

# Hands-On Activity: Introduction to BigQuery

TOTAL POINTS 2

---

1.



## Activity overview

By now, you have been introduced to BigQuery, a data warehouse on Google Cloud that data analysts can use to query, filter large datasets, aggregate results, and perform complex operations. In this activity, you will explore the BigQuery interface; upload public data to your console; and write some simple SQL queries using SELECT, FROM, and WHERE.

By the time you complete this activity, you will be more familiar with writing queries in the BigQuery interface. This will enable you to practice SQL, which is important for working with databases in your career as a data analyst.

## Explore BigQuery

For this activity, you will need a BigQuery account. If you haven't made one already, you can follow the instructions from the [Using BigQuery reading](#). Once you have your account, you can start exploring!

### Open your console

1. Log in to [BigQuery](#).

2. Then, click the **Go to console** button on the BigQuery homepage. This will open a new tab with your console.

## BigQuery

### Benefits

#### Key features

#### Customers

#### What's new

## Documentation

## Use cases

Migrating data warehouses to BigQuery

Predictive analytics

Bring any data into BigQuery

## All features

# BigQuery

Serverless, highly scalable, and cost-effective multicloud data warehouse designed for business agility.

New customers get \$300 in free credits to spend on Google Cloud during the first 90 days. All customers get 10 GB storage and up to 1 TB queries/month, completely free of charge.

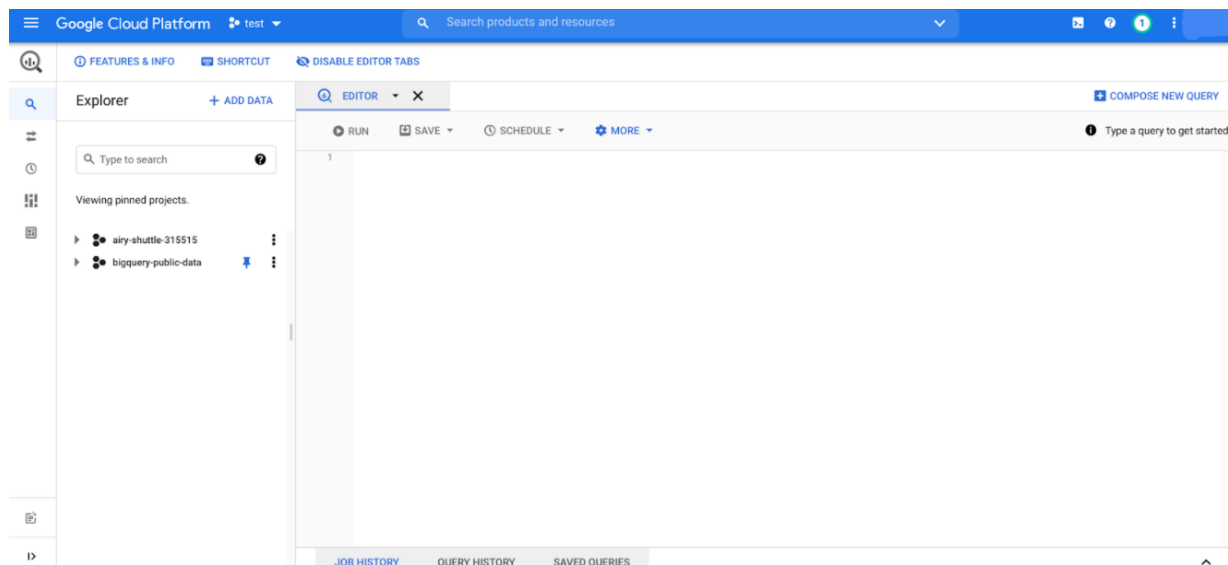
[Go to console](#)[Contact sales](#)

✓ Democratize insights with a secure and scalable platform with built-in machine learning

✓ Power business decisions from data across clouds with a flexible, multicloud analytics solution

✓ Run analytics at scale with 26%–34% lower three-year TCO than [cloud](#)

