BigQuery: Qwik Start - Console

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.

1. Click **Activate Cloud Shell** ![icon] at the top of the Google Cloud console.

2. Click through the following windows:

   - Continue through the Cloud Shell information window.
   - Authorize Cloud Shell to use your credentials to make Google Cloud API calls.

When you are connected, you are already authenticated, and the project is set to your **Project_ID**, `qwiklabs-gcp-04-cfee0b383173`. The output contains a line that declares the **Project_ID** for this session:

```
Your Cloud Platform project in this session is set to qwiklabs-gcp-04-cfee0b383173
```

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

3. (Optional) You can list the active account name with this command:
```
gcloud auth list
```

4.
5.
6. Click **Authorize**.

Output:

```
ACTIVE: *
ACCOUNT: student-01-86a32db08d08@qwiklabs.net

To set the active account, run:
    $ gcloud config set account `ACCOUNT`
```

5. (Optional) You can list the project ID with this command:
```
gcloud config list project
```

Output:

```
[core]
project = qwiklabs-gcp-04-cfee0b383173
```
Note: For full documentation of `gcloud`, in Google Cloud, refer to the gcloud CLI overview guide.

Task 1. Open BigQuery

The BigQuery console provides an interface to query tables, including public datasets offered by BigQuery. The query you will run accesses a table from a public dataset that BigQuery provides. It uses standard query language to search the dataset, and limits the results returned to 10.

Open the BigQuery console

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

2. Click **Done**.
The BigQuery console opens.

Task 2. Query a public dataset

1. Click + (SQL query) to create a new query. Copy and paste the following query into the BigQuery Query editor:
```
#standardSQL
SELECT
```

```
  weight_pounds, state, year, gestation_weeks
FROM
 `bigquery-public-data.samples.natality`
ORDER BY weight_pounds DESC LIMIT 10;
```

This data sample holds information about US natality (birth rates).

A green or red check displays depending on whether the query is valid or invalid. If the query is valid, the validator also describes the amount of data to be processed after you run the query.

This information helps determine the cost to run a query.

2. Click the **Run** button.

Your query results should resemble the following:

Query results   ⬇ SAVE RESULTS ▾

Query complete (2.626 sec elapsed, 3.49 GB processed)

Job information   **Results**   JSON   Execution details

Row	weight_pounds	state	year	gestation_weeks
1	18.0007436923	KY	2004	47
2	18.0007436923	OR	1972	40
3	18.0007436923	null	2007	39
4	18.0007436923	null	2008	null
5	18.0007436923	TX	1969	null
6	18.0007436923	null	2005	40
7	18.0007436923	null	2007	45
8	18.0007436923	null	2005	null
9	18.0007436923	GA	1979	34
10	18.0007436923	null	2007	38

Task 3. Create a new dataset

To load custom data into a table, you first need to create a BigQuery dataset.
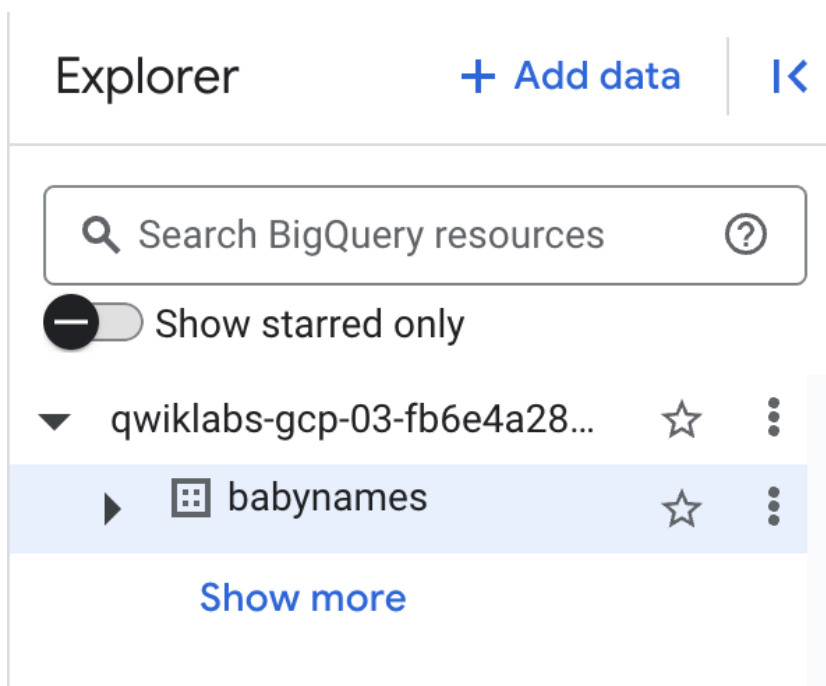
Datasets help control access to tables and views in a project. This lab uses only one table, but you still need a dataset to hold the table.

