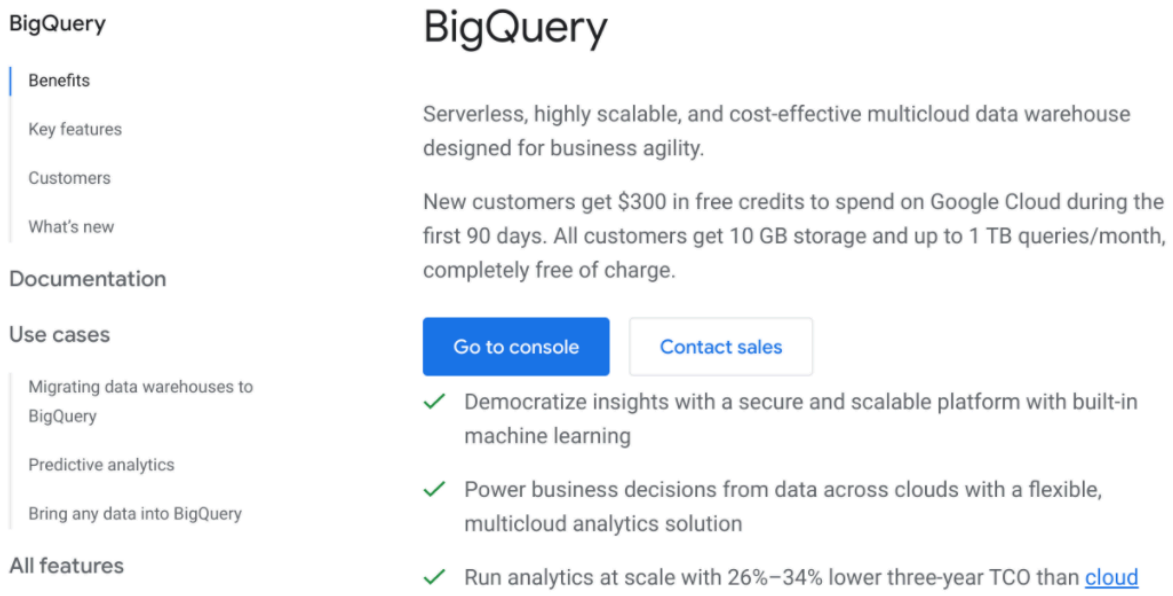
## Activity overview

You have recently been introduced to BigQuery, a data warehouse on Google Cloud that data analysts 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`**.