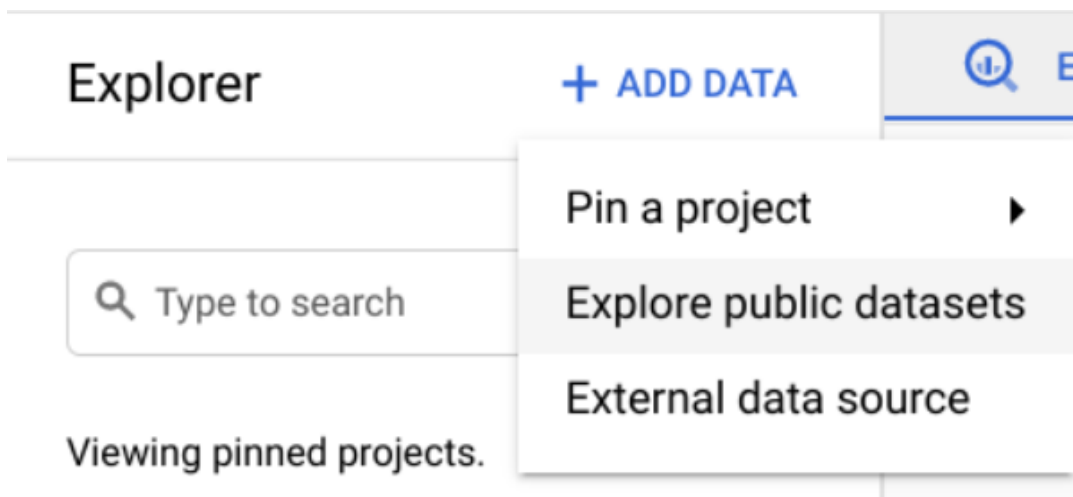
3. Take a moment to explore your console. On the left side, you will find the **Explorer** menu; this includes a search bar you can use to find resources, pinned projects, and the **+ ADD DATA** button. On the right side, you will find the **Query Editor**. This is where you will input queries and view datasets. You can also find your Job History, Query History, and Saved Queries here.



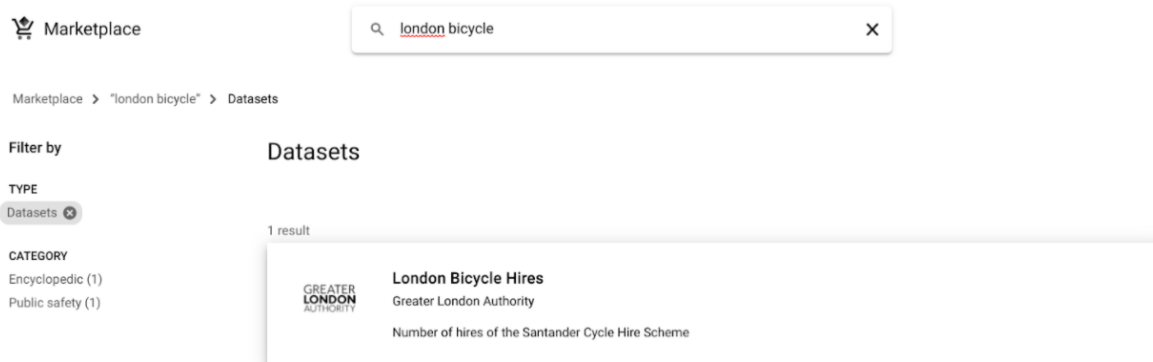
## Access public data in BigQuery

In order to actually start writing queries, you will need some data to work with. Once you're familiar with the BigQuery interface, you can access a public dataset directly from your console.

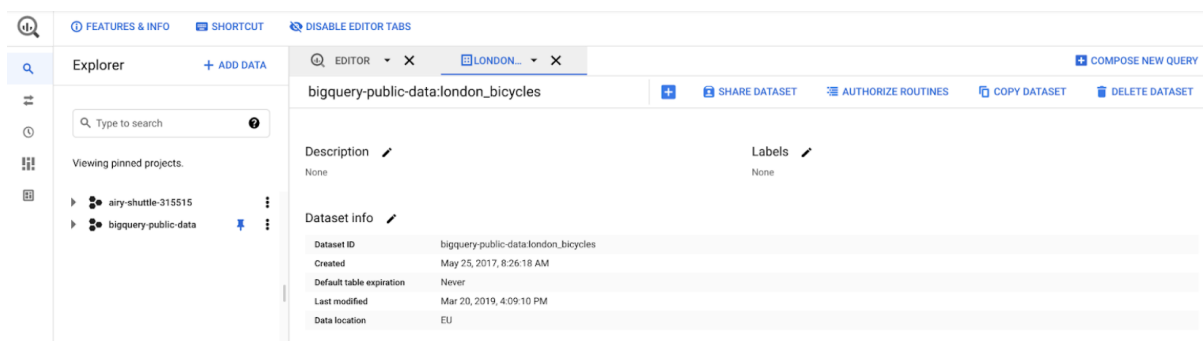
1. Click on the **+ ADD DATA** button in the **Explorer menu pane** and select **Explore public datasets**. This will open a new menu where you can search public datasets that are already available through Google Cloud.



