BigQuery: Qwik Start - Console

GSP072

Overview

Storing and querying massive datasets can be time consuming and expensive without the right hardware and infrastructure. BigQuery is an enterprise data warehouse that solves this problem by enabling super-fast SQL queries using the processing power of Google's infrastructure. Simply move your data into BigQuery and let us handle the hard work. You can control access to both the project and your data based on your business needs, such as giving others the ability to view or query your data.
You can access BigQuery in the Console, the classic Web UI or a command-line tool, or by making calls to the BigQuery REST API using a variety of client libraries such as Java, .NET, or Python. There are also a variety of third-party tools that you can use to interact with BigQuery, such as visualizing the data or loading the data.
This hands-on lab shows you how to use the Web UI to query public tables and load sample data into BigQuery.

Setup and Requirements

Qwiklabs setup

Before you click the Start Lab button

Read these instructions. Labs are timed and you cannot pause them. The timer, which starts when you click **Start Lab**, shows how long Google Cloud resources will be made available to you.

This Qwiklabs hands-on lab lets you do the lab activities yourself in a real cloud environment, not in a simulation or demo environment. It does so by giving you new, temporary credentials that you use to sign in and access Google Cloud for the duration of the lab.

What you need

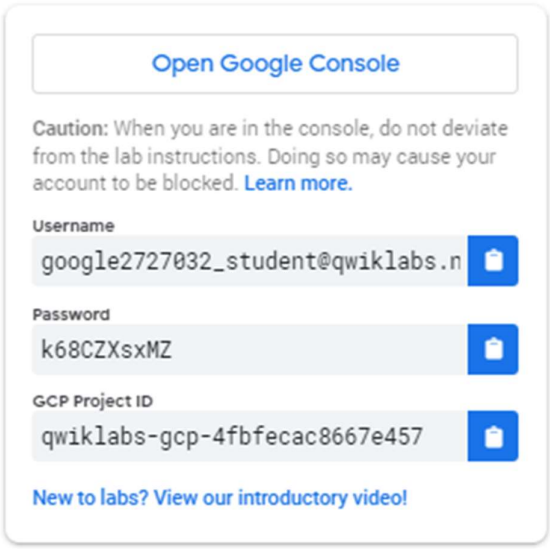
To complete this lab, you need:

- Access to a standard internet browser (Chrome browser recommended).
- Time to complete the lab.
  Note: If you already have your own personal Google Cloud account or project, do not use it for this lab.

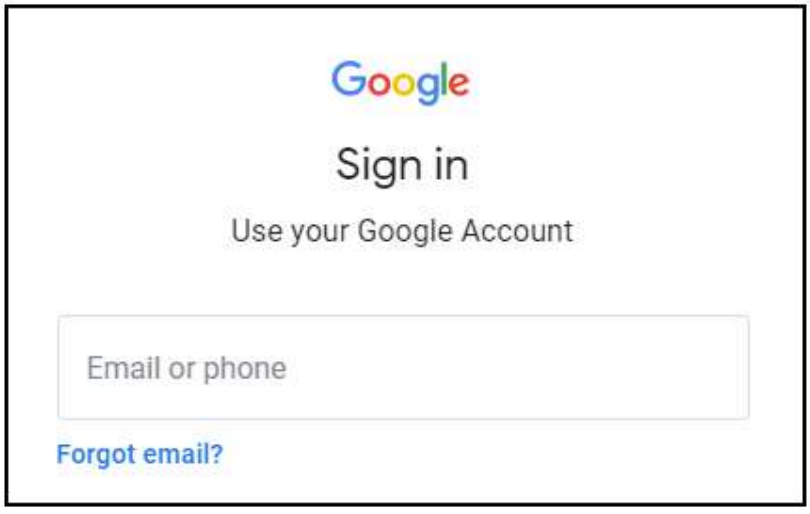
Note: If you are using a Pixelbook, open an Incognito window to run this lab.

How to start your lab and sign in to the Google Cloud Console

1. Click the **Start Lab** button. If you need to pay for the lab, a pop-up opens for you to select your payment method. On the left is a panel populated with the temporary credentials that you must use for this lab.

