Introduction

Part of working with billing data is working with data exported to BigQuery. In this lab, we are going to view sample billing exports maintained by Google, and conduct queries against a public dataset of billing exports. This will be a fun lab in that we can play around with SQL queries in BigQuery and see what kind of results we can get. Let's get started!

Once you are in your lab project, we will need to get into the BigQuery web console. From the top left menu, scroll down to Big Data, and select BigQuery.

Now that we are in BigQuery, let's look at the sample dataset we are going to work with. We are going to view all columns in our example table to see what fields are included. From the large Query Editor box, copy and paste the following query, then click the Run button:

SELECT \* FROM `cloud-training-prod-bucket.arch\_infra.billing\_data`

The field of cloud-training-prod-bucket.arch\_infra.billing\_data is the public dataset we are working with.

If we click the Results tab underneath, we can view the entire table we are going to work with. Feel free to experiment with other queries such as ordering by cost or usage amount by adding the below string to your query to sort by the column of your choice:

SELECT \* FROM `cloud-training-prod-bucket.arch\_infra.billing\_data` ORDER BY cost DESC

In this query, we are bringing up the entire table contents, but sorting by the highest cost first. You can experiment with other fields as well.

Let's now do some specific queries. In the same Query editor box, delete the existing contents, and enter the below query to find all charges that were more than 3 dollars:

SELECT product, resource\_type, start\_time, end\_time, cost, project\_id, project\_name, project\_labels\_key, currency, currency\_conversion\_rate, usage\_amount, usage\_unit FROM `cloud-training-prod-bucket.arch\_infra.billing\_data` WHERE (cost > 3)

Next let’s find which product had the highest total number of records:

SELECT product, COUNT(\*) FROM `cloud-training-prod-bucket.arch\_infra.billing\_data` GROUP BY product LIMIT 200

Looks like Pub/Sub is pretty popular here...

Finally, let’s see which product most frequently cost more than a dollar:

SELECT product, cost, COUNT(\*) FROM `cloud-training-prod-bucket.arch\_infra.billing\_data` WHERE (cost > 1) GROUP BY cost, product LIMIT 200

That will conclude this lab. You can quit anytime, however you are also free to experiment with other query types as well. The great thing about working with BigQuery is the flexibility of discovering what you can learn from your data!