

Quick tip: Review the prerequisites before you run the lab

Start Lab

01:20:00

# Preparing and Aggregating Data for Visualizations using Cloud Dataprep

🧪 Lab ⌚ 1 hour 20 minutes

🎥 1 Credit ✓ Introductory



This lab may incorporate AI tools to support your learning.

*This lab was developed with our partner, [Alteryx Designer Cloud \(Trifacta\)](#). Your personal information may be shared with Trifacta, the lab sponsor, if you have opted-in to receive product updates, announcements, and offers in your Account Profile.*



OSCAR . SILVA ABARCA

Member since 2025

301 Credits

[Buy credits or subscription](#)

[Dashboard](#)

[Activities](#)

[Settings](#)

[Sign Out](#)

[Privacy](#)

[Terms](#)

Task 6. Optional - Export your flow

Congratulations!

Google Cloud

qwiklabs-gcp-04-61e05ae8abd6

🔍 ✨ 📄 1 ? ⋮ S

Cloud Storage

Open project picker (Ctrl Q)

Bucket details

Go to path Refresh Learn

Overview

Buckets

Monitoring

Settings

Storage Intelligence

Insights datasets

Configuration

Marketplace

Release Notes

<|

oscar\_silva

Location: us-central1 (Iowa) Storage class: Standard Public access: Subject to object ACLs Protection: Soft Delete

Objects Configuration Permissions Protection Lifecycle

Buckets > oscar\_silva

Create folder Upload Transfer data Other services

Filter by name prefix only Filter Show Live objects only

<input type="checkbox"/>	Name	Size	Type	Created	Storage class	Last modified
No rows to display						

(qwiklabs-gcp-04-61e05ae8abd6) Open Editor

```
KET/
Creating gs://OSCAR_SILVA/...
BadRequestException: 400 Invalid bucket name: 'OSCAR_SILVA'
student_00_a8a8813866c3@cloudshell:~ (qwiklabs-gcp-04-61e05ae8abd6) $ PROJECT_ID=$(gcloud config get-value project)
BUCKET=oscar_silva
gsutil mb -l us-central1 gs://$BUCKET/
Your active configuration is: [cloudshell-20820]
Creating gs://oscar_silva/...
student_00_a8a8813866c3@cloudshell:~ (qwiklabs-gcp-04-61e05ae8abd6) $ gsutil cp ufo_sighting.csv gs://$BUCKET/
CommandException: No URLs matched
student_00_a8a8813866c3@cloudshe
```

Uploads and qwiklabs-gcp-04-61e05...operations

ufo\_sighting.csv

Google Cloud

qwiklabs-gcp-04-61e05ae8abd6

🔍 ✨ 📄 1 ? ⋮ S

Cloud Storage

Open project picker (Ctrl Q)

Bucket details

Go to path Refresh Learn

Overview

Buckets

Monitoring

Settings

Storage Intelligence

Insights datasets

Configuration

Marketplace

Release Notes

<|

oscar\_silva

Location: us-central1 (Iowa) Storage class: Standard Public access: Subject to object ACLs Protection: Soft Delete

Objects Configuration Permissions Protection Lifecycle

Buckets > oscar\_silva

Create folder Upload Transfer data Other services

Filter by name prefix only Filter Show Live objects only

Name	Size	Type	Created	Storage class	Last modified
No rows to display					

(qwiklabs-gcp-04-61e05ae8abd6) Open Editor

```
KET/
Creating gs://OSCAR_SILVA/...
BadRequestException: 400 Invalid bucket name: 'OSCAR_SILVA'
student_00_a8a8813866c3@cloudshell:~ (qwiklabs-gcp-04-61e05ae8abd6) $ PROJECT_ID=$(gcloud config get-value project)
BUCKET=oscar_silva
gsutil mb -l us-central1 gs://$BUCKET/
Your active configuration is: [cloudshell-20820]
Creating gs://oscar_silva/...
student_00_a8a8813866c3@cloudshell:~ (qwiklabs-gcp-04-61e05ae8abd6) $ gsutil cp ufo_sighting.csv gs://$BUCKET/
CommandException: No URLs matche
student_00_a8a8813866c3@cloudshe
```

Uploads and qwiklabs-gcp-04-61e05...operations

ufo\_sighting.csv

+ New

Edit

Import Data and Add to Flow

Search... /

Upload

Cloud Storage

Google Sheets

BigQuery

Choose a file or folder

CloudStorage / oscar\_silva

Create Dataset with Parameters

Search...