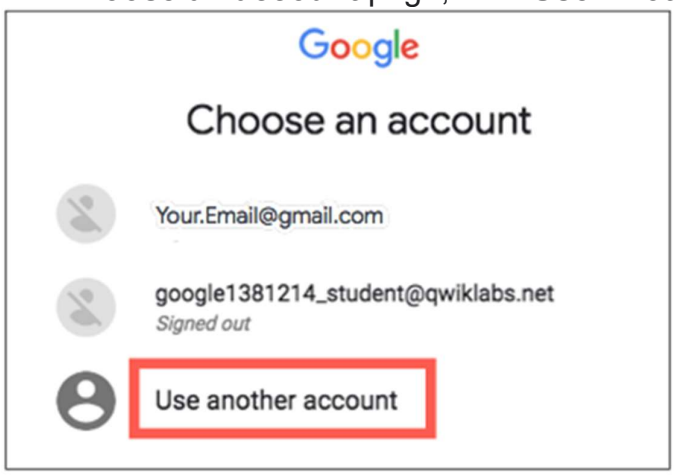


2. Copy the username, and then click **Open Google Console**. The lab spins up resources, and then opens another tab that shows the **Sign in** page.



Tip: Open the tabs in separate windows, side-by-side.

If you see the **Choose an account** page, click **Use Another Account**.



3. In the **Sign in** page, paste the username that you copied from the Connection Details panel. Then copy and paste the password.

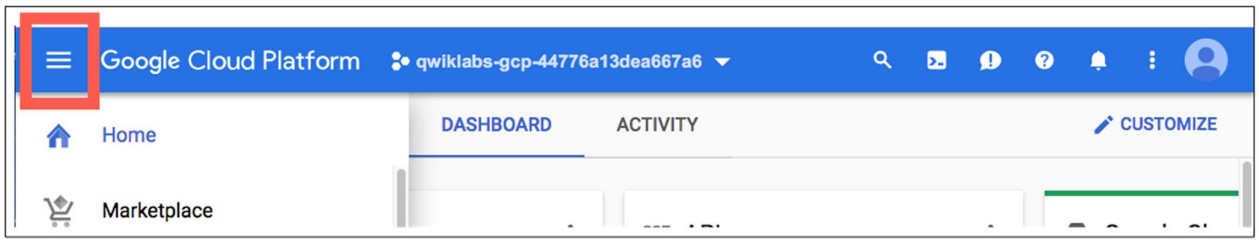
   Important: You must use the credentials from the Connection Details panel. Do not use your Qwiklabs credentials. If you have your own Google Cloud account, do not use it for this lab (avoids incurring charges).

4. Click through the subsequent pages:

   - Accept the terms and conditions.
   - Do not add recovery options or two-factor authentication (because this is a temporary account).
   - Do not sign up for free trials.

After a few moments, the Cloud Console opens in this tab.

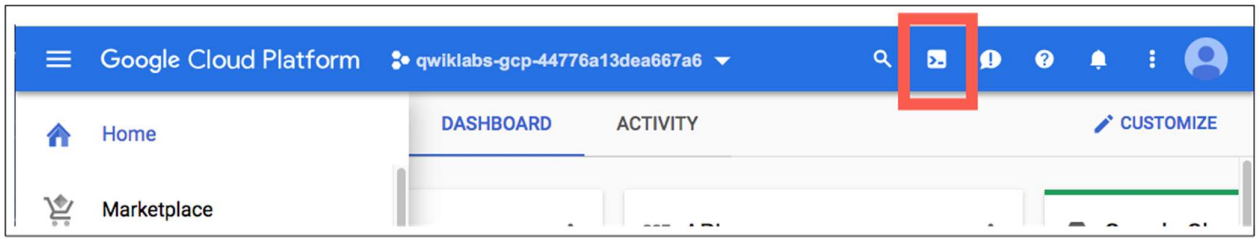
Note: You can view the menu with a list of Google Cloud Products and Services by clicking the **Navigation menu** at the top-left.



