

Using BigQuery and Cloud Logging to Analyze BigQuery Usage

 [qwiklabs.com/focuses/6100](https://www.qwiklabs.com/focuses/6100)

GSP617



Google Cloud Self-Paced Labs

Overview

Cloud Logging allows you to store, search, analyze, monitor, and alert on log data and events from the Google Cloud including BigQuery. Stackdriver also provides the ability to export certain logs to sinks such as Cloud Pub/Sub, Cloud Storage or BigQuery.

In this lab you view the BigQuery logs inside Cloud Logging, setup a sink to export them back into BigQuery, and then use SQL to analyze the logs.

Setup and Requirements

Before you click the Start Lab button

Read these instructions. Labs are timed and you cannot pause them. The timer, which starts when you click **Start Lab**, shows how long Google Cloud resources will be made available to you.

This Qwiklabs hands-on lab lets you do the lab activities yourself in a real cloud environment, not in a simulation or demo environment. It does so by giving you new, temporary credentials that you use to sign in and access Google Cloud for the duration of the lab.

What you need

To complete this lab, you need:

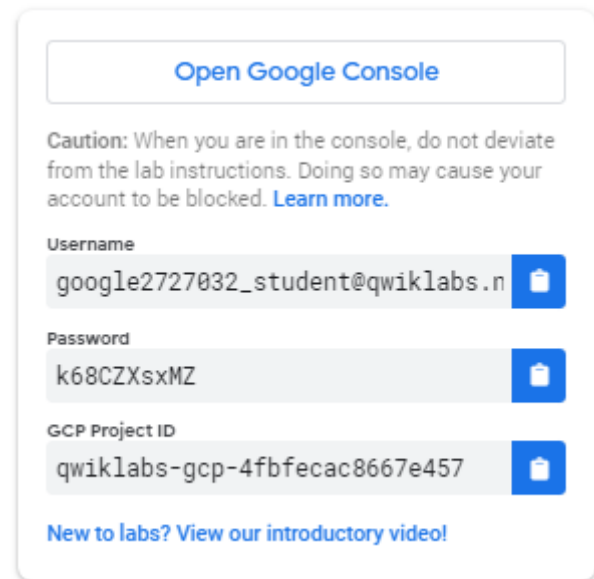
- Access to a standard internet browser (Chrome browser recommended).
- Time to complete the lab.

Note: If you already have your own personal Google Cloud account or project, do not use it for this lab.

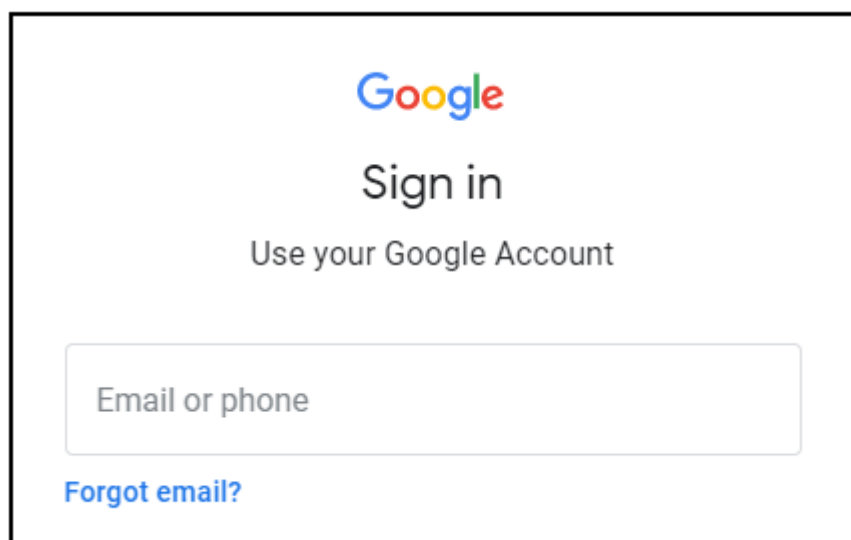
Note: If you are using a Pixelbook, open an Incognito window to run this lab.

How to start your lab and sign in to the Google Cloud Console

1. Click the **Start Lab** button. If you need to pay for the lab, a pop-up opens for you to select your payment method. On the left is a panel populated with the temporary credentials that you must use for this lab.
2. Copy the username, and then click **Open Google Console**. The lab spins up resources, and then opens another tab that shows the **Sign in** page.