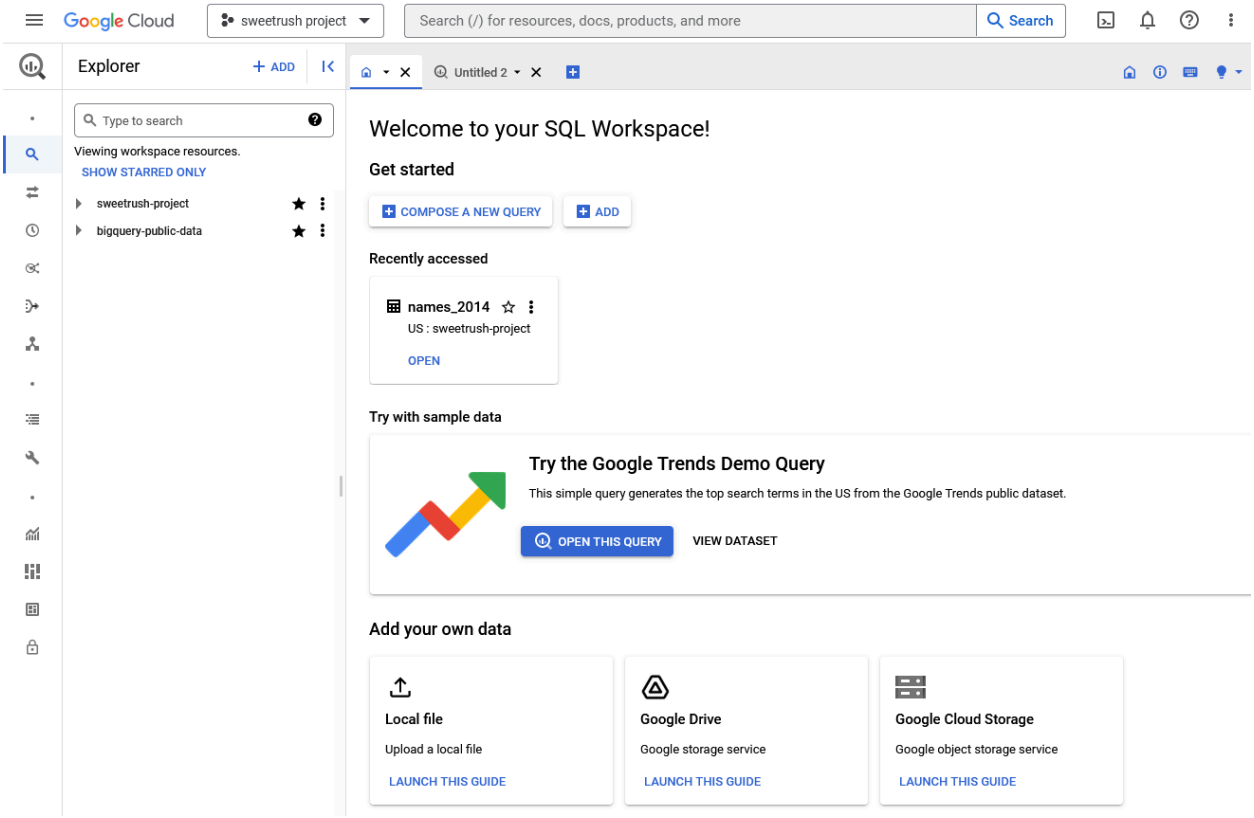
By the time you complete this activity, you will be more familiar with writing queries in the BigQuery interface. This will enable you to practice SQL, which is important for working with databases in your career as a data analyst.

# Step 2: Open your BigQuery console

1. Log in to BigQuery.
2. Select the Go to console button on the BigQuery homepage. This will open a new tab with your console.



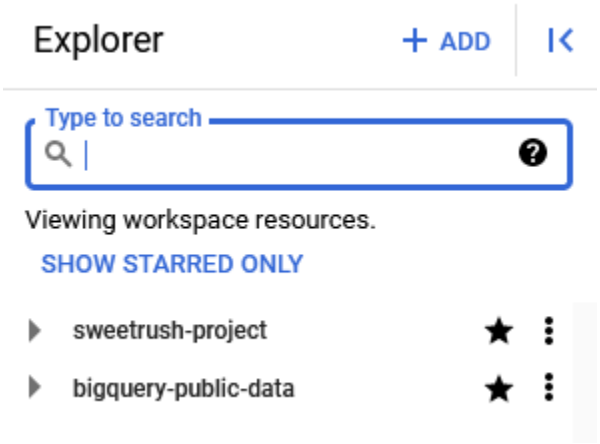

3. Take a moment to explore your console. The Explorer menu includes a search bar you can use to find resources, pinned projects, and the + ADD button for adding data. The Editor welcome page is where you will navigate to a query editor, try sample data, add local data files, add Google cloud storage, or add other external connections. You can also find your job history, query history, and saved queries here.