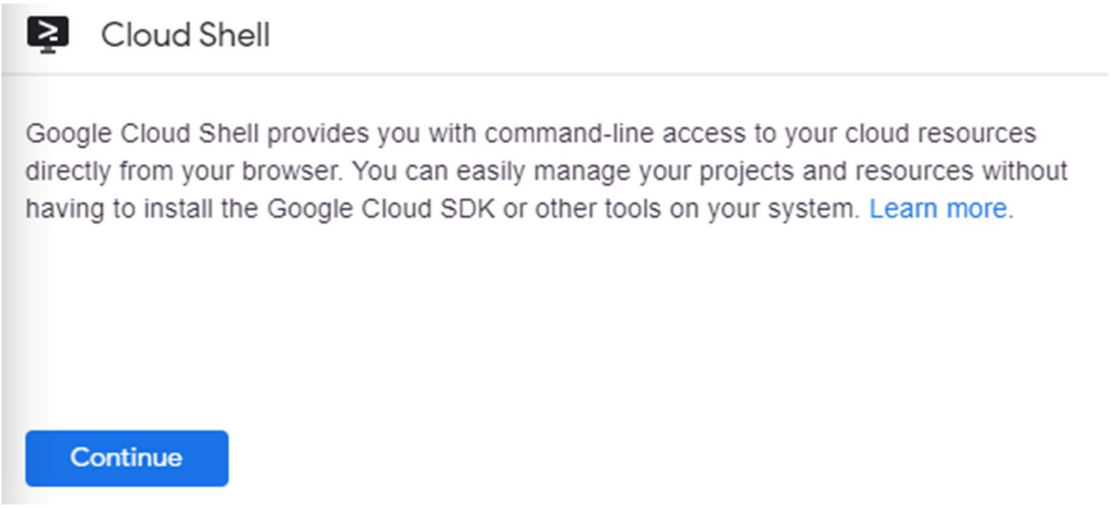
Activate Cloud Shell

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.

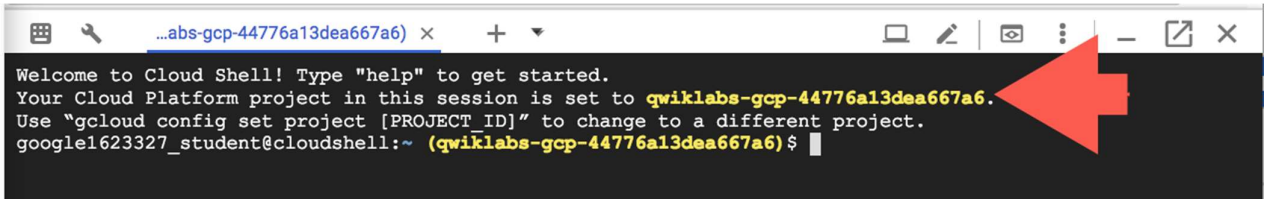
In the Cloud Console, in the top right toolbar, click the **Activate Cloud Shell** button.



Click **Continue**.



It takes a few moments to provision and connect to the environment. When you are connected, you are already authenticated, and the project is set to your *PROJECT_ID*. For example:



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

You can list the active account name with this command:

```
gcloud auth list
```

(Output)

```
Credentialed accounts:
 - <myaccount>@<mydomain>.com (active)
```

(Example output)

```
Credentialed accounts:
 - google1623327_student@qwiklabs.net
```

You can list the project ID with this command:

```
gcloud config list project
```

(Output)

```
[core]
project = <project_ID>
```

(Example output)

```
[core]
project = qwiklabs-gcp-44776a13dea667a6
```

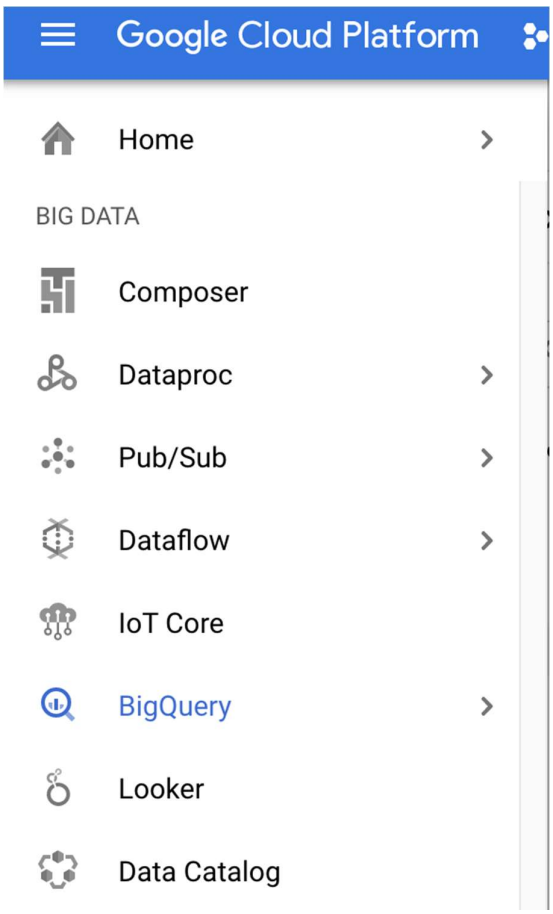
For full documentation of `gcloud` see the gcloud command-line tool overview.