This panel, titled "Open Google Console", contains a caution message: "Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked. [Learn more.](#)". Below this, it lists three credentials with copy icons: Username "google2727032_student@qwiklabs.n", Password "k68CZXsxMZ", and GCP Project ID "qwiklabs-gcp-4fbfecac8667e457". At the bottom, there is a link: "New to labs? View our introductory video!"



The Google Sign in page features the Google logo at the top, followed by the text "Sign in" and "Use your Google Account". Below this is a text input field labeled "Email or phone". At the bottom left, there is a link that says "Forgot email?".

Tip: Open the tabs in separate windows, side-by-side.

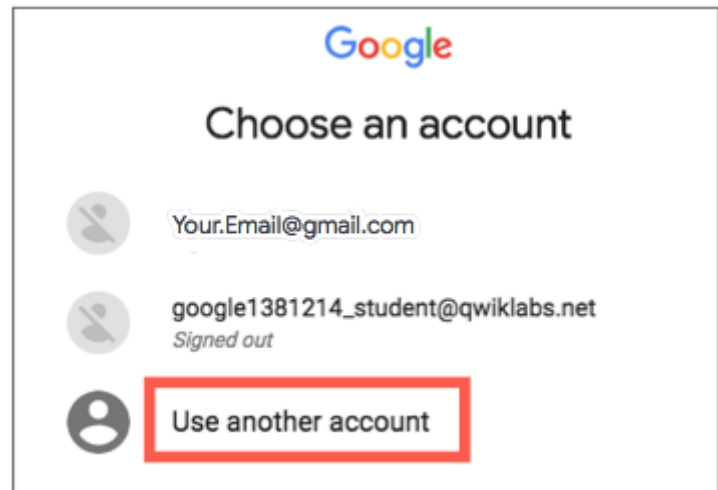
If you see the **Choose an account** page, click **Use Another Account**.

3. In the **Sign in** page, paste the username that you copied from the Connection Details panel. Then copy and paste the password.

Important: You must use the credentials from the Connection Details panel. Do not use your Qwiklabs credentials. If you have your own Google Cloud account, do not use it for this lab (avoids incurring charges).

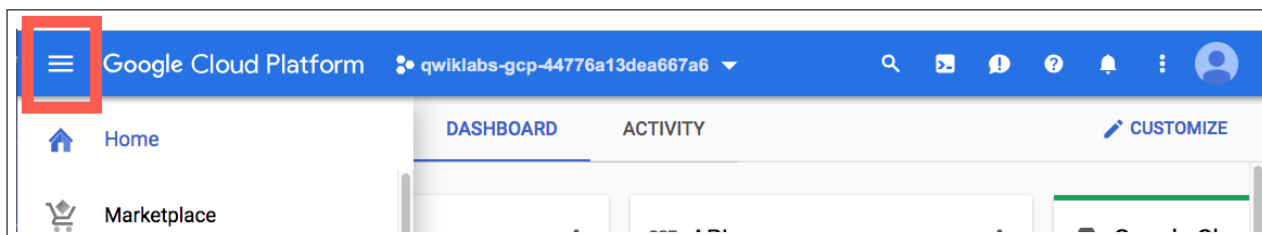
4. Click through the subsequent pages:

- Accept the terms and conditions.
- Do not add recovery options or two-factor authentication (because this is a temporary account).
- Do not sign up for free trials.



After a few moments, the Cloud Console opens in this tab.

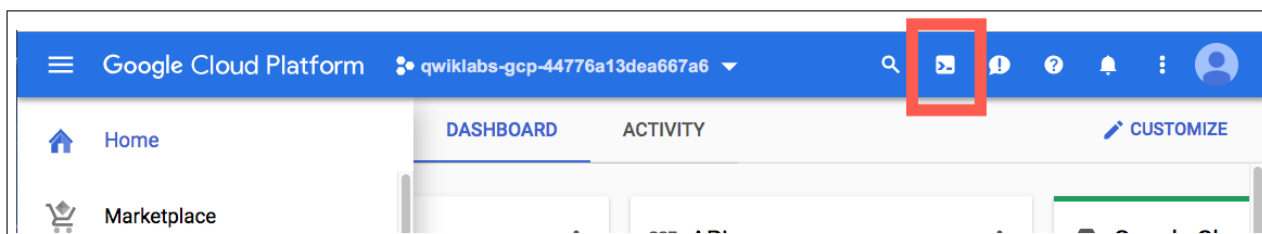
Note: You can view the menu with a list of Google Cloud Products and Services by clicking the **Navigation menu** at the top-left.