Show hidden

	NAME	SIZE	LAST UPDATED
+ [icon]	ufo_sighti...	12.4MB	Today at 7:57 PM

Close X

1 New Dataset Clear All

[icon] ufo\_sighting.csv

Add a Description

Preview not available.

Edit settings

**FROM CLOUD STORAGE TO BIGQUERY**

**Initial data**

**Run**

**Recipe**

Date_time	city	state/province	country
1906 - 2014	17,277 Categories	52 Categories	6 Categories
Date_time	city	state/province	country
10-10-1949 20:30	san marcos	tx	us
10-10-1949 21:00	lackland afb	tx	
10-10-1955 17:00	chester (uk/england)		gb
10-10-1956 21:00	edna	tx	us
10-10-1968 20:00	kaneohe	hi	us
10-10-1961 19:00	bristol	tn	us
10-10-1965 21:00	penarth (uk/wales)		gb
10-10-1965 23:45	norwalk	ct	
10-10-1966 20:00	pell city	al	us
10-10-1966 21:00	live oak	fl	us
10-10-1968 13:00	hawthorne	ca	us
10-10-1968 19:00	brevard	nc	us
10-10-1970 16:00	bellmore	ny	us
10-10-1970 19:00	manchester	ky	us
10-10-1971 21:00	lexington	nc	us
10-10-1972 19:00	harlan county	ky	us
10-10-1972 22:30	west bloomfield	mi	us
10-10-1973 19:00	niantic	ct	us
10-10-1973 23:00	bermuda nas		
10-10-1974 19:30	hudson	ma	us
10-10-1974 21:30	cardiff (uk/wales)		gb
10-10-1974 23:00	hudson	ks	us
10-10-1975 17:00	north charleston	sc	us
10-10-1976 20:30	washougal	wa	us
10-10-1976 22:00	stoke mandeville (uk/england)		gb
10-10-1977 12:00	san antonio	tx	us
10-10-1977 22:00	louisville	ky	us
10-10-1978 2:00	elmont	ny	us
10-10-1979 0:00	poughkeepsie	ny	us
10-10-1979 22:00	saddle lake (canada)	ab	
10-10-1979 22:00	san diego	ca	us
10-10-1979 22:00	security	co	us
10-10-1980 19:00	houston	tx	us

Show / Hide data grid options 11 Columns 64,781 Rows 5 Data Types

The recipe is empty

Add New Step



From Cloud Storage to BigQuery > ufo\_sighting  
Job 32538556  
Finished Today at 5:39 PM

[View BigQuery job](#) [...](#)

[Overview](#) [Output destinations](#) [Profile](#) [Dependency graph](#) [Data sources](#)

	pacific ocean (western)
	pacific ocean (no state)
1971-02-01T20:00:00	pacific ocean
	pacific ocean (in-flight; from tokyo airport to honolulu)

11 columns 80332 rows This is a preview of the current data in your destination. It might not reflect the output from this particular job run.

[View on BigQuery](#) [View details](#)

Execution stages

✓ **Schema validation**  
Completed Today at 5:39 PM, started Today at 5:39 PM • Ran for 7 sec

Datasets

ufo_sighting.csv	✓ No schema changes found
------------------	---------------------------

[View all](#)

✓ **Transform with profile**  
Completed Today at 5:39 PM, started Today at 5:39 PM • Ran for 28 sec

Environment BigQuery

92% valid values 0.4% mismatching values 7% missing values

[View steps and dependencies](#) [View profile](#) [View BigQuery job](#)

✓ **Publish**  
Completed Today at 5:39 PM, started Today at 5:39 PM • Ran for <1 sec

