Visualizing Billing Data with Google Data Studio

Introduction

Google Data Studio allows you to unlock the power of your data with interactive dashboards and beautiful reports that inspire smarter business decisions.

With Data Studio, you can:

Connect: easily access a wide variety of data. With built in and partner connectors, you can connect to virtually any type of data stream.
Visualize: turn your data into compelling stories of data visualization art. You can quickly build dashboards with Data Studio's web-based reporting tools.
Share: share your reports and dashboards with individuals, teams, or the world. Collaborate in real time. Embed your report on any page.

In this lab, you will learn how to build data visualizations with Google Data Studio. You will first explore a sample Google Cloud bill and learn how to export the billing data to BigQuery—Google's serverless, highly scalable enterprise data warehouse that is designed to make data analysts more productive with unmatched price-performance.

After running a few SQL queries on your billing data, you will export those metrics to Data Studio, where you will explore the service's chief features and build your own billing data visualizations.

Objectives

In this lab, you will learn how to:

- Use the billing service in the Cloud Console to explore projects and their consumption of cloud computing resources.
- Export billing data to BigQuery.
- Explore your billing data in BigQuery.

- Run SQL queries to better understand a project's consumption of Google Cloud services.
- Export your queried data to Data Studio.
- Explore Data Studio tools and generate visualizations with your queried data.

Task 1. Explore your billing data in BigQuery

The billing data was automatically exported into BigQuery when the lab spun up. In this section, you use an SQL query in BigQuery to see what information is available.

1- On the Navigation menu, scroll down and click BigQuery.

2- If prompted with a "Welcome to BigQuery in the Cloud Console" message, click Done. BigQuery opens.

3- Under Select a recent project, click the tile with your Project ID.

Note: Your Project ID is the left panel in your lab instructions.
The BigQuery console opens. In the left panel, under Explorer, you see your project. Your project name is your Project ID.

4- Type or paste the following SQL query into the Query editor to view data information:

```
SELECT * FROM `ctg-storage.bigquery_billing_export.gcp_billing_export_v1_01150A_B8F62B_47D999`
```

SELECT * returns all column values from a specified table.

5- Click Run to run the query.

You should receive similar results:

Query results SAVE RESULTS EXPLORE WITH DATA STUDIO

Query complete (0.0 sec elapsed, cached)

Job information | Results | JSON | Execution details

Row	billing_account_id	service.id	service.description	sku.id	sku.description
1	01150A-B8F62B-47D999	1F14-4801-0E16	Cloud Scheduler	A5D6-43A2-28A7	Jobs
2	01150A-B8F62B-47D999	95FF-2EF5-5EA1	Cloud Storage	1C42-50D6-1B1D	Coldline Data Retrieval
3	01150A-B8F62B-47D999	95FF-2EF5-5EA1	Cloud Storage	3953-E539-F17F	Multi-Regional Coldline Storage Asia (Early Delete)
4	01150A-B8F62B-47D999	95FF-2EF5-5EA1	Cloud Storage	2D3D-C91F-7E7C	Multi-Regional Coldline Storage Asia
5	01150A-B8F62B-47D999	95FF-2EF5-5EA1	Cloud Storage	9ADA-9AED-1B24	Class A Request Multi-Regional Storage
6	01150A-B8F62B-47D999	95FF-2EF5-5EA1	Cloud Storage	6B37-399C-BF69	GCP Storage egress between NA and APAC
7	01150A-B8F62B-47D999	F17B-412E-CB64	App Engine	8AE6-668C-FBCB	Task Queue Storage
8	01150A-B8F62B-47D999	95FF-2EF5-5EA1	Cloud Storage	0D5D-6E23-4250	Multi-Regional Storage US
9	01150A-B8F62B-47D999	95FF-2EF5-5EA1	Cloud Storage	A1F1-940B-6E1A	Class A Request Coldline Storage

Task 2. Run SQL queries in BigQuery and build data visualizations with Data Studio

You've explored a sample billing account in BigQuery that has thousands of rows of information. For this information to be useful you must be able to analyze the data to provide specific information. In BigQuery, you run SQL queries to answer questions to provide that specific information.