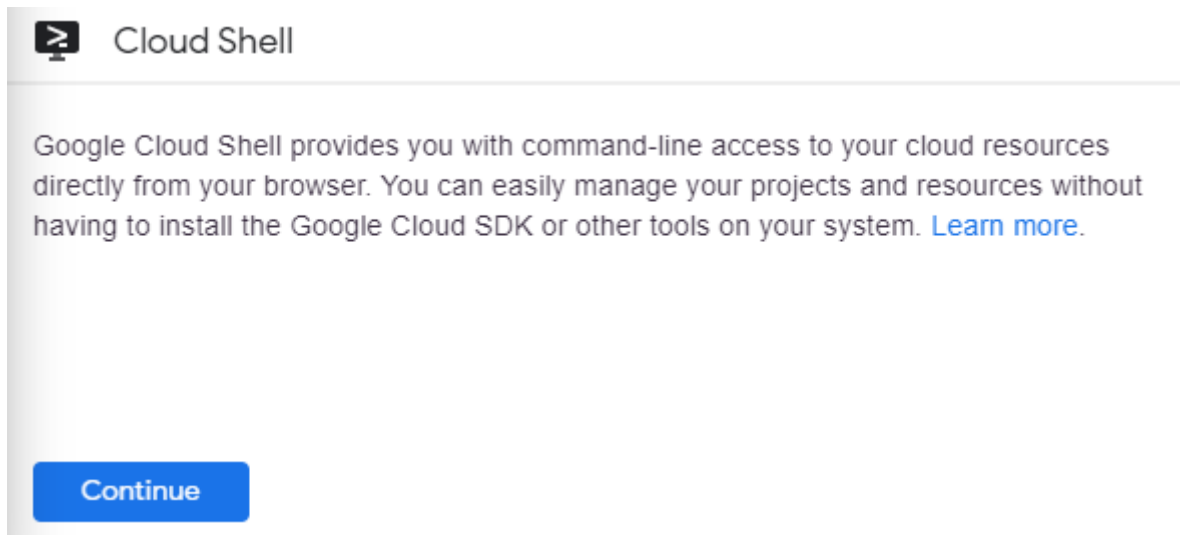
Activate Cloud Shell

Cloud Shell is a virtual machine that is loaded with development tools. It offers a persistent 5GB home directory and runs on the Google Cloud. Cloud Shell provides command-line access to your Google Cloud resources.

In the Cloud Console, in the top right toolbar, click the **Activate Cloud Shell** button.



Click **Continue**.



It takes a few moments to provision and connect to the environment. When you are connected, you are already authenticated, and the project is set to your *PROJECT_ID*. For example:

`gcloud` is the command-line tool for Google Cloud. It comes pre-installed on Cloud Shell and supports tab-completion.

You can list the active account name with this command:

```
gcloud auth list
```

(Output)

```
Credentialed accounts:
- <myaccount>@<mydomain>.com (active)
```

(Example output)

```
Credentialed accounts:
- google1623327_student@quiklabs.net
```

You can list the project ID with this command:

```
gcloud config list project
```

(Output)

```
[core]
project = <project_ID>
```

(Example output)

```
[core]  
project = qwiklabs-gcp-44776a13dea667a6
```

For full documentation of `gcloud` see the [gcloud command-line tool overview](#).

Open BigQuery

Open BigQuery Console

In the Google Cloud Console, select **Navigation menu** > **BigQuery**:



Home



BIG DATA



Composer



Dataproc



Pub/Sub



Dataflow



IoT Core



BigQuery



Looker

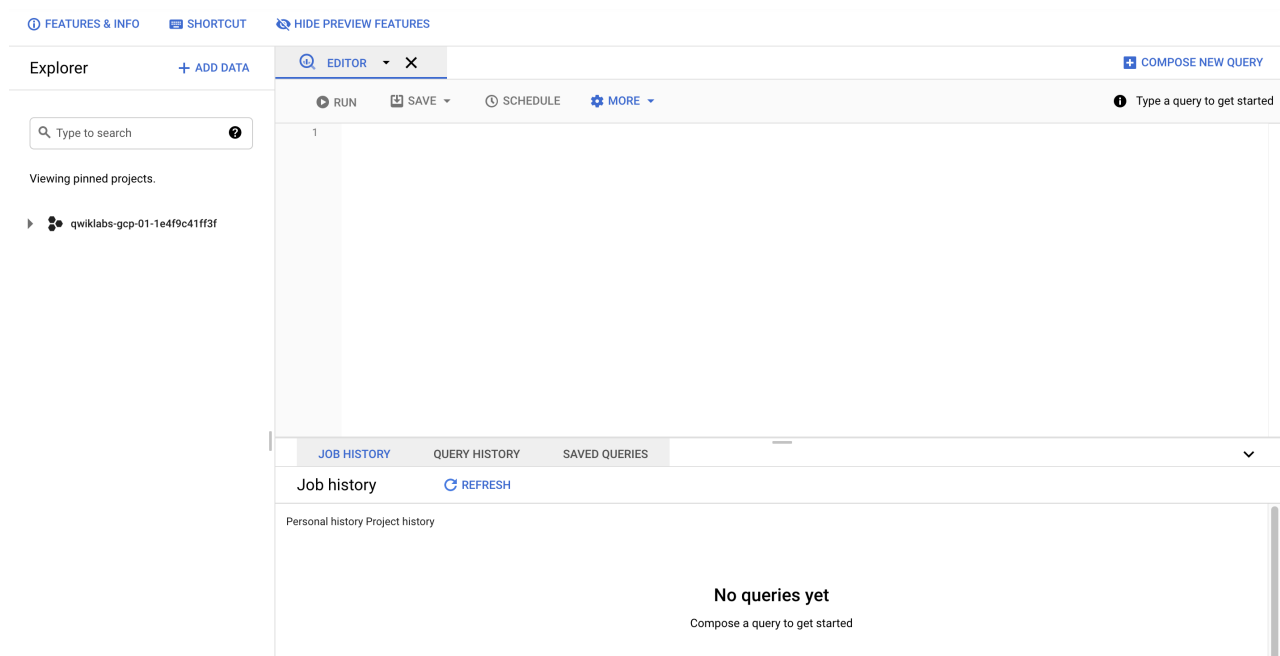


Data Catalog

The **Welcome to BigQuery in the Cloud Console** message box opens. This message box provides a link to the quickstart guide and the release notes.

Click **Done**.

The BigQuery console opens.



Create a Dataset

1. Under the **Explorer** section, hover over on resource starting with *qwiklabs-gcp-*.
2. Click **View actions** icon and select **CREATE DATASET**.
3. Set *Dataset ID* to **bq_logs**.
4. Click **Create dataset**.

Click *Check my progress* to verify the objective. Create a Dataset

Run a query

First, run a simple query, which generates a log in Stackdriver. Later you use this log to easily setup the log export from Stackdriver to BigQuery.

1. Copy and paste the following query into the BigQuery Query editor:

```
SELECT current_date
```

2. Click **Run**.

Setup Log Export from Cloud Logging

1. In the Cloud Console, select **Navigation menu > Logging > Logs Explorer**.

Note: If prompted, Click LEAVE for Unsaved work.

2. In **Resource** drop-down, select **Bigquery** and click **Add**.

The screenshot shows the Google Cloud Logs Explorer interface. On the left, the 'Operations Logging' sidebar has 'Logs Explorer' selected. The main panel displays the 'Query builder' with a 'Resource' dropdown menu open. The dropdown menu shows 'RECENT' resources with 'BigQuery' selected, and 'ALL RESOURCE TYPES' including 'Audited Resource', 'BigQuery', 'BigQuery Dataset', 'BigQuery Project', and 'GCE Project'. At the bottom of the dropdown, a 'String Preview' section shows 'resource.type="bigquery..."' and an 'Add' button.

3. Now, click **Run Query**.

A few log entries from our query should appear.

Look for the entry that contains the word "jobcompleted".

