6gIMj2c4...	7MsrZplVRYIw+vrAPIQI/17wq...	APAC	588316.63	154141.89
13	A+9Wg9+byWnuF08b6gIMj2c4...	7MsrZplVRYIw+vrAPIQI/17wq...	APAC	412082.67	42303.84
14	A+9Wg9+byWnuF08b6gIMj2c4...	7MsrZplVRYIw+vrAPIQI/17wq...	APAC	632110.93	96746.87
15	A+9Wg9+byWnuF08b6gIMj2c4...	7MsrZplVRYIw+vrAPIQI/17wq...	AMER	561590.42	87140.62
16	cBPU+PEPm0D562AW64wYLL...	7MsrZplVRYIw+vrAPIQI/17wq...	APAC	218734.58	51921.44

## Core Concepts

Project, Dataset, Table

Partitioned & Clustered table

Table & View

Authorized view

Row-level access control

Policy Tag

Data Lineage

## Fine-grained Reader Role (Grant on IAM page)

demo\_finance

Query

Chat

Open in

Share

Copy

+

+

+

Schema

Details

Preview

Table Explorer

Preview

Insights

Lineage

Data Profile

Data Quality

Row	customer	sales	region	revenue	margin
1	Alpine Logistics	David Chen	AMER	808585.22	173593.12
2	Alpine Logistics	David Chen	APAC	151675.12	23147.33
3	Alpine Logistics	David Chen	EMEA	527471.23	113096.77
4	Desert Oasis Ltd	David Chen	APAC	916885.92	202511.48
5	Desert Oasis Ltd	David Chen	AMER	153416.53	26817.16
6	Desert Oasis Ltd	David Chen	APAC	968805.84	110177.68
7	Desert Oasis Ltd	David Chen	APAC	143594.66	39670.66
8	Desert Oasis Ltd	David Chen	APAC	328163.43	54515.41
9	Doha Dynamics	David Chen	AMER	801037.4	128467.01
10	Doha Dynamics	David Chen	AMER	463728.05	136980.64
11	Doha Dynamics	David Chen	EMEA	619900.52	150356.65

## Core Concepts

Project, Dataset, Table  
Partitioned & Clustered table

Table & View

Authorized view

Row-level access control

Policy Tag

Data Lineage

The screenshot displays the Google Cloud Data Catalog interface for the table `total_green_trips_22_21`. The **Lineage** tab is selected, showing a data lineage graph. The graph consists of three nodes: `nyc_green_trips_2022`, `nyc_green_trips_2021`, and `total_green_trips_22_21`. Arrows indicate data flow from the source tables to the target table. The target table node is expanded to show its columns: `vendor_id` and `number_of_trips`. The **Lineage Explorer** sidebar on the left contains the following configuration options:

- Column Level Lineage**: A dropdown menu for selecting a column name.
- Direction**: Two checked checkboxes, **Upstream** and **Downstream**, are highlighted with a red box.
- Time Range**: Three radio buttons: **None** (selected), **Start Time**, and **End Time**. Below the radio buttons is a text input field containing `None`, a `EAT` button, and a calendar icon.

At the bottom of the sidebar is an **Apply** button. The top navigation bar includes options for Query, Chat, Open in, Share, Copy, Snapshot, Delete, and Export. The sidebar also includes options for Schema, Details, Preview, Table Explorer, Insights, Lineage, Data Profile, and Data Quality.

# Build Pipelines Data Canvas Dataform

The screenshot displays a data canvas interface with the following components:

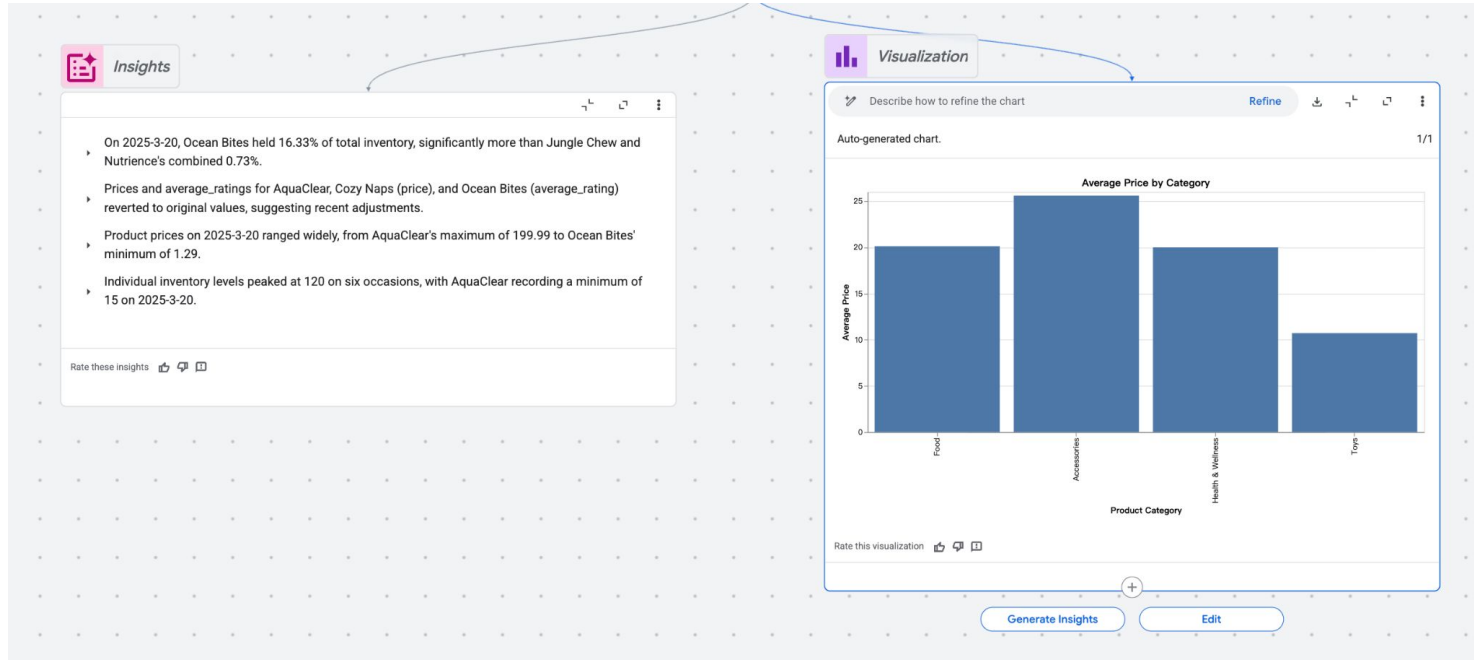
- Top Bar:** Search, zoom (44%), and navigation icons.
- Left Panel:** A sidebar with a search bar and a list of tables.
- Schema Views:**
  - cymbal\_products:** A table with columns: product\_id (INTEGER), product\_name (STRING), brand (STRING), category (STRING), subcategory (STRING), animal\_type (STRING), search\_keywords (STRING), price (FLAT), description (STRING), inventory\_level (INTEGER), supplier\_id (INTEGER), and average\_rating (FLAT).
  - cymbal\_product\_images:** A table with columns: url (STRING), generation (INTEGER), content\_type (STRING), size (INTEGER), md5\_hash (STRING), updated (TIMESTAMP), metadata (RECORD), and ref (RECORD).
- SQL Editor:** A central window with a text area containing a SQL query, a 'Run' button, and a 'Query completed' status.
- Query Results:** A table showing the output of the SQL query with columns: product\_id, product\_name, brand, and category. The results include:

Row	product_id	product_name	brand	category
1	11	Nibble Time Hamster Food	Nibble Time	Food
2	23	Purfect Perch Cat Scratcher	Purfect Perch	Accessories
- Bottom Bar:** A navigation bar with buttons for 'Query these results', 'Generate Insights', 'Visualize', 'Set Destination', and 'Join'. A blue circular button with a plus sign is also present.

# Build Pipelines

## Data Canvas

### Dataform



## Build Pipelines

Data Canvas

Dataform

Development Workspaces

Workflow Execution Logs

Releases & scheduling

Settings

+ Create development workspace

Delete selected workspaces

Development workspaces contain an editable copy of your team's repository. Using development workspaces, you can develop code without affecting other users, commit changes, and push commits to your remote Git repository. If your repository is connected to a remote Git provider, changes from your development workspace will be pushed to a remote branch named after your development workspace. Otherwise, your changes will be pushed to your default branch. [Learn more](#)

Name ↑

[jackson](#)

[paggie](#)

branch

Development Workspaces




Workflow Execution Logs

Releases & scheduling

Settings

 Connect with Git

 Configure private npm packages

Name	demo
Location	me-central1
Service account 	Default Dataform service account 
actAs permission checks	
Enforcement	Don't enforce
Encryption Settings	
Type	Google-managed

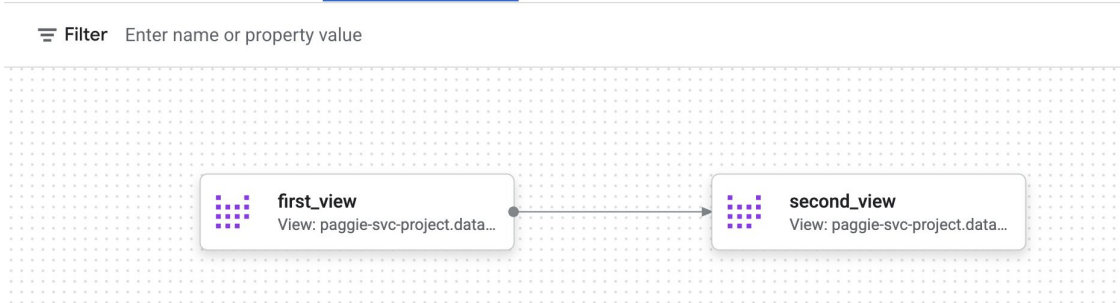
# Build Pipelines

Data Canvas

## Dataform

The screenshot shows the Dataform editor interface. At the top, there are navigation tabs: 'Code' (selected), 'Compiled graph', 'Executions', 'Start execution', and 'Ask Agent'. Below the tabs is a file explorer on the left with a search bar and a list of files: '\*definitions', '\*first\_view.sqlx', '\*second\_view.sqlx' (highlighted), 'includes', '\*.gitignore', and '\*workflow\_settings.yaml'. The main area is a code editor showing the content of 'definitions/second\_view.sqlx':

```
1 config { type: "view" }
2
3 -- Use the ref() function to manage dependencies.
4 -- Learn more about ref() and other built in functions here:
5
6 SELECT test from ${ref("first_view")}
7
```



## Augment with BigQuery ML

Scenario 1: [Remote Model] Document AI with BQML

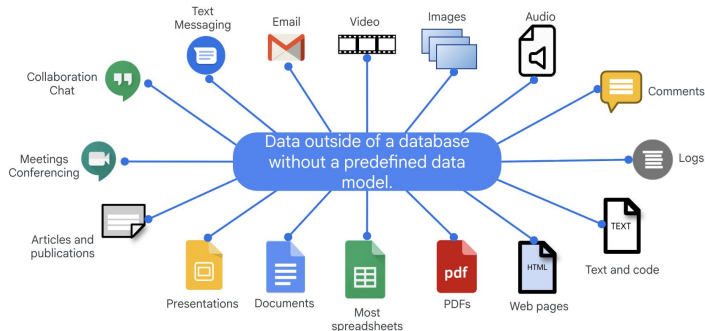
Scenario 2: [Remote Model] Image Data Summarization with BQML