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 BigQuery Console

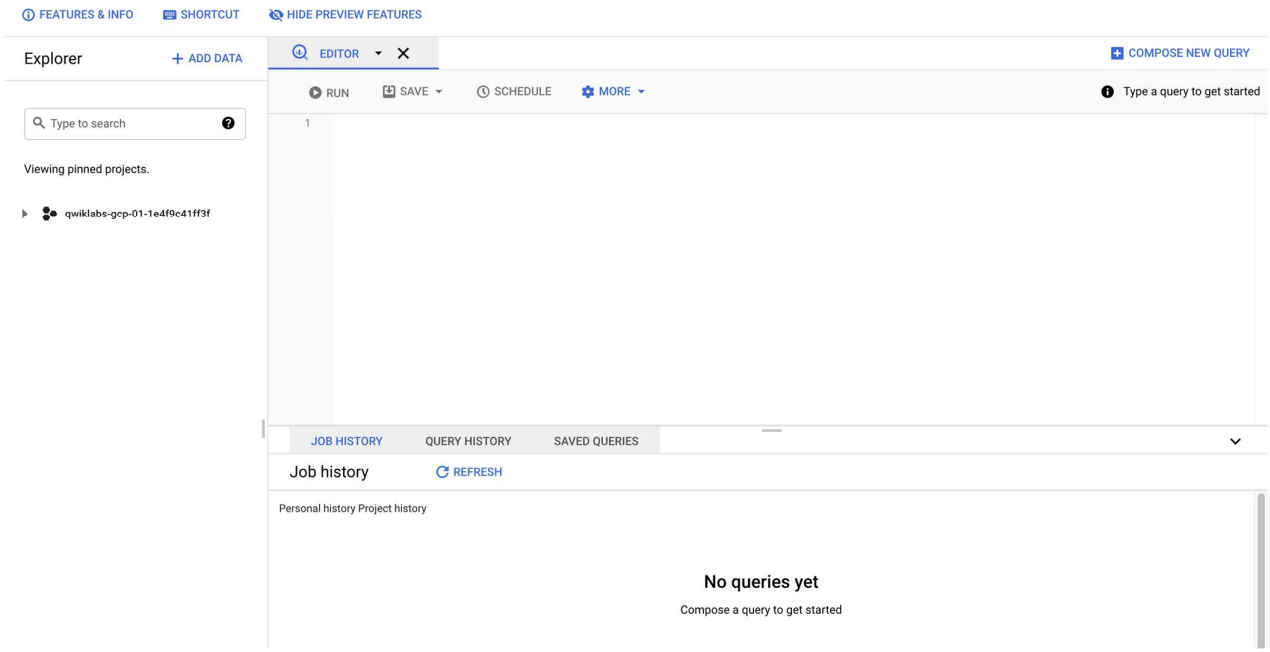
In the Google Cloud Console, select **Navigation menu** > **BigQuery**:



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.

Click **Done**.

The BigQuery console opens.