In this section, you ask two sample questions and use BigQuery to get that information. You then use Google Data Studio to build reports with data visualizations to share those insights.

Question 1: Which service types are most and least used

To find which service types are most and least used, you must determine:

- What types of services do the 4 projects use.
- Which service types are most and least used.

For the answers, run SQL queries on the billing data you have hosted in BigQuery.

Query to identify service types

1. Click **Compose New Query** to clear the **Query editor**.
2. Type or paste the following into the **Query editor** to find out which services the four projects consume.

```
SELECT service.description FROM
`ctg-storage.bigquery_billing_export.gcp_billing_export_v1_01150A_B8F62B_47D999` GROUP BY service.description
```

The **service.description** column tells you what Google Cloud service is associated with each log.
The **GROUP BY** keyword aggregates result-set rows that share common criteria (in this case the service description) and returns all of the unique entries found for such criteria.

3. Click Run.

The results show that the four projects use 8 different types of Google Cloud services.

Query to find which service types are most and least used

1. Click Compose New Query to clear the Query editor.

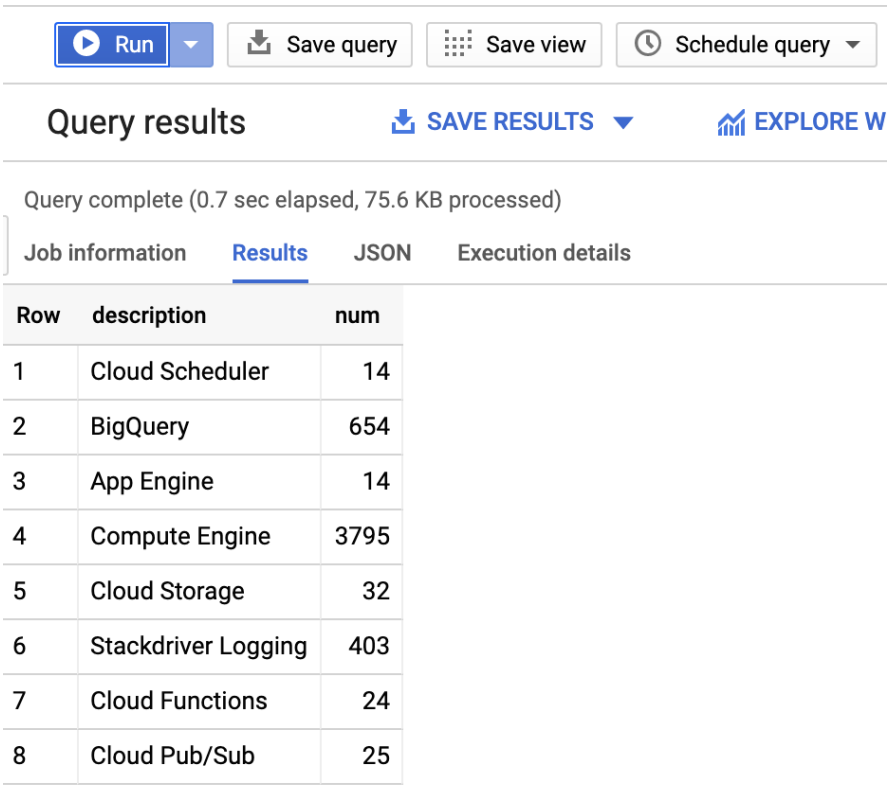
2. Type or paste in the following query into the Query editor to determine which service types are most and least used:

```
SELECT service.description, COUNT(*) AS num FROM
`ctg-storage.bigquery_billing_export.gcp_billing_export_v1_01150A_B8F62B_47D999`
GROUP BY service.description
```

The COUNT(*)function returns the number of rows that share the same criteria (in this case the service description).

3. Click Run.

Your results should look similar to the example table below, but your actual query output will be different.



Row	description	num
1	Cloud Scheduler	14
2	BigQuery	654
3	App Engine	14
4	Compute Engine	3795
5	Cloud Storage	32
6	Stackdriver Logging	403
7	Cloud Functions	24
8	Cloud Pub/Sub	25

In this example, and in terms of service logs, the projects use Compute Engine the most (3795 logs) and Cloud Scheduler and App Engine the least (14 logs). Remember, your query output will be different!