Scenario 3: [Native Model] Custom BQML Model

# The Unstructured Challenge

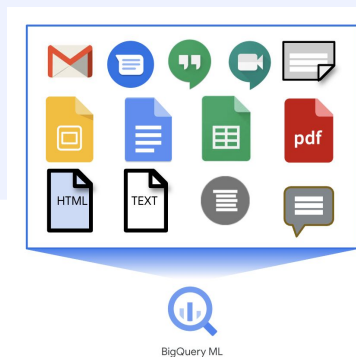
## 80% of enterprise data

is unstructured, trapped in images, PDFs, and audio files. Traditionally, this data was inaccessible to SQL-based analytics.



## Multimodal BQML

We bridge the gap between data engineering and cognitive inference, enabling massive scalability without leaving the BigQuery ecosystem.



# Introducing Object Tables

An **Object Table** is a specialized **external table** in BigQuery that manages **ObjectRef** metadata for unstructured data stored in Cloud Storage, let you analyze unstructured data in Cloud Storage.



## Consistent Metadata

Automatic tracking of size, URI, and timestamps.



## Unified Security

IAM permissions govern both SQL access and file reads.



## AI Native

Directly compatible with ML.GENERATE\_TEXT.

# Introducing Object Tables

An **Object Table** is a specialized **external table** in BigQuery that manages **ObjectRef** metadata for unstructured data stored in Cloud Storage, let you analyze unstructured data in Cloud Storage.

☆ doc\_object\_ta... Lakehouse 🔍 Query 🗨 Chat Open in ▾

Schema Details Insights Lineage Data Profile Data Quality

☰ Filter Enter property name or value

<input type="checkbox"/>	↕ Field name	Type	Mode	Description
<input type="checkbox"/>	uri	STRING	NULLABLE	-
<input type="checkbox"/>	generation	INTEGER	NULLABLE	-
<input type="checkbox"/>	content_type	STRING	NULLABLE	-
<input type="checkbox"/>	size	INTEGER	NULLABLE	-
<input type="checkbox"/>	md5_hash	STRING	NULLABLE	-
<input type="checkbox"/>	updated	TIMESTAMP	NULLABLE	-
<input type="checkbox"/>	▶ metadata	RECORD	REPEATED	-
<input type="checkbox"/>	▶ ref	RECORD	NULLABLE	-

Edit schema

View row access policies

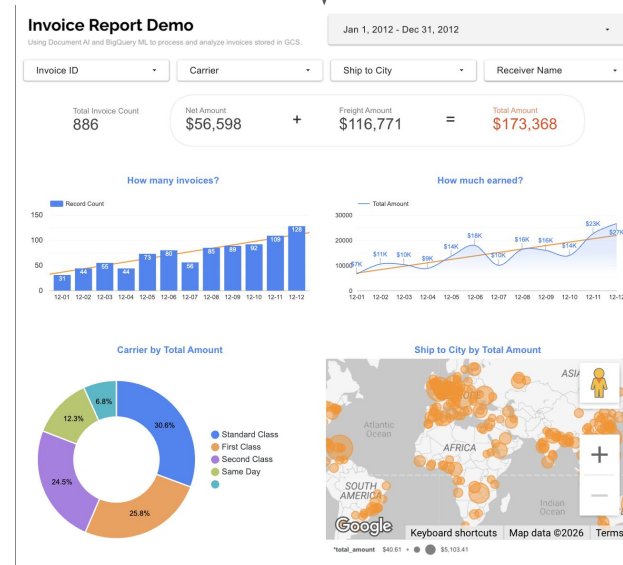
Describe data

# Augment with BigQuery ML

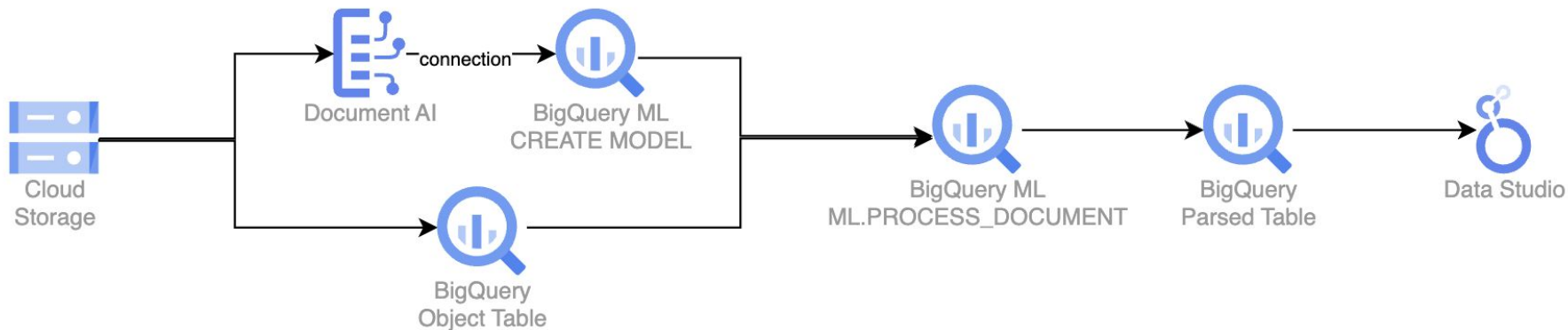
Scenario 1: [Prebuilt API] Document AI with BQML