FEATURES & INFO
SHORTCUT
HIDE PREVIEW FEATURES
Explorer
ADD DATA
Type to search
Viewing pinned projects.
qwiklabs-gcp-01-1e4f9c41ff3f
EDITOR
COMPOSE NEW QUERY
RUN
SAVE
SCHEDULE
MORE
Type a query to get started
1
JOB HISTORY
QUERY HISTORY
SAVED QUERIES
Job history
REFRESH
Personal history Project history
No queries yet
Compose a query to get started

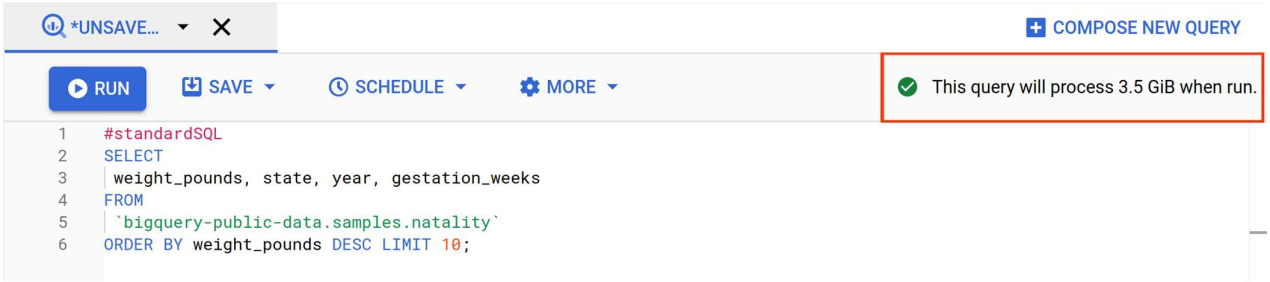
Query a public dataset

1. 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

Test Completed Task

Click **Check my progress** to verify your performed task. If you have successfully query against public dataset, you'll see an assessment score.

You can browse the schema of other public datasets in BigQuery by clicking **+ ADD DATA** > **Explore pubic datasets**, then search for "bigquery public data" in the Search field.

Load custom data into a table

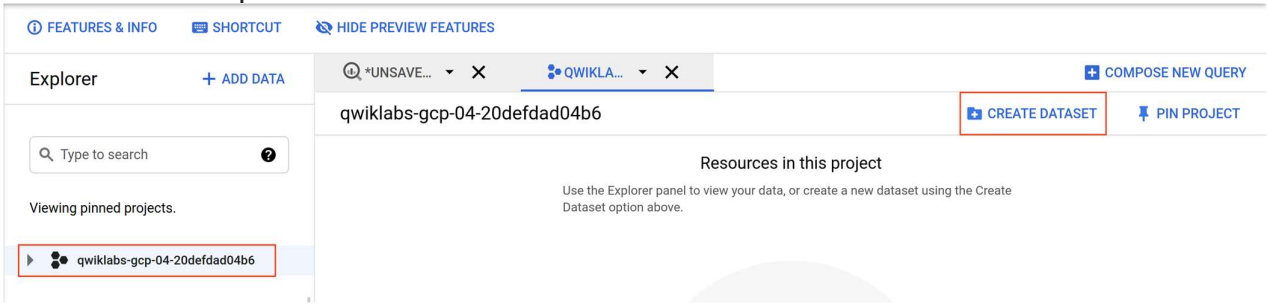
To load custom data into a table, you perform the following tasks:

- Create a dataset
- Create a table
- Add data to your project (to a storage bucket)
- Load the data from the bucket to the table you created

Create a 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.

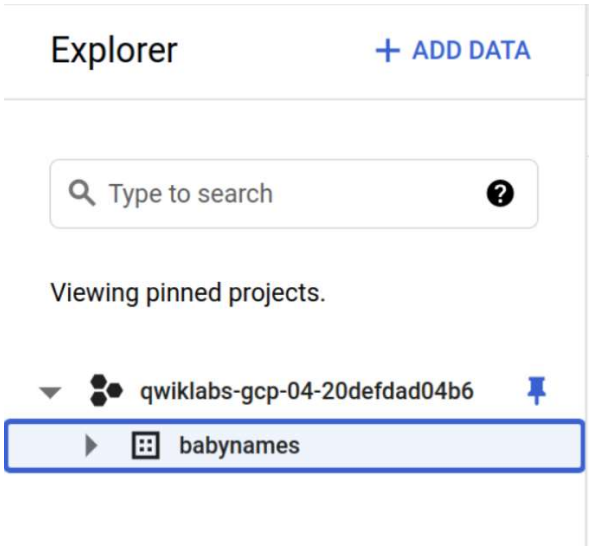
1. In the left pane, click your project name in the **Explorer** section, then click **Create Dataset**. You may have to widen your browser window to see the **Create Dataset** option.