2. In the dataset menu you just opened, **type *london bicycle*** in the search box at the top; this will return the London Bicycle Hires dataset from the Greater London Authority. **Click the dataset** for more details.



3. From the dataset information page, **click the blue VIEW DATASET** button. This will open your console in a new tab with this dataset loaded.



You'll notice that **bigquery-public-data** is now pinned in your Explorer pane. You can now explore and query these public datasets.

4. Click on **the arrow** next to **bigquery-public-data** and scroll down the list of public datasets until you find the **london\_bicycles** data. When you click on the dataset, it will list two tables. Click on **cycle\_hire**.



Viewing pinned projects.

▶		hud_zipcode_crosswalk	⋮
▶		human_genome_variants	⋮
▶		human_variant_annotation	⋮
▶		immune_epitope_db	⋮
▶		iowa_liquor_sales	⋮
▶		irs_990	⋮
▶		labeled_patents	⋮
▶		libraries_io	⋮
▼		london_bicycles	⋮
		cycle_hire	⋮
		cycle_st	⋮
		cycle_hire	⋮
▶		london_crime	⋮
▶		london_fire_brigade	⋮
▶		medicare	⋮
▶		ml_datasets	⋮

This will pull up a new tab in your Query Editor with information about the table schema.

5. After checking out the table schema, you can take a peek into what data the cycle\_hire table contains by clicking on the **Preview** tab. This will give you a better idea of what kind of data you'll be working with.

EDITOR

×

LONDON...

×

CYCLE\_H...

×

COMPOSE NEW QUERY

cycle\_hire

QUERY

SHARE

COPY

DELETE

LOAD DATA

EXPORT

SCHEMA

DETAILS

PREVIEW

Row	rental_id	duration	bike_id	end_date	end_station_id	end_station_name	start_date	start_station_id	start_station_name
1	47469109	3180	7054	2015-09-03 12:45:00 UTC	111	Park Lane , Hyde Park	2015-09-03 11:52:00 UTC	300	Serpentine Car Park, Hyde Park
2	46915469	7380	3792	2015-08-16 11:59:00 UTC	407	Speakers' Corner 1, Hyde Park	2015-08-16 09:56:00 UTC	407	Speakers' Corner 1, Hyde Park
3	65899423	2040	3038	2017-06-09 18:30:00 UTC	165	Orsett Terrace, Bayswater	2017-06-09 17:56:00 UTC	579	Queen Street 2, Bank
4	64280726	2280	10868	2017-04-22 10:14:00 UTC	553	Regent's Row , Haggerston	2017-04-22 09:36:00 UTC	519	Teviot Street, Poplar
5	59235489	2340	7183	2016-10-09 04:31:00 UTC	100	Albert Embankment, Vauxhall	2016-10-09 03:52:00 UTC	612	Wandsworth Rd, Isley
6	55248935	2160	7619	2016-06-26 07:26:00 UTC	465	Pitfield Street North,Hoxton	2016-06-26 06:50:00 UTC	459	Gunmakers Lane, Old
7	43015438	5400	2779	2015-04-27 22:50:00 UTC	772	Binfield Road, Stockwell	2015-04-27 21:20:00 UTC	157	Wright's Lane, Kensin
8	61090882	1980	1710	2016-12-15 21:16:00 UTC	487	Canton Street, Poplar	2016-12-15 20:43:00 UTC	564	Somerset House, Stra
9	47150505	2400	6748	2015-08-22 17:15:00 UTC	291	Claverton Street, Pimlico	2015-08-22 16:35:00 UTC	248	Triangle Car Park, Hyc
10	54154089	3000	12290	2016-05-23 22:03:00 UTC	785	Aquatic Centre, Queen Elizabeth Olympic Park	2016-05-23 21:13:00 UTC	785	Aquatic Centre, Queer
11	52607737	6540	11430	2016-04-03 16:39:00 UTC	304	Cumberland Gate, Hyde Park	2016-04-03 14:50:00 UTC	382	Farm Street, Mayfair
12	49135790	1980	9338	2015-10-27 14:29:00 UTC	209	Denyer Street, Knightsbridge	2015-10-27 13:56:00 UTC	127	Wood Street, Guildhal
13	50251732	2580	9160	2015-12-12 13:12:00 UTC	229	Whitehall Place, Strand	2015-12-12 12:29:00 UTC	421	Southwark Station 2, !