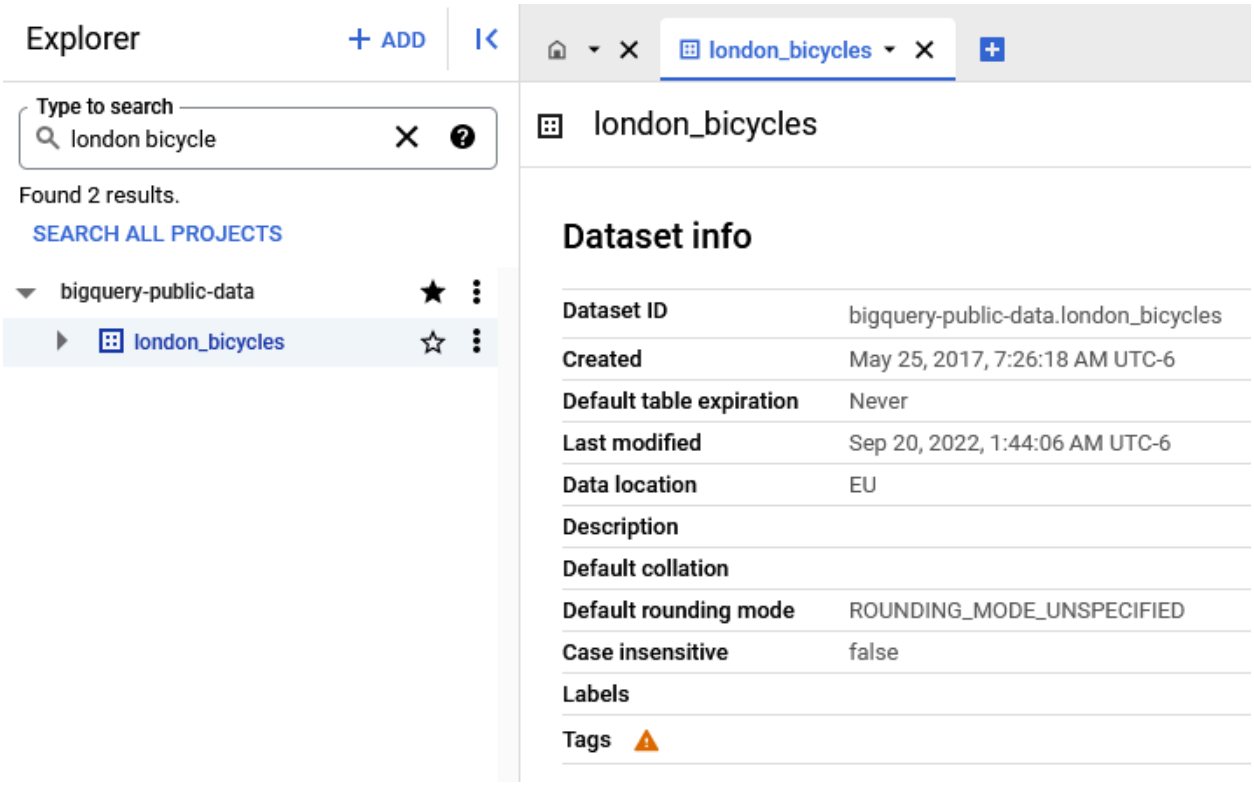

# Step 3: Access public data in BigQuery

In order to 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.

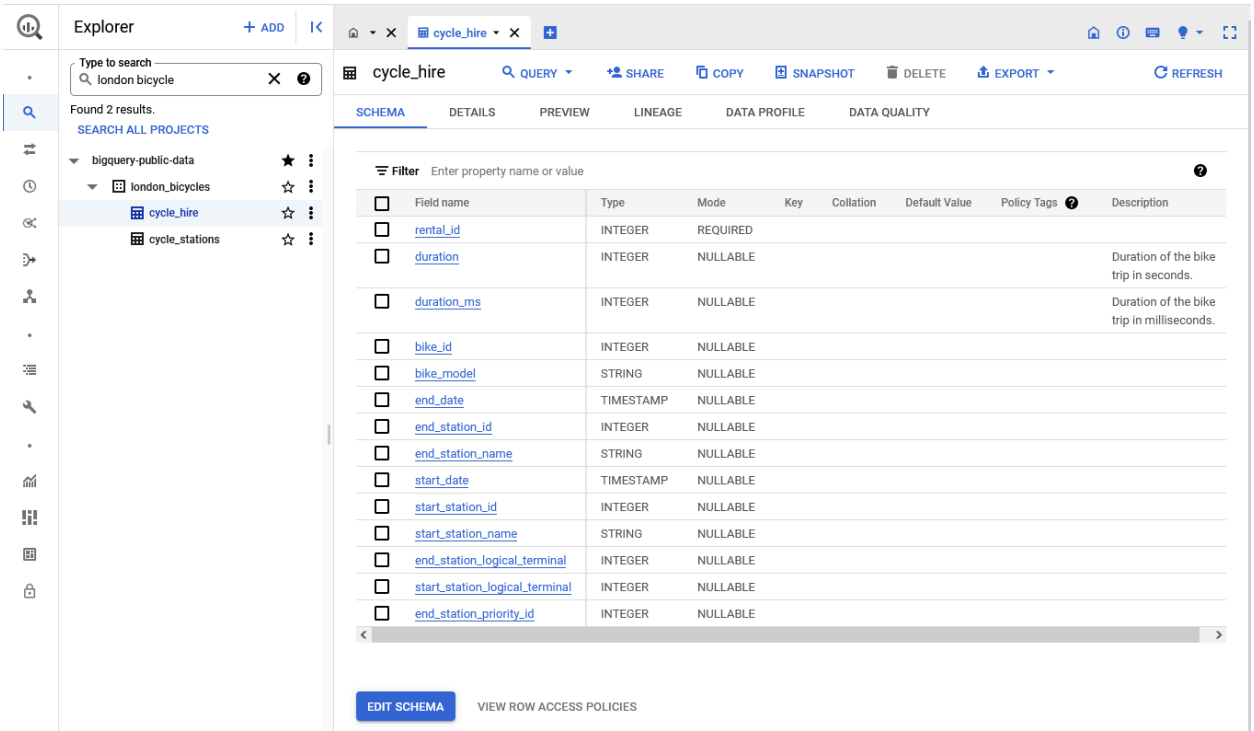
1. Select the search bar in the Explorer pane.