2. Set **Dataset ID** to **babynames**.

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

Now you have a dataset.



Test Completed Task

Click **Check my progress** to verify your performed task. If you have successfully created BigQuery dataset, you'll see an assessment score.

Add custom data

The custom data file you'll use contains approximately 7 MB of data about popular baby names, provided by the US Social Security Administration. You'll add the zip file to your project then create a storage bucket for the specific file that you'll need to query against.

In Cloud Shell, run the following commands to add the data files to your project:

```
gsutil cp gs://spls/gsp072/baby-names.zip .

unzip baby-names.zip
```

Create a Cloud Storage bucket

Now create a Cloud Storage bucket to hold the data files you downloaded.

1. In the Cloud Console, select **Navigation menu** > **Storage** > **Browser**, and then click **Create bucket**.

2. Give your bucket a universally unique name, then click **Create**.

Test Completed Task

Click **Check my progress** to verify your performed task. If you have successfully created a storage bucket, you'll see an assessment score.

3. In Cloud Shell, run the following to move file `yob2014.txt` into your bucket. Replace `<your_bucket>` with the name of the bucket you just created:

```
gsutil cp yob2014.txt gs://<your_bucket>
```

Test Completed Task

Click **Check my progress** to verify your performed task. If you have successfully uploaded object in cloud storage bucket, you'll see an assessment score.

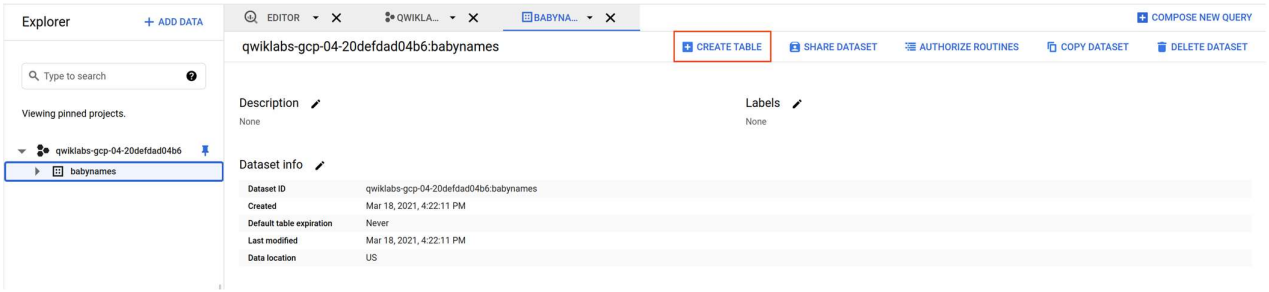
Now you can tell BigQuery where to find the data to query against.

Load the 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.