> ⓘ	2020-12-09 11:38:43.994 IST	bigquery.googleapis.com	datasetservice.insert	projects/qwiklabs-gcp-00-ee69bed6bdbf/datasets	st
> ⓘ	2020-12-09 11:39:13.433 IST	bigquery.googleapis.com	jobservice.insert	projects/qwiklabs-gcp-00-ee69bed6bdbf/jobs	student-00
> ⓘ	2020-12-09 11:39:14.401 IST	bigquery.googleapis.com	jobservice.insert	projects/qwiklabs-gcp-00-ee69bed6bdbf/jobs	student-00
> ⓘ	2020-12-09 11:39:14.545 IST	bigquery.googleapis.com	jobservice.jobcompleted	projects/qwiklabs-gcp-00-ee69bed6bdbf/jobs/bquxjo	
> ⓘ	2020-12-09 11:39:15.659 IST	bigquery.googleapis.com	jobservice.getqueryresults	projects/qwiklabs-gcp-00-ee69bed6bdbf/queries/	
> ⓘ	2020-12-09 11:41:41.643 IST	bigquery.googleapis.com	datasetservice.update	projects/qwiklabs-gcp-00-ee69bed6bdbf/datasets/bq_1	

4. Click on the triangle on the left to open up the entry and then click on **Expand nested fields** on the right hand side.

> ⓘ	2020-12-09 11:38:43.994 IST	bigquery.googleapis.com	dataset.service.insert	projects/qwiklabs-gcp-00-ee69bed6bdf/datasets	st
> ⓘ	2020-12-09 11:39:13.433 IST	bigquery.googleapis.com	job.service.insert	projects/qwiklabs-gcp-00-ee69bed6bdf/jobs	student-00
> ⓘ	2020-12-09 11:39:14.401 IST	bigquery.googleapis.com	job.service.insert	projects/qwiklabs-gcp-00-ee69bed6bdf/jobs	student-00
> ⓘ	2020-12-09 11:39:14.545 IST	bigquery.googleapis.com	job.service.jobcompleted	projects/qwiklabs-gcp-00-ee69bed6bdf/jobs/bquxjc	
> ⓘ	2020-12-09 11:39:15.659 IST	bigquery.googleapis.com	job.service.getqueryresults	projects/qwiklabs-gcp-00-ee69bed6bdf/queries/	
> ⓘ	2020-12-09 11:41:41.643 IST	bigquery.googleapis.com	dataset.service.update	projects/qwiklabs-gcp-00-ee69bed6bdf/datasets/bq_1	

This shows the full JSON log entry, scroll down and have a look at the different fields.

- Then scroll back up to the header of the entry, click on 'jobcompleted' and choose **Show Matching Entries**.

> ⓘ	2020-12-09 11:38:43.994 IST	bigquery.googleapis.com	dataset.service.insert	projects/qwiklabs-gcp-00-ee69bed6bdf/datasets	st
> ⓘ	2020-12-09 11:39:13.433 IST	bigquery.googleapis.com	job.service.insert	projects/qwiklabs-gcp-00-ee69bed6bdf/jobs	student-00
> ⓘ	2020-12-09 11:39:14.401 IST	bigquery.googleapis.com	job.service.insert	projects/qwiklabs-gcp-00-ee69bed6bdf/jobs	student-00
> ⓘ	2020-12-09 11:39:14.545 IST	bigquery.googleapis.com	job.service.jobcompleted	projects/qwiklabs-gcp-00-ee69bed6bdf/jobs/bquxjc	
> ⓘ	2020-12-09 11:39:15.659 IST	bigquery.googleapis.com	job.service.getqueryresults	projects/qwiklabs-gcp-00-ee69bed6bdf/queries/	
> ⓘ	2020-12-09 11:41:41.643 IST	bigquery.googleapis.com	dataset.service.update	projects/qwiklabs-gcp-00-ee69bed6bdf/datasets/bq_1	

Showing logs for last 1 hour ending at 12/9/20, 11:58 AM. [Extend](#)

- Show matching entries
- Hide matching entries
- Show entries with matching substring

This sets up the search with the correct terms.

Logs Explorer
OPTIONS
REFINE SCOPE
Project

Query builder
Recent (3)
Saved (0)
Suggested (0)

Resource
Log name
Severity

```

1 resource.type="bigquery_resource"
2 protoPayload.methodName="job.service.jobcompleted"

```

Histogram

Query results

SEVERITY
TIMESTAMP
IST
SUMMARY

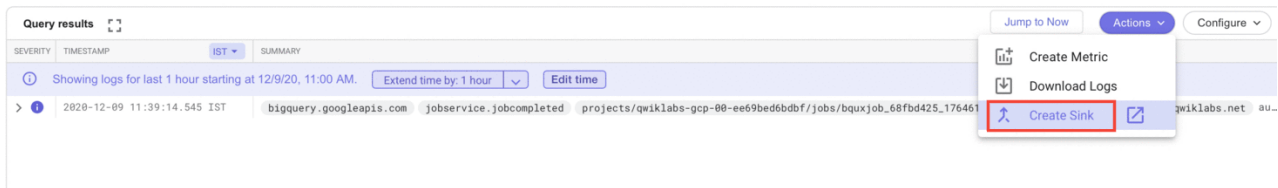
Showing logs for last 1 hour starting at 12/9/20, 11:26 AM.
Extend time by: 1 hour
Edit time

> ⓘ
2020-12-09 11:39:14.545 IST
bigquery.googleapis.com
job.service.jobcompleted
projects/qwiklab

Create Sink

Now you have the logs you need, it is easy to setup the sink.