Scenario 2: [Vertex AI] Image Data Summarization with BQML


Scenario 3: [Native Model] Custom BQML Model




# Architecture Overview





# Step1. Upload invoice PDFs to GCS


Buckets > bqml-invoice-bucket 

[Create folder](#) [Upload](#) [Transfer data](#) [Create batch operations](#)

Filter by name prefix only  |  Filter Filter objects and folders

<input type="checkbox"/>	Name	Size	Type	Created 
<input type="checkbox"/>	 <a href="#">1000+ PDF_Invoice_Folder/</a>	—	Folder	—

## Step2. Create Document AI Processor: Invoice Parser



**Invoice Parser**

Extracts 30+ fields from Invoices:  
ID, amount, line items, etc.

Document AI / Processor gallery

Processor gallery


Search processors

Features	
Trainable	9

Type	
General	3
Specialized	13


Access status	
Public	13
Private	7

Region	
asia-south1	6
asia-southeast1	7
australia-southeast1	7
eu	19
europa-west2	5




**Bank Statement Parser**

Extract from Bank Statements including name, transactions, etc.




**Expense Parser**

Extracts from Receipts, including supplier, total amount, tip, etc.




**Identity Document Proofing**

Predict the validity of ID documents using multiple signals




**Investment and Retirement Statement Parser**



**Invoice Parser**

Extracts 30+ fields from Invoices: ID, amount, line items, etc.



**Lending Doc Splitter/Classifier**

Identify documents in a large file & classify known lending doc types

# Step2. Create Document AI Processor: Invoice Parser

← invoice\_parser [Disable processor](#) [Activity](#) [Get Code](#) [Learn](#)

[Processor details](#) Train Evaluate & test Manage versions Monitoring

### Basic information

Name	invoice_parser
ID	c4c2e8b155f5543f
Status	✔ Enabled
Processor Type	Invoice Parser
Created	Apr 17, 2026, 2:13:48 PM
Encryption Type	Google-managed
Region	us

### Prediction

Prediction endpoint <a href="#">?</a>	<a href="https://us-documentai.googleapis.com/v1/projects/304589458397/locations/us/processors/c4c2e8b155f5543f:process">https://us-documentai.googleapis.com/v1/projects/304589458397/locations/us/processors/c4c2e8b155f5543f:process</a> <a href="#">📄</a>
---------------------------------------	---

## Scenario 1: Document AI with BQML

# Step3. Create BigQuery External Connection & Assign IAM

docc-connection

 Query

 Share

 Delete

 Refresh

## Connection info

 Edit Details

Connection ID projects/paggie-svc-project/locations/us/connections/docc-connection

Friendly name

Created Apr 17, 2026, 2:16:14 PM UTC+3

Last modified Apr 17, 2026, 2:16:14 PM UTC+3

Data location us

Description Document AI

Connection type Vertex AI remote models, remote functions, Lakehouse and Spanner (Cloud Resource)

Service account id bqcx-304589458397-ij5i@gcp-sa-bigquery-condel.iam.gserviceaccount.com

Principal 

Name

Role

bqcx-304589458397-ij5i@gcp-sa-bigquery-condel.iam.gserviceaccount.com

Document AI Viewer (Beta)  
Storage Object Viewer

# Step 4. Create BigQuery Model

SQL Editor -  
BigQuery

```
CREATE OR REPLACE MODEL
```

```
`<Project_Name>.<Dataset_Name>.doc_model`
```

```
REMOTE WITH CONNECTION `<Project_Name>.us.doc_conn`
```

```
OPTIONS (
```

```
  REMOTE_SERVICE_TYPE = 'CLOUD_AI_DOCUMENT_V1',
```

```
  DOCUMENT_PROCESSOR = '<Processor_ID>'
```

```
);
```

Query ready to run

Specifies the service to use to create the model:

- [Cloud Natural Language API](#)
- [Cloud Translation API](#)
- [Cloud Vision API](#)
- [Document AI API](#)
- [Speech-to-Text API](#)

# Step 5. Create Object Table

SQL Editor - BigQuery

```
CREATE EXTERNAL TABLE
```

```
`<Project_Name>.<Dataset_Name>.doc_object_table`
```

```
WITH CONNECTION `<Project_Name>.us.doc_conn`
```

```
OPTIONS(
```

```
  object_metadata = 'SIMPLE',
```

```
  uris = ['gs://<Bucket_Name>/*']
```

```
);
```

Query ready to run

paggie-svc-project / Datasets / demo\_bqml / Tables / doc\_object\_table

☆ doc\_object\_ta... Lakehouse 🔍 Query 🗨 Chat 🗄 Open in ▾

[Schema](#) [Details](#) [Insights](#) [Lineage](#) [Data Profile](#) [Data Quality](#)

☰ Filter Enter property name or value

<input type="checkbox"/>	↕ Field name	Type	Mode	Description
<input type="checkbox"/>	uri	STRING	NULLABLE	-
<input type="checkbox"/>	generation	INTEGER	NULLABLE	-
<input type="checkbox"/>	content_type	STRING	NULLABLE	-
<input type="checkbox"/>	size	INTEGER	NULLABLE	-
<input type="checkbox"/>	md5_hash	STRING	NULLABLE	-
<input type="checkbox"/>	updated	TIMESTAMP	NULLABLE	-
<input type="checkbox"/>	▶ metadata	RECORD	REPEATED	-
<input type="checkbox"/>	▶ ref	RECORD	NULLABLE	-