1. In the **Explorer** pane, near your project id, click on **View actions** (⋮) then click **Create dataset**.
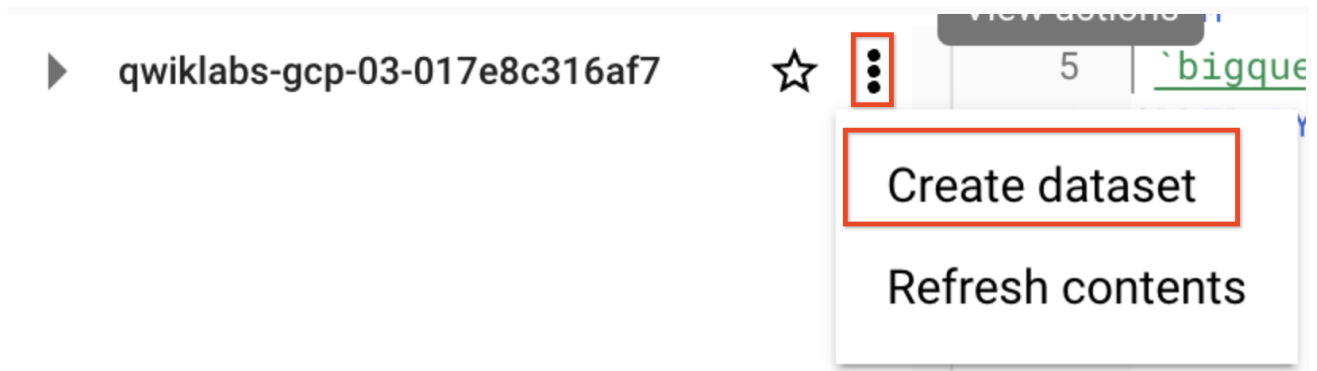
2. Set **Dataset ID** to **babynames**.

3. Leave all other fields at their default settings. Click **Create dataset**.

Now you have a dataset.

Task 4. Load data into a new table

Next you create a table inside the **babynames** dataset, then load the data file from your storage bucket into the new table.

The custom data file you'll use contains approximately 7 MB of data about popular baby names, provided by the US Social Security Administration.

1. In the Cloud Console, select **Navigation menu** > **BigQuery** to return to the BigQuery console.

2. Navigate to the **babynames** dataset, by clicking **View actions** (⋮) near your dataset then click **Create table**.

3. In the Create table dialog, set the following fields, leave all others at the default value:

Field	Value
Create table from	**Google Cloud Storage**
Select file from GCS bucket	`spls/gsp072/baby-names/yob2014.txt`
File format	**CSV**
Table	`names_2014`

Schema > **Edit as text**	Slide on, then add the following in the textbox: `name:string,gender:string,count:integer`

4. Click the **Create table** button.

When BigQuery is finished creating the table and loading the data, you see the `names_2014` table under the `babynames` dataset.

Task 5. Preview the table

Check your table! View the first few rows of the data.

* Click the `names_2014` table in the left panel, then click **Preview**.

Your table is ready for queries.

Task 6. Query a custom dataset

Running a query against custom data is identical to the [querying a public dataset](#) that you did earlier, except that now you're querying your own table instead of a public table.

1. In BigQuery, click the + (SQL query) icon at the top.

2. Paste or type the following query into the query **Editor**.

Note: If your table name is something other than **babynames**, update the code with your table name.

```
#standardSQL
SELECT
 name, count
FROM
 `babynames.names_2014`
WHERE
 gender = 'M'
ORDER BY count DESC LIMIT 5;
```

3. Click the **Run** button. The query displays the top 5 boys names for the year of data (2014) you loaded into the table.