4. To the right of Query results, click Explore Data, and then Explore with Data Studio. Google Data Studio opens.

5. Step through the subsequent pages:

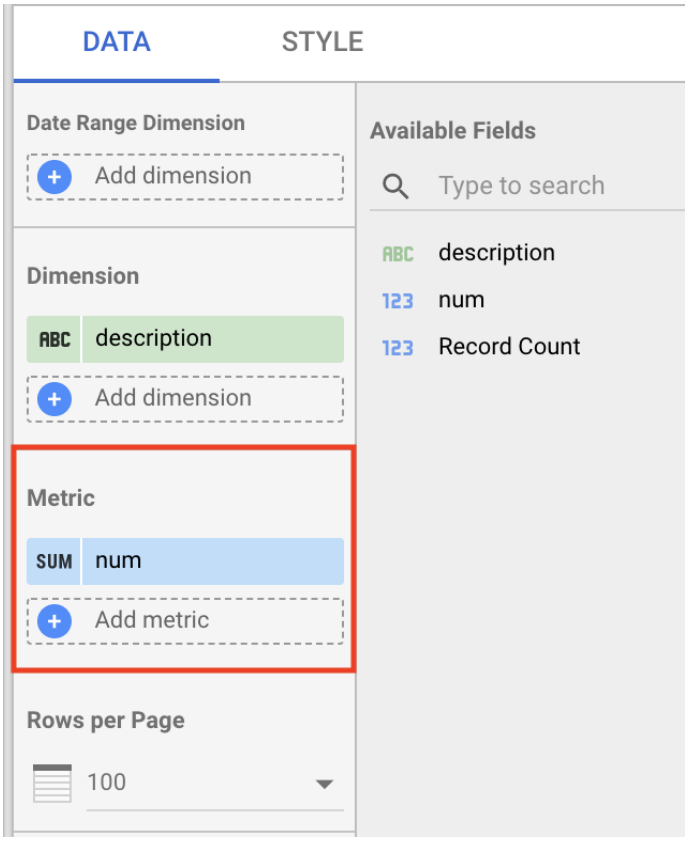
- Click Get Started.
- Click Authorize.

6. In the upper left, click Untitled Explorer and rename this report to Services Breakdown.

7. In the right panel, in the Data tab, find the Metric section. Hover over Record Count and click X to remove that metric.

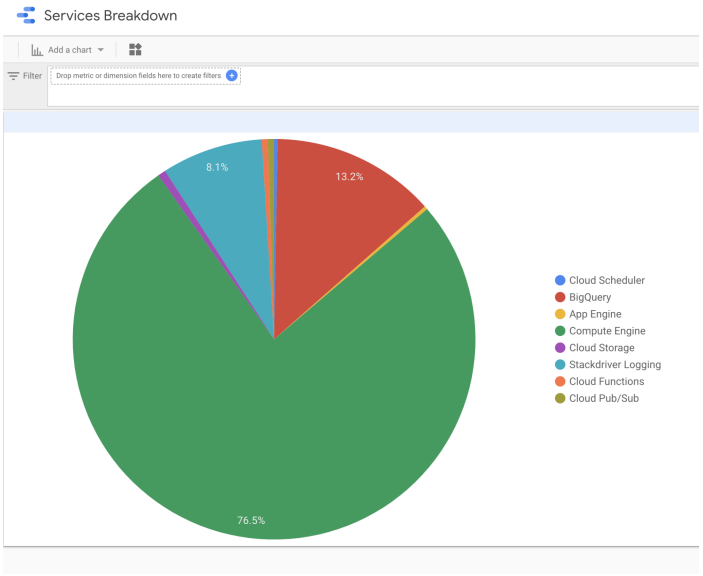
8. Click Add Metric and select num.

You may have to scroll down to see the num menu option.



9. In the right panel, in the Chart > Table section, select the Pie Chart table.

Data Studio generates a pie chart on the use of services:



10. Click the SAVE button in the top-right corner. If a "Welcome to Google Data Studio" popup appears, do the following for the subsequent dialogs:

- Click Get Started.
- Check the Acknowledge the terms and conditions box and click Accept.
- Check the No radio button for all four tips and recommendations.
- Click Continue.