Edit schema

View row access policies

Describe data

## Step 6. ML.PROCESS\_DOCUMENT

SQL Editor - BigQuery

```
CREATE OR REPLACE TABLE `AS  
SELECT *  
FROM ML.PROCESS_DOCUMENT(  
  MODEL `  TABLE `);
```

Query ready to run

☆ parsed\_table 🔍 Query 🗨 Chat 📄 Open in ▾ 👤 Share ▾

Schema Details Preview Table Explorer Preview Insights

Filter Enter property name or value

<input type="checkbox"/>	Field name	Type	Mode
<input type="checkbox"/>	ml_process_document_result	JSON	NULLABLE
<input type="checkbox"/>	ml_process_document_status	STRING	NULLABLE
<input type="checkbox"/>	invoice_type	STRING	NULLABLE
<input type="checkbox"/>	currency	STRING	NULLABLE
<input type="checkbox"/>	due_date	STRING	NULLABLE
<input type="checkbox"/>	invoice_date	STRING	NULLABLE
<input type="checkbox"/>	invoice_id	STRING	NULLABLE
<input type="checkbox"/>	net_amount	STRING	NULLABLE
<input type="checkbox"/>	purchase_order	STRING	NULLABLE
<input type="checkbox"/>	receiver_name	STRING	NULLABLE
<input type="checkbox"/>	receiver_tax_id	STRING	NULLABLE
<input type="checkbox"/>	supplier_iban	STRING	NULLABLE
<input type="checkbox"/>	supplier_name	STRING	NULLABLE
<input type="checkbox"/>	supplier_payment_ref	STRING	NULLABLE
<input type="checkbox"/>	supplier_registration	STRING	NULLABLE
<input type="checkbox"/>	supplier_tax_id	STRING	NULLABLE

Edit schema View row access policies Describe data

# Scenario 1: Document AI with BQML

## Step 7. Data Studio Dashboard

### Invoice Report Demo

Using Document AI and BigQuery ML to process and analyze invoices stored in GCS.

Jan 1, 2012 - Dec 31, 2012

Invoice ID  Carrier  Ship to City  Receiver Name

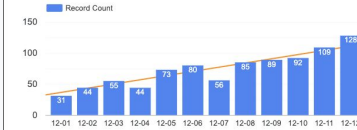
Total Invoice Count  
886

Net Amount  
\$56,598

+ Freight Amount  
\$116,771

= Total Amount  
\$173,368

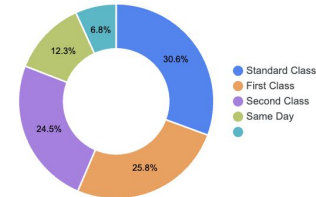
How many invoices?



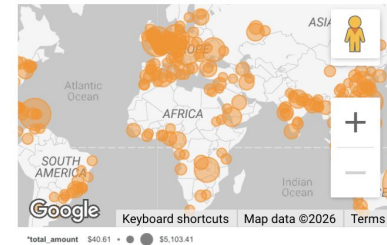
How much earned?