2. Enter "london bicycle" in the search box and press enter; this will return the `london_bicycles` database from the Greater London Authority. Select the database for more details. If you cannot find it, make sure you're searching in all projects. The `london_bicycles` database is in the `bigquery-public-data` project.



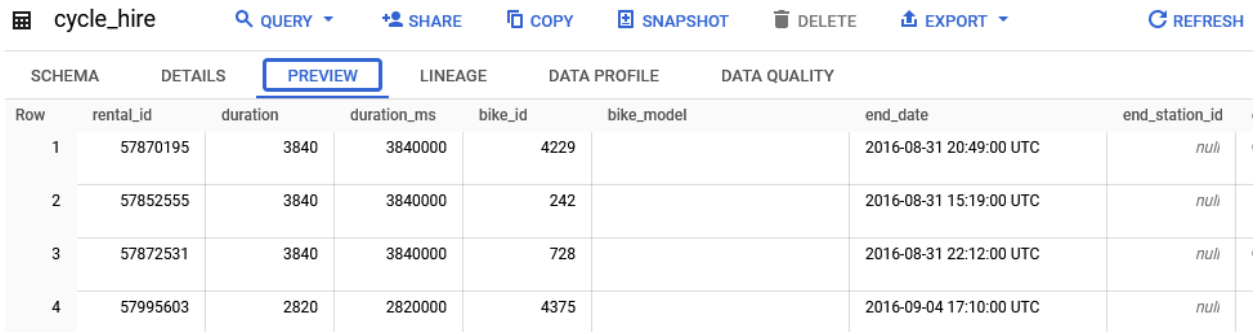

3. Select the arrow to the left of the **`london_bicycles`** database name. This expands the dataset to reveal two table names: **`cycle_hire`** and **`cycle_stations`**. Select the **`cycle_hire`** table name within the Explorer pane.




Screenshot of BigQuery. The titles cycle_hire and cycle_stations are shown beneath london_bicycles in the explorer pane. Cycle_hire is selected.

This will pull the **`cycle_hire`** schema into the console. Take a moment to explore the field names and the associated information.

4. Now, select the PREVIEW tab to find a sample of the data that you'll be working with.




| Row | rental_id | duration | duration_ms | bike_id | bike_model | end_date | end_station_id |
|---|---|---|---|---|---|---|---|
| 1 | 57870195 | 3840 | 3840000 | 4229 | | 2016-08-31 20:49:00 UTC | null |
| 2 | 57852555 | 3840 | 3840000 | 242 | | 2016-08-31 15:19:00 UTC | null |
| 3 | 57872531 | 3840 | 3840000 | 728 | | 2016-08-31 22:12:00 UTC | null |
| 4 | 57995603 | 2820 | 2820000 | 4375 | | 2016-09-04 17:10:00 UTC | null |

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

# Step 4: Review basic parts of a query

So far, you've learned three basic parts of a query: **`SELECT`**, **`FROM`**, and **`WHERE`**. As a refresher:

- **`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. You must provide a full path to the table. The path includes the project name, database name, and table name, each separated by a period.
- **`WHERE`** is the section of a query that indicates any filters you'd like to apply to your table.