11. Click SAVE once more.

Close this tab open and return to the BigQuery Console. You are now ready to answer the second question.

Question 2: Which regions are most and least used?

To find which regions are most and least used across all four projects, you must determine:

- In what regions the Google Cloud services run.
- Which regions are most and least used.

Query to determine the regions in which the Google Cloud services ran

1. Click Compose New Query to clear Query editor.

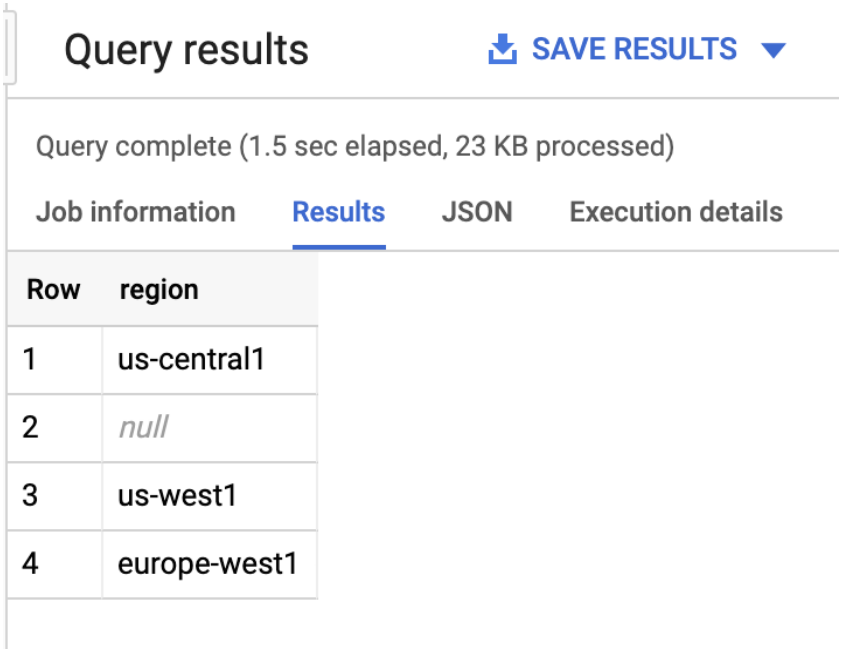
2. Type or paste the following SQL query:

```
SELECT location.region FROM
`ctg-storage.bigquery_billing_export.gcp_billing_export_v1_01150A_B8F62B_47D999`
GROUP BY location.region
```

The location.region column provides in what region the Google Cloud service runs.

3. Click Run.

You should see similar results. Remember, your results will be different:



Row	region
1	us-central1
2	*null*
3	us-west1
4	europe-west1

The example results above show that project resources ran in the us-central1, us-west1, and europe-west1 regions, or the region is null. A null region means the region is not known.

Query to determine which regions are used the most and least

1. In the BigQuery Console, click Compose New Query to clear the Query editor.

2. Type or paste in the following query into the Query editor:

```
SELECT location.region, COUNT(*) AS num FROM `ctg-storage.bigquery_billing_export.gcp_billing_export_v1_01150A_B8F62B_47D999` GROUP BY location.region
```

3. Click Run.

Your results should look similar to the example table below, but your actual query output will be different.

Row	region	num
1	us-central1	1269
2	*null*	1089
3	europe-west1	1298
4	us-west1	1305

The example results show that the projects use all regions somewhat evenly, and the "null" region is used a little less so. Remember, your results will be different.