Carrier by Total Amount



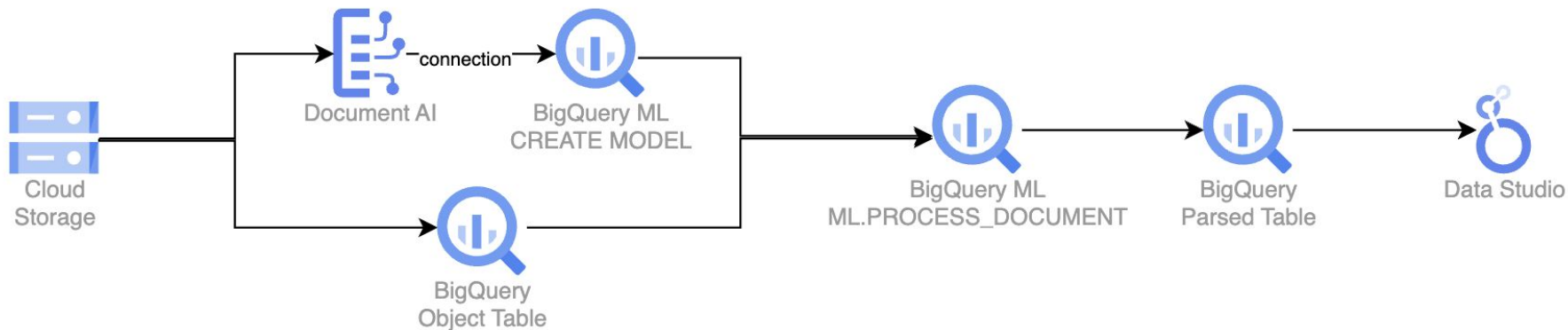
Ship to City by Total Amount



Invoice Overview

Invoice ID	Invoice Title	Net Amount	Freight Amount	Total Amount
1. 42837	invoice_Anna Gayman_42837.pdf	\$922.95	\$282.09	\$1,205.04
2. 48022	invoice_Clay Cheatham_48022.pdf	\$855.15	\$325.04	\$1,180.19
3. 17190	invoice_Ralph Arnett_17190.pdf	\$957.66	\$141.92	\$1,099.52
4. 21585	invoice_Jason Klamczynski_21585.pdf	\$985.92	\$80.66	\$1,066.58
5. 28157	invoice_Alex Russell_28157.pdf	\$942.01	\$113.45	\$1,055.46
6. 5434	invoice_Yana Sorensen_5434.pdf	\$964.96	\$66.7	\$1,031.66
7. 26495	invoice_Sylvia Foulston_26495.pdf	\$947.03	\$70.64	\$1,017.67
8. 48563	invoice_Dennis Bolton_48563.pdf	\$848.4	\$97.54	\$945.94
9. 22559	invoice_Michael Dominguez_22559.pdf	\$862.34	\$81.63	\$943.97
10. 42483	invoice_Carol Darley_42483.pdf	\$829.32	\$99.68	\$929
11. 42617	invoice_Peter McVee_42617.pdf	\$642.69	\$285.83	\$928.52
12. 8440	invoice_Seán O'Donnell_8440.pdf	\$842.55	\$84.63	\$927.18
13. 8265	invoice_Toby Ritter_8265.pdf	\$785.87	\$141.21	\$927.08
14. 50119	invoice_Rose O'Brian_50119.pdf	\$787.92	\$110.6	\$898.52
15. 30570	invoice_Toby Swindell_30570.pdf	null	\$894.77	\$894.77
<b>Grand total</b>		<b>\$56,597.53</b>	<b>\$116,770.73</b>	<b>\$173,368.26</b>

# Architecture Overview



## Augment with BigQuery ML

Scenario 1: [Prebuilt API] Document AI with BQML

**Scenario 2: [Vertex AI] Image Data Summarization with BQML**

Scenario 3: [Native Model] Custom BQML Model



<https://storage.googleapis.com/cloud-samples-data/bigquery/tutorials/cymbal-pets/images/aquaclear-aquarium-co2-system.png>



Image Description

## Scenario 2: Image Data Summarization with BQML

# Step 1. Create Connection

SQL Editor - BigQuery

**CREATE OR REPLACE CONNECTION**

```
`us.cymbal_connection` OPTIONS( connection_type  
= 'CLOUD_RESOURCE' );
```

Query ready to run



This creates a **Resource Connection Service Account**. Grant this account the **Vertex AI User** role in IAM.

# Step 2. Load Product Info

SQL Editor - BigQuery

```
LOAD DATA OVERWRITE demo_bqml.cymbal_products
FROM
FILES(
  format = 'avro',
  uris = [
'gs://cloud-samples-data/bigquery/tutorials/cymbal-pets/tables/products/products_*.avro' ] );
```

Query ready to run

Field name	Type	Mode
product_id	INTEGER	NULLABLE
product_name	STRING	NULLABLE
brand	STRING	NULLABLE
category	STRING	NULLABLE
subcategory	STRING	NULLABLE
animal_type	STRING	NULLABLE
search_keywords	STRING	REPEATED
price	FLOAT	NULLABLE
description	STRING	NULLABLE
inventory_level	INTEGER	NULLABLE
supplier_id	INTEGER	NULLABLE
average_rating	FLOAT	NULLABLE
uri	STRING	NULLABLE

## Scenario 2: Image Data Summarization with BQML

# Step 3. Create Product Image Object Table

SQL Editor - BigQuery

```
CREATE EXTERNAL TABLE
demo_bqml.cymbal_product_images
  WITH CONNECTION `us.cymbal_connection`
  OPTIONS (
    object_metadata = 'SIMPLE',
    uris =
['gs://cloud-samples-data/bigquery/tutorials/cymbal-
pets/images/*.png'],
    max_staleness = INTERVAL 30 MINUTE,
    metadata_cache_mode = AUTOMATIC);
```

Query ready to run

Field name	Type	Mode
uri	STRING	NULLABLE
generation	INTEGER	NULLABLE
content_type	STRING	NULLABLE
size	INTEGER	NULLABLE
md5_hash	STRING	NULLABLE
updated	TIMESTAMP	NULLABLE
▶ metadata	RECORD	REPEATED
▶ ref	RECORD	NULLABLE

## Scenario 2: Image Data Summarization with BQML

### Step 4. Inner Join both

SQL Editor - BigQuery

```
CREATE OR REPLACE TABLE
demo_bqml.cymbal_products_mm
AS
SELECT cymbal_products.* EXCEPT (uri), ot.ref AS
image FROM demo_bqml.cymbal_products
INNER JOIN demo_bqml.cymbal_product_images ot
ON ot.uri = cymbal_products.uri;
```

Query ready to run



## Scenario 2: Image Data Summarization with BQML

# Step 5. Register Gemini Model

SQL Editor - BigQuery

```
CREATE OR REPLACE MODEL `demo_bqml.gemini`  
  REMOTE WITH CONNECTION `us.cymbal_connection`  
  OPTIONS (ENDPOINT = 'gemini-2.0-flash');
```

Query ready to run

### Vertex AI Integration

This entry in the **Model Registry** allows BigQuery to send inference requests to Vertex AI endpoints via standard SQL syntax.

## Scenario 2: Image Data Summarization with BQML

# Step 6. Generate image\_description by inferencing Gemini

```
CREATE OR REPLACE TABLE demo_bqml.cymbal_products_mm
AS
SELECT
  product_id,
  product_name,
  brand,
  category,
  subcategory,
  animal_type,
  search_keywords,
  price,
  description,
  inventory_level,
  supplier_id,
  average_rating,
  image,
  image_description
FROM
  AI.GENERATE_TABLE(
    MODEL `demo_bqml.gemini`,
    (
      SELECT
        ('Can you describe the following image? ', OBJ.GET_ACCESS_URL(image, 'r')) AS prompt,
        *
      FROM
        demo_bqml.cymbal_products_mm
    ),
    STRUCT('image_description STRING' AS output_schema));
```

The `AI.GENERATE_TABLE` function is a BigQuery SQL command that allows you to use generative AI model to transform unstructured data, such as raw text or images, directly into a structured table format.