# Step 5: 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. Select the Blue + button or QUERY - In new tab to start a new query.
2. 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`**.
3. After you have indicated which column you are selecting, write your **`FROM`** clause. 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:

```sql
SELECT
    end_station_name
FROM
    `bigquery-public-data.london_bicycles.cycle_hire`;
```

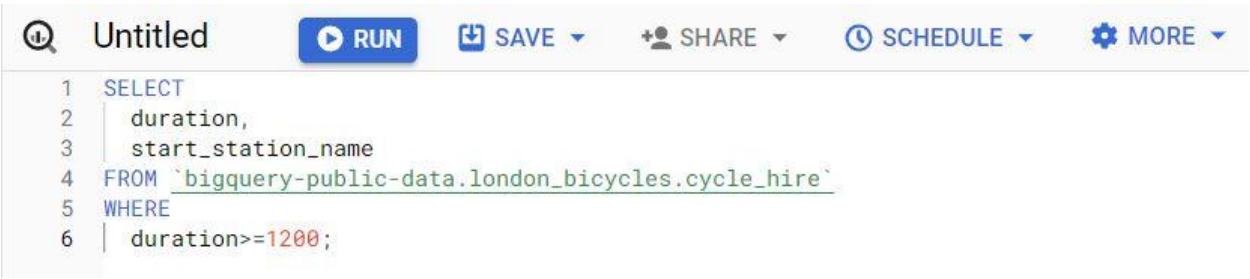
4. Run your completed query by selecting the blue RUN button.
This query may take a few seconds to execute. Once it has finished, you will find the list of station names you requested under the Query Results console pane.

## Step 6: Write a query to answer a question

After running the first basic query, try answering a specific question about the data. For example, how many bike trips lasted for 20 minutes or longer?
1. Select the Blue + button or QUERY - In new tab to start a new query. Start with your **`SELECT`** statement again. This time, include the two columns **`duration`** and **`start_station_name`** in the query. The data in these columns will tell where the trip started and the duration of the trip. Be sure to separate each column name with a comma.
2. Next, add your **`FROM`** statement. You will be using the same table as the previous query: **`FROM `bigquery-public-data.london_bicycles.cycle_hire`;`**. Note: The backticks around the table in this line of code are optional.
3. Finally, 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. Write that as **`WHERE duration >= 1200;`**
Your completed query should be written like this:

```sql
SELECT
    duration,
    start_station_name
FROM
    `bigquery-public-data.london_bicycles.cycle_hire`
WHERE
    duration >= 1200;
```

4. Run your completed query by clicking the RUN button.




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 millions of rows with bike trips that are 20 minutes or longer!

# Optional Step 7: Up for a challenge?

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

- What is the name of the station whose **`start_station_id`** is 111?
- Return all the **`rental_id`**s, station IDs, and station names that **`bike_id`** 1710 started from.
- What is the **`bike_model`** of **`bike_id`** 58782?

## Reflection

Run another query on your table:

```
SELECT
    end_station_name
FROM
    `bigquery-public-data.london_bicycles.cycle_hire`
WHERE
    rental_id = 57635395;
```

- Southwark Street, Bankside
- Notting Hill Gate Station, Notting Hill
- East Village, Queen Elizabeth Olympic Park
- Tower Gardens, Tower

In this activity, you had an opportunity to get more familiar with BigQuery and writing SQL queries. In the text box below, write 2-3 sentences (40-60 words) in response to each of the following questions:

- How do you think you can use public datasets in BigQuery to help develop your data analysis skills?
- How do you think understanding basic query syntax will help you write more complex queries in the future?

Congratulations on completing this hands-on activity! You explored BigQuery, uploaded public data to your console, and constructed some queries. An effective response would note that BigQuery public datasets help you practice writing SQL. In upcoming activities, you will continue exploring working with databases and writing queries with SQL—an essential tool in every data analyst’s toolkit.