Activity

avistamientos	✓ Completed
---------------	-------------

[View all](#)

Job summary

Job ID 32538556  
Job status Completed  
Flow [From Cloud Storage to BigQue...](#)  
Output ufo\_sighting

Execution summary

Job type Manual  
User student 95a1774b  
Start time May 31st 2025, 5:39 pm  
Finish time May 31st 2025, 5:39 pm  
Last update May 31st 2025, 5:39 pm  
Duration a few seconds  
memory usage 0.063963136 GB  
Environment BigQuery

Optimization summary

Optimization Enabled

Google Cloud

qwiklabs-gcp-00-d19a8daae6f2

Search (/) for resources, docs, products, and more

Search

1

S

BigQuery

+ Add data

Q \*Untitled...ery

Run

Save

Download

Share

Studio

Query resources

Pipelines & I...

Data transfe...

Dataform

Scheduled q...

Scheduling

Governance

Sharing (An...

Policy tags

Metadata cu...

Administrati...

Monitoring

Jobs explorer

Capacity ma...

BI Engine

Disaster rec...

Recommend...

Migration

Assessment

Partner Center

Settings

Release Notes

Query resources

ed only

00-d19a8daa...

ories

oks

invases

eparations

is

il connections

ap

aprep\_32538...

tamientos

review

repository selected

repository and a

to view its content.

repositories

Untitled query

Run

Save

Download

Share

```
1 SELECT
2   UFO_shape,
3   COUNT(*) AS Sightings
4 FROM
5   `ovni.avistamientos`
6 WHERE
7   country = 'us'
8 GROUP BY
9   UFO_shape
10 ORDER BY
11   Sightings DESC
12 LIMIT 5;
13 #oscarsilva
```

This query will process 2.69 MB when run.

Query results

Save results

Open in

Job information

Results

Chart

JSON

Execution details

Execution graph

Row	UFO_shape	Sightings
1	light	40419
2	triangle	19647
3	circle	18354
4	fireball	15444
5	unknown	13701

Results per page: 50

1 - 5 of 5

<

>

>>

Job history

Refresh

Google Cloud | qwiklabs-gcp-00-d19a8daae6f2 | Search (/) for resources, docs, products, and more

Explorer | + Add data | Search BigQuery resources

Query Editor: **Untitled query**

```

1 SELECT
2   EXTRACT(YEAR FROM DATETIME(Date_time)) AS Year,
3   COUNT(*) AS Sightings
4 FROM
5   `ovni.avistamientos`
6 WHERE
7   country = 'us'
8 GROUP BY
9   Year
10 ORDER BY
11   Sightings DESC;
12 #oscarsilva

```

Query completed

Query results | Save results | Open in

Row	Year	Sightings
1		116442
2	2012	7971
3	2013	7359
4	2011	5352
5	2010	4845
6	2008	4557
7	2007	4365
8	2009	4158
9	2005	3867

Results per page: 50 | 1 - 50 of 79

Job history | Refresh

Looker Studio Reporting - 5/31/25, 6:26 PM

avistamientos

city	length_of_encounter_sec...
1. birmingham (uk/england)	293,677,527
2. ottawa (canada)	248,479,299
3. greenbrier	198,839,745
4. englewood	157,955,517
5. fisley	157,872,480
6. dont know	157,869,608
7. somerset (uk/england)	75,750,885
8. virginia beach	31,815,861
9. oklahoma	31,585,635
10. menifee	31,584,633
11. new canaan	28,412,280

1 - 50 / 19900

length\_of\_encounter\_seconds

country

Let's get started

Drag a field from the Data Panel to the canvas to add a new chart or select a component on the report canvas to edit it.

Data Panel:

- avistamientos - 5/31/25, 6:26 PM
- city
- country
- data\_documented
- data\_type
- described\_duration\_of\_encounter
- description
- latitude
- length\_of\_encounter\_seconds
- longitude
- province
- UFO\_shape
- Record Count

Actions:

- Add a field
- Add a parameter
- Add Data