# Step 7. Review



<https://storage.googleapis.com/cloud-samples-data/bigquery/tutorials/cymbal-pets/images/aquaclear-aquarium-co2-system.png>

### Gemini Generated Image Description:

*The image displays a laboratory setup with a clear glass Aquaclear brand device featuring a spiral internal structure. A transparent tube connects this device to a regulator with a gauge and a black adjustment knob, all set against a gradient gray background.*

## Augment with BigQuery ML

Scenario 1: [Prebuilt API] Document AI with BQML

Scenario 2: [Vertex AI] Image Data Summarization with BQML

**Scenario 3: [Native Model] Custom BQML Model**

# Demo BQML–Predict NYC taxi fares using BQML

In this demo, we'll build a model to predict taxi fares in NYC using one of BigQuery's public datasets `nyc-tlc.yellow.trips`.

Category	Key Columns	Description
<b>Time</b>	<code>pickup_datetime</code> , <code>dropoff_datetime</code>	The precise date and time when the meter was engaged and disengaged.
<b>Location</b>	<code>pickup_longitude</code> , <code>pickup_latitude</code> / <code>PULocationID</code>	The longitude/latitude (older data) or TLC Taxicab Boundary Zone (newer data) where the trip began/ended.
<b>Fare</b>	<code>fare_amount</code> , <code>tip_amount</code> , <code>tolls_amount</code>	The cost components: base fare, credit card tips, and bridge/tunnel tolls.
<b>Trip Details</b>	<code>passenger_count</code> , <code>trip_distance</code>	The number of passengers (driver-entered) and the elapsed trip distance reported by the taximeter.
<b>Payment</b>	<code>payment_type</code>	A numeric code signifying how the passenger paid (e.g., 1=Credit Card, 2=Cash).

## Scenario 3: Custom BQML Model

# Step 1. Create Dataset

```
#standardSQL
create or replace table `taxi.taxi300k` as
WITH taxi_preproc AS (
SELECT
  ABS(MOD(FARM_FINGERPRINT(STRING(pickup_datetime)), 10000)) AS dataset,
  (tolls_amount + fare_amount) AS fare_amount,
  pickup_datetime,
  EXTRACT(DAYOFWEEK FROM pickup_datetime) AS dayofweek,
  EXTRACT(HOUR FROM pickup_datetime) AS hourofday,
  pickup_longitude AS pickuplon,
  pickup_latitude AS pickuplat,
  dropoff_longitude AS dropofflon,
  dropoff_latitude AS dropofflat,
  passenger_count
FROM
  `nyc-tlc.yellow.trips`
WHERE
  trip_distance > 0
  AND fare_amount >= 2.5
  AND fare_amount < 200
  AND pickup_longitude > -78
  AND pickup_longitude < -70
  AND dropoff_longitude > -78
  AND dropoff_longitude < -70
  AND pickup_latitude > 37
  AND pickup_latitude < 45
  AND dropoff_latitude > 37
  AND dropoff_latitude < 45
  AND passenger_count > 0
  AND ABS(MOD(FARM_FINGERPRINT(STRING(pickup_datetime)), 10000)) < 3
)
SELECT
  dataset,
  fare_amount,
  pickup_datetime,
  hourofday,
  dayofweek,
  CAST(dayofweek * 24 + hourofday AS STRING) AS dayhour,
  pickuplon,
  pickuplat,
  dropofflon,
  dropofflat,
  SQRT(POW((pickuplon - dropofflon),2) + POW(( pickuplat - dropofflat), 2)) AS dist,
  #Euclidean distance between pickup and drop off
  pickuplon - dropofflon AS londiff,
  pickuplat - dropofflat AS latdiff,
  passenger_count
FROM taxi_preproc
WHERE dataset < 3
```

# Step 2. Create/Train Model

```
CREATE OR REPLACE MODEL
  taxi.taxifare_dnn OPTIONS (model_type='dnn_regressor',
    hidden_units=[144, 89, 55],
    labels=['fare_amount']) AS
SELECT
  fare_amount,
  hourofday,
  dayofweek,
  pickuplon,
  pickuplat,
  dropofflon,
  dropofflat,
  passenger_count
FROM
  `taxi.taxi300k`
WHERE
  dataset = 0;
```

# Step 3. Evaluation

paggie-svc-project / Datasets / demo\_bq

☆ taxifare\_dnn [Refresh](#)

Details Training metrics Evaluation

Mean absolute error	2.5286
Mean squared error	14.7226
Mean squared log error	0.0813
Median absolute error	2.0137
R squared	0.8204

Metric	Definition & Meaning
<b>Mean absolute error (MAE)</b>	The average of the absolute differences between the predicted fare and the actual fare. On average, your predictions are off by <b>\$2.53</b> .
<b>Mean squared error (MSE)</b>	The average of the squared differences. Because it squares the error, it penalizes large outliers more heavily than MAE.
<b>Mean squared log error</b>	Measures the relative error between the log of the predicted and actual values. It is useful when the target values have a wide range.
<b>Median absolute error</b>	The median of all error magnitudes. 50% of your predictions have an error of less than <b>\$2.01</b> . It is more robust to outliers than MAE.
<b>R squared</b>	The "Coefficient of Determination". It indicates that <b>82.04%</b> of the variance in the taxi fare is explained by your model. Closer to 1.0 is better.