1. Click **Create Sink** from the **Actions** drop-down.



2. Fill in the fields as follows:

- Sink name: **JobComplete** and click next.
- Select sink service: **BigQuery dataset**
- Select Bigquery dataset (Destination): **bq_logs** (The dataset you setup previously)

3. Click **Create Sink**.

4. Click **CLOSE**.

Any subsequent log entries from BigQuery are now exported to a table in the **bq_logs** dataset.

Click *Check my progress* to verify the objective. Create a Sink

Run example queries

To populate your new table with some logs, you need to run some example queries.

Navigate to **Cloud Shell**, then add each of the following BigQuery command line commands into Cloud Shell:

```
bq query --location=us --use_legacy_sql=false --use_cache=false \  
'SELECT fullName, AVG(CL.numberofYears) avgyears  
FROM `bigquery-samples.nested.persons_living`, UNNEST(citiesLived) as CL  
GROUP BY fullName'  
  
bq query --location=us --use_legacy_sql=false --use_cache=false \  
'select month, avg(mean_temp) as avgtemp from `bigquery-samples.weather_geo.gsod`  
where station_number = 947680  
and year = 2010  
group by month  
order by month'  
  
bq query --location=us --use_legacy_sql=false --use_cache=false \  
'select CONCAT(departure_airport, "-", arrival_airport) as route, count(*) as  
numberflights  
from `bigquery-samples.airline_ontime_data.airline_id_codes` ac,  
`bigquery-samples.airline_ontime_data.flights` fl  
where ac.code = fl.airline_code  
and regexp_contains(ac.airline , r"Alaska")  
group by 1  
order by 2 desc  
LIMIT 10'
```

You should see the results of each query returned.

Click *Check my progress* to verify the objective. Run example queries

Viewing the logs in BigQuery

1. Navigate back to BigQuery (**Navigation menu > BigQuery**).
2. Expand your resource starting with the name **qwiklabs-gcp-** and inspect your dataset **bq_logs**.

The name may vary, but you should see a "cloudataudit_googleapis_com_data_access_2019-06-19" table.

3. Inspect the schema of the table and note it has a very large number of fields.

If you tried to Preview and wondered why it doesn't show the logs for our recently run queries it is because the logs are streamed into the table, which means that the new data can be queried but won't show up in Preview for a little while.

To make the table more usable, you create a **VIEW**, which pulls out subset of fields and also perform some calculations to derive a metric for query time.

4. In the BigQuery Query editor, run the following SQL after replacing with the name of your project (the Project ID is easily copied from the left side of the lab page).

```

CREATE OR REPLACE VIEW
  bq_logs.v_querylogs AS
SELECT
  resource.labels.project_id,
  protopayload_auditlog.authenticationInfo.principalEmail,

  protopayload_auditlog.servicedata_v1_bigquery.jobCompletedEvent.job.jobConfiguratio

  protopayload_auditlog.servicedata_v1_bigquery.jobCompletedEvent.job.jobConfiguratio

  protopayload_auditlog.servicedata_v1_bigquery.jobCompletedEvent.job.jobStatus.error

  protopayload_auditlog.servicedata_v1_bigquery.jobCompletedEvent.job.jobStatistics.s

  protopayload_auditlog.servicedata_v1_bigquery.jobCompletedEvent.job.jobStatistics.e

  TIMESTAMP_DIFF(protopayload_auditlog.servicedata_v1_bigquery.jobCompletedEvent.job.

  protopayload_auditlog.servicedata_v1_bigquery.jobCompletedEvent.job.jobStatistics.s
    MILLISECOND)/1000 AS run_seconds,

  protopayload_auditlog.servicedata_v1_bigquery.jobCompletedEvent.job.jobStatistics.t

  protopayload_auditlog.servicedata_v1_bigquery.jobCompletedEvent.job.jobStatistics.t

  ARRAY(SELECT as STRUCT datasetid, tableId FROM
  UNNEST(protopayload_auditlog.servicedata_v1_bigquery.jobCompletedEvent.job.jobStati
    as tables_ref,

  protopayload_auditlog.servicedata_v1_bigquery.jobCompletedEvent.job.jobStatistics.t

  protopayload_auditlog.servicedata_v1_bigquery.jobCompletedEvent.job.jobStatistics.q

  severity
FROM
  `<YOUR-PROJECT-ID>.bq_logs.cloudaudit_googleapis_com_data_access_*`
ORDER BY
  startTime