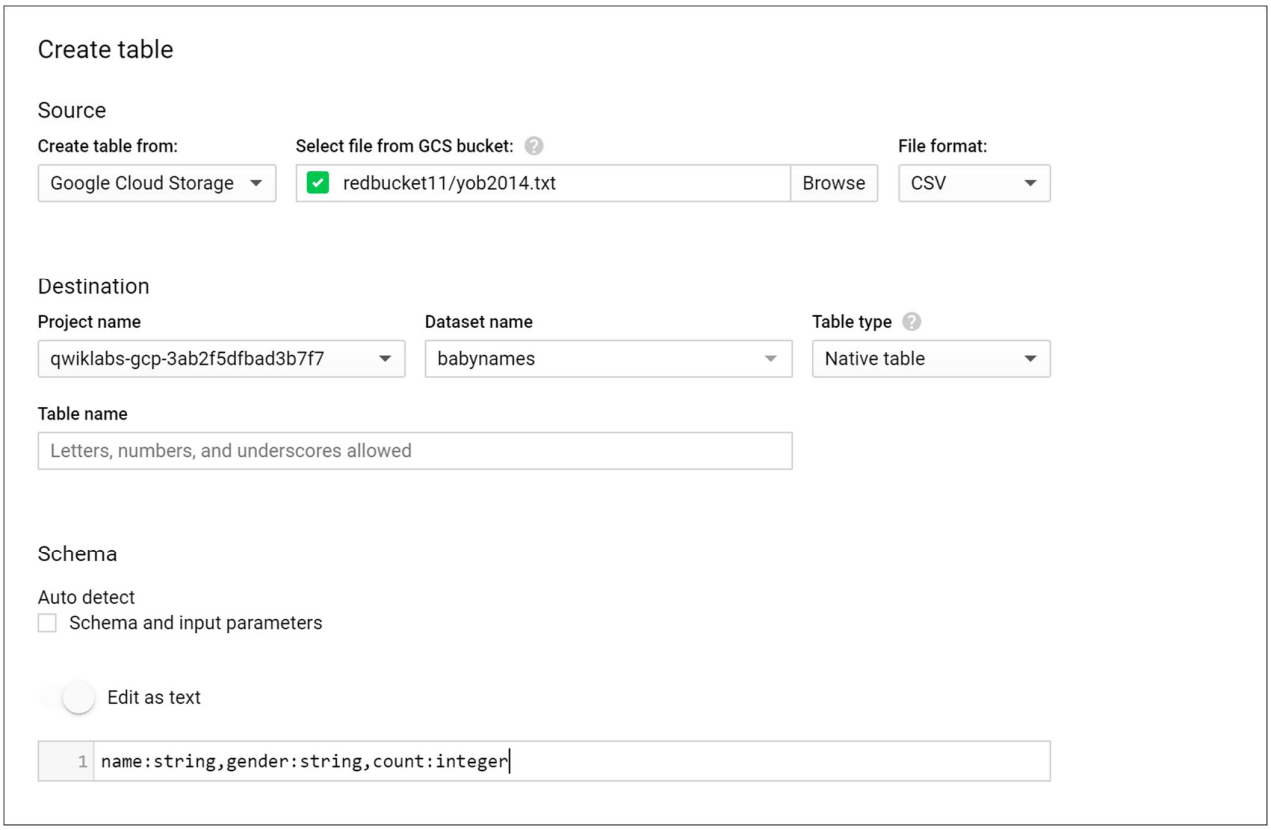
1. In the Cloud Console, select **Navigation menu** > **BigQuery** to return to the BigQuery console.

2. Navigate to the **babynames** dataset, then click **Create table**. You may have to widen your browser window to see the **Create table** option.



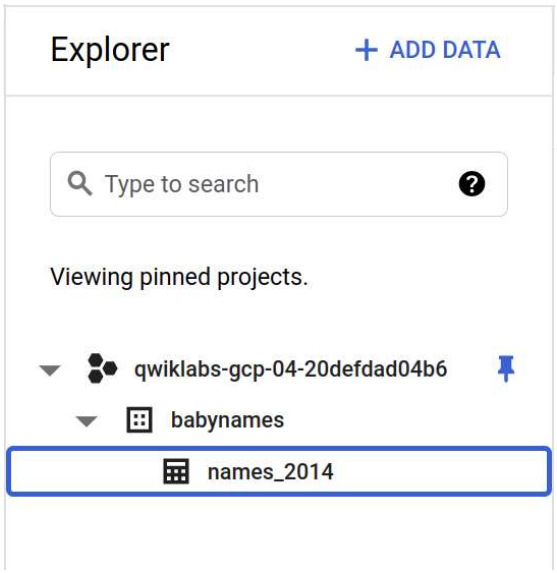
3. In the Create table dialog, set the following fields, leave all others at the default value:

Field	Value
Create table from:	**Google Cloud Storage**
Select file from GCS bucket:	`<bucket_name>/yob2014.txt`, replace <bucket_name> with the name of the bucket you created earlier.
File format:	**CSV**
Table name	`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.



Test Completed Task

Click **Check my progress** to verify your performed task. If you have successfully load data in dataset table, you'll see an assessment score.