## Scenario 3: Custom BQML Model

# Step 4. Predict

```
SELECT
  fare_amount,
  predicted_fare_amount,
  hourofday,
  dayofweek,
  pickuplon,
  pickuplat,
  dropofflon,
  dropofflat,
  passenger_count
FROM
  ML.PREDICT(MODEL taxi.taxifare_dnn, (SELECT
  fare_amount,
  hourofday,
  dayofweek,
  pickuplon,
  pickuplat,
  dropofflon,
  dropofflat,
  passenger_count
FROM
  `taxi.taxi300k`
WHERE
  dataset = 2
  ))
LIMIT 10
```

1 SELECT  
2 fare\_amount,  
3 predicted\_fare\_amount,  
✓ This query will process 23.22 MB when run.  
Using on-demand processing quota Processing location: US X

Query results [+ Create conversation](#) [Save results](#) [Open in](#)

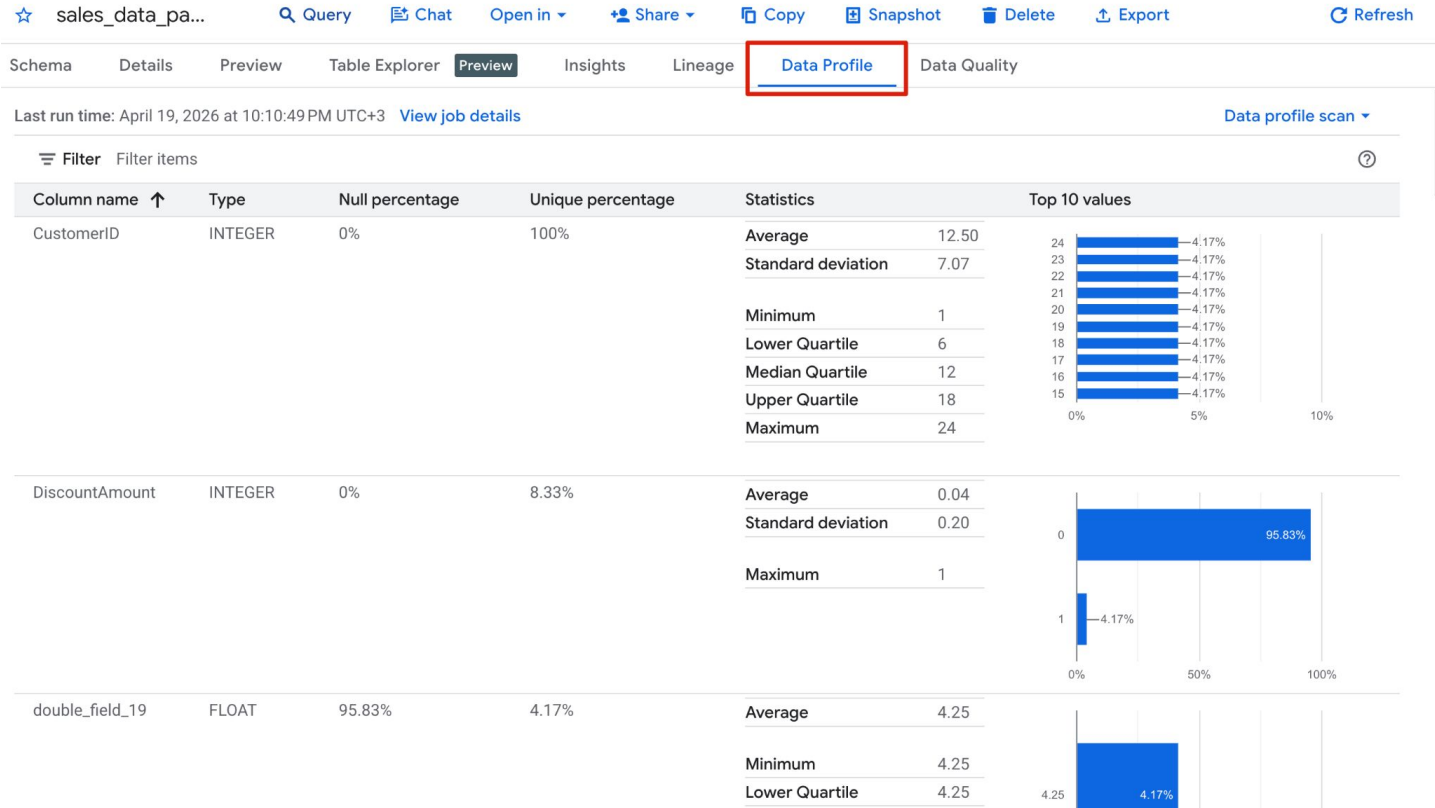
Job information **Results** Visualization JSON Execution details Execution graph

Row	fare_amount	predicted_fare_a...	hourofday	dayofweek	pickuplon	pickuplat
1	4.6	7.020443916320...	4	4	-73.958026	40.77446
2	12.2	14.52977752685...	4	4	-74.006158	40.74366
3	9.8	12.53529262542...	4	4	-73.999302	40.72825
4	6.1	8.04678726196289	4	4	-74.005826	40.7402
5	4.1	7.401981353759...	4	4	-73.962267	40.76062
6	6.9	10.7291841506958	4	4	-74.006501	40.73972
7	5.7	10.26519107818...	4	4	-73.949382	40.80229
8	4.9	9.40158748626709	4	4	-73.978439	40.7413
9	4.1	9.242655754089...	4	4	-73.987827	40.74882
10	4.5	7.433818340301...	4	4	-73.960693	40.77812

# Data Governance

## Data Profiling

### Data Quality Scan



# Data Governance

Data Profiling

Data Quality Scan

Select rules for editing



Choose columns \*

OrderID X

Browse

Choose rule types

NULL check

Delete

Filter Filter items



<input type="checkbox"/>	Column name ↑	Rule name	Rule type	Rule template	Source	Source type	Evaluation
<input type="checkbox"/>	OrderID	-	Null check	-	Predefined rule type	-	Per row

Thank You!