```

Click *Check my progress* to verify the objective. Viewing the logs in BigQuery

5. Now query the VIEW. Clear the previous query and run the following command:

```
SELECT * FROM bq_logs.v_querylogs
```

6. Explore the results.

View the three queries that you executed in the previous step, similar to the image below.

Query results

Query complete (2.4 sec elapsed, 0 B processed)

Job information Results JSON Execution details

Row	project_id	principal_email	query	statement_type	message	startTime	endTime	run_seconds	totalProcessedBps	totalStats	datasetId	tableId	totalTableProcessed	queryLogFlowCount
1	qwklabs-gcp-2950cf8d...	gprtaging71855_student@qwklabs.net	select month, avg(mean_temp) as avgtemp from `bigquery-samples.weather_geo.good` where station_number = 847880 and year < 2010 group by month order by month	SELECT	null	2019-06-20 03:11:09.716 UTC	2019-06-20 03:11:12.481 UTC	2.775	3889392448	19950	weather_geo	good	1	1
2	qwklabs-gcp-2950cf8d...	gprtaging71855_student@qwklabs.net	select CONCAT(departure_airport, ',', arrival_airport) as route, count(*) as numberflights from `bigquery-samples.airline_ontime_data.airline_id_codes` as `bigquery-samples.airline_ontime_data.flights` # where ac_code = 'E' and regop_code and regop_contains(ac.airline, 'Alaska') group by 1 order by 2 desc LIMIT 10	SELECT	null	2019-06-20 03:11:20.256 UTC	2019-06-20 03:11:22.899 UTC	2.643	1200056843	59464	airline_ontime_data	airline_id_codes	2	1
3	qwklabs-gcp-2950cf8d...	gprtaging71855_student@qwklabs.net	select CONCAT(departure_airport, ',', arrival_airport) as route, count(*) as numberflights from `bigquery-samples.airline_ontime_data.airline_id_codes` as `bigquery-samples.airline_ontime_data.flights` # where ac_code = 'E' and regop_code and regop_contains(ac.airline, 'Alaska') group by 1 order by 2 desc LIMIT 10	SELECT	null	2019-06-20 03:11:20.256 UTC	2019-06-20 03:11:22.899 UTC	2.643	1200056843	59464	airline_ontime_data	flights	2	1

Congratulations!

This concluded the self-paced lab, Using BigQuery and Stackdriver to Analyze BigQuery Usage. You successfully exported BigQuery logs from Stackdriver and then returned to BigQuery for easy analysis with SQL.



Continue your Quest

This self-paced lab is part of the Qwiklabs Cloud Logging Quest. A Quest is a series of related labs that form a learning path. Completing this Quest earns you the badge above to recognize your achievement. You can make your badge public and link to them in your online resume or social media account. Enroll in this Quest and get immediate completion credit if you've taken this lab. See other available Qwiklabs Quests.

Take your next lab

Look at the quest to learn more about BigQuery and Stackdriver:

[Google Cloud's Operations Suite](#)

Next steps / learn more

- Lean more about [Cloud Logging](#)
- Get complete information about [BigQuery](#)

Google Cloud Training & Certification

...helps you make the most of Google Cloud technologies. Our classes include technical skills and best practices to help you get up to speed quickly and continue your learning journey. We offer fundamental to advanced level training, with on-demand, live, and virtual options to suit your busy schedule. Certifications help you validate and prove your skill and expertise in Google Cloud technologies.

Manual Last Updated: May 7, 2021

Lab Last Tested: May 7, 2021

Copyright 2021 Google LLC All rights reserved. Google and the Google logo are trademarks of Google LLC. All other company and product names may be trademarks of the respective companies with which they are associated.

Ready for more?

Here's another lab we think you'll like.

Lab

A Tour of Qwiklabs and Google Cloud