Test your Understanding

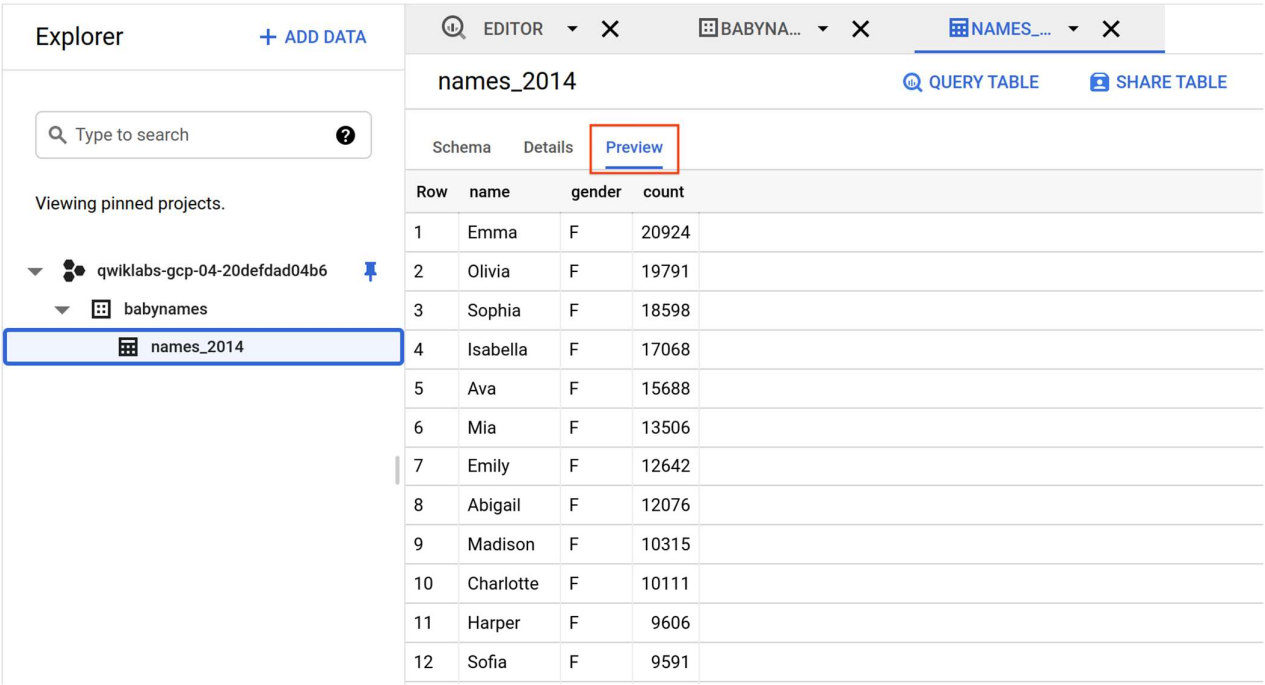
Below are multiple choice-questions to reinforce your understanding of this lab's concepts. Answer them to the best of your abilities.

BigQuery is fully-managed enterprise data warehouse that enable super-fast SQL queries.
True

Preview the table

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

Click the `names_2014` table in the left-hand menu, then click **Preview**.



Your table is ready for queries.

Query a custom dataset

Running a query against custom data is identical to querying a public dataset that you did earlier, except that now you're querying your own table instead of a public table.
In BigQuery, click the **Compose New query** button in the top right corner to clear out your previous query.

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;
```

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

Test Completed Task

Click **Check my progress** to verify your performed task. If you have successfully query against custom dataset, you'll see an assessment score.

Congratulations!

You used the BigQuery Web UI to query public tables and load sample data into BigQuery.

Finish Your Quest

This self-paced lab is part of the Qwiklabs BigQuery Basics for Data Analysts Quest. A Quest is a series of related labs that form a learning path. Completing this Quest earns you the badge above, to recognize your achievement. You can make your badge (or badges) public and link to them in your online resume or social media account. Enroll in a Quest and get immediate completion credit if you've taken this lab. See other available Qwiklabs Quests.

Next Steps / Learn More

This lab is part of a series of labs called Qwik Starts. These labs are designed to give you a little taste of the many features available with Google Cloud. Search for "Qwik Starts" in the lab catalog to find the next lab you'd like to take!

Google Cloud Training & Certification

...helps you make the most of Google Cloud technologies. Our classes include technical skills and best practices to help you get up to speed quickly and continue your learning journey. We offer fundamental to advanced level training, with on-demand, live, and virtual options to suit your busy schedule. Certifications help you validate and prove your skill and expertise in Google Cloud technologies.
Manual Last Updated: March 18, 2021
Lab Last Tested: March 18, 2021