Rows per page: 100

1 - 100 of 24369201

First page <

<

>

> Last page

JOB HISTORY

QUERY HISTORY

SAVED QUERIES

Once you have finished previewing the data, you can write a query!

## Query your data

So far, you've learned three basic parts of a query: **SELECT**, **FROM**, and **WHERE**. As a refresher, here are what those basic parts represent in the query:

- **SELECT** is the section of a query that indicates what data you want SQL to return to you
- **FROM** is the section of a query that indicates which table the desired data comes from.
- **WHERE** is the section of a query that indicates any filters you'd like to apply to your dataset

### Write a basic query

Now, construct a simple command using the basic parts of a query you have already learned! For example, you can select a specific column from the `cycle_hire` table, such as the `end_station_name` column.

1. Start your query with a **SELECT clause** and indicate which column you want to select from the table; in this case, you'll input `end_station_name`.
2. After you have indicated which column you are selecting, write your **FROM clause**. You will need to specify the table you are querying from by inputting the following location: ``bigquery-public-data.london_bicycles.cycle_hire``;

The completed query should appear like this:

```
SELECT
  end_station_name
FROM
```

```
`bigquery-public-data.london_bicycles.cycle_hire`;
```

3. Run your completed query by clicking on the blue **RUN** button.

This query may take a few seconds to execute. Once it has finished, you will find the list of station names you requested under the **Query Results** console pane.

Write a query to answer a question

After running the first basic query, try answering a specific question about the data. For example, what were the rental\_ids for all bike trips that lasted for 20 minutes or longer?

1. Click **COMPOSE NEW QUERY** to start a new query. Start with your **SELECT statement** again. This time, you will want to include all of the columns in the table for this query. You can use an asterisk to indicate that you are selecting all of the data like this: `SELECT COUNT(*) AS num_of_trips`
2. Then you will add your **FROM statement**. You will be using the same dataset as the previous query: `FROM `bigquery-public-data.london_bicycles.cycle_hire``.
3. Finally, you'll add a **WHERE** statement to specify that you want to filter for only bike rides 20 minutes or longer. If you check the preview of this data, you might notice that the duration is recorded in seconds, so you'll specify 1200 seconds in your query. You can write that as `WHERE duration >= 1200`;

Your completed query should be written like this:

**SELECT**

```
COUNT(*) AS num_of_trips
```





**FROM**

```
`bigquery-public-data.london_bicycles.cycle_hire`
```

**WHERE**

```
duration >= 1200;
```


4. Run your completed query by clicking on the blue **RUN** button.



```
1  SELECT
2    COUNT(*) AS num_of_biketrips
3  FROM
4    `bigquery-public-data.london_bicycles.cycle_hire`
5  WHERE
6    duration >= 1200;
```

This query may take a few seconds to execute. Once it has finished, you will find a list of rides from this table that fit your criteria. There are more than 7 million rows with bike trips that are 20 minutes or longer!

### Up for a challenge?

 If you're comfortable using queries to answer questions, try creating and running queries to answer any of the questions below:

- What are the names of the stations that bike\_id 1710 started from?
- How many bike\_ids have ended at "Moor Street, Soho"?
- What is the station\_id for "Canton Street, Poplar"?
- What is the name of the station whose ID is 111?
- How many distinct bike\_ids had trip durations greater than 2400 seconds (or 40 minutes)?

You can use the solutions doc to check your work: [Intro to BigQuery Solutions](#)

Or download the file directly here:

Intro to BigQuery solutions.docx