Create a Data Studio visualization for regions

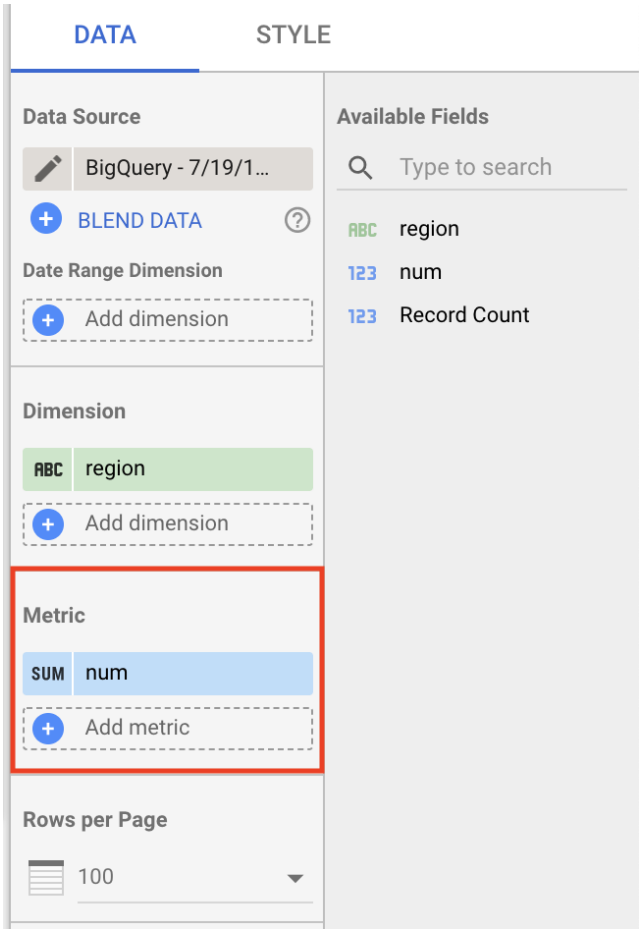
1. Still in the BigQuery console, to the right of Query results, click Explore Data, and then click Explore with Data Studio.
This takes you to the Data Studio Console.

2. Click in the "Untitled Explorer" field and rename it Regions Breakdown.

3. In the right panel, find the Metric field.

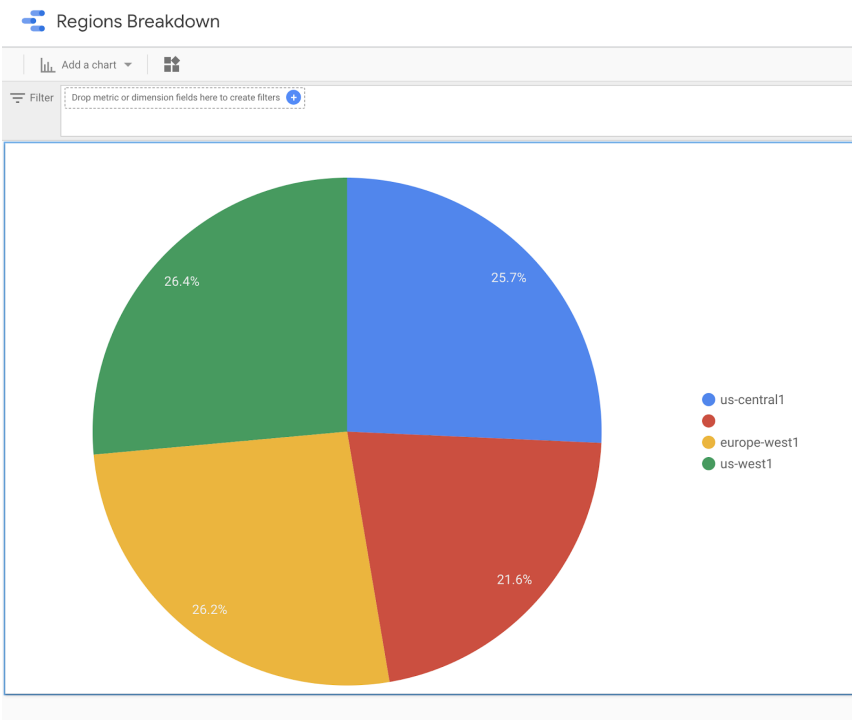
4. Hover over Record Count and click X to remove that metric.

5. Click Add Metric and select num.



6. In the right menu, in the Chart > Table section, select the Pie Chart table.

Data Studio generates a pie chart on the use of services:



7. Click SAVE in the top-right corner to save your visualization.

You have successfully created two data visualizations from your billing data queries.

8. To see the visualizations you just created, click the Data Studio icon to the right of Regions Breakdown:

The Data Studio console opens listing the 2 visualizations you created.