Hands-On Activity: Create a custom table in BigQuery ← Back Practice Quiz • 1h • 2 total points

Congratulations! You passed!

Grade received 100% To pass 80% or higher

Go to next item

1/1 point



Activity overview

What you will need

Recently, you've been thinking about identifying good data sources that would be useful for analysis. You also spent some time in a previous activity exploring a public dataset in BigQuery and writing some basic SQL queries. In addition to using public data on BigQuery, you will need to be able to import data from other sources. In this activity, you will create a custom table and dataset, which you'll load into a new table and query.

By the time you complete this activity, you will be able to load your own data into BigQuery for analysis. This will enable you to import your own data sources into BigQuery, which is a skill you will need in order to analyze data from different sources.

To get started, download the baby names data zip file. This file contains about 7 MB of data about popular baby names from the US Social Security Administration website.

Click the link to the baby names data zip file to download it.

Link to baby names data: names.zip

Create a custom table

and a custom table. Step 1: Unzip the file

Once you have the zip file downloaded, you can import it into BigQuery to query and analyze. In order to do that, you will need to create a new dataset

You will need to unzip the file you downloaded onto your computer in order to access it on BigQuery. Once you have unzipped the file, you will find a .pdf file titled NationalReadMe that contains more information about the dataset. This dataset tracks the popularity of baby names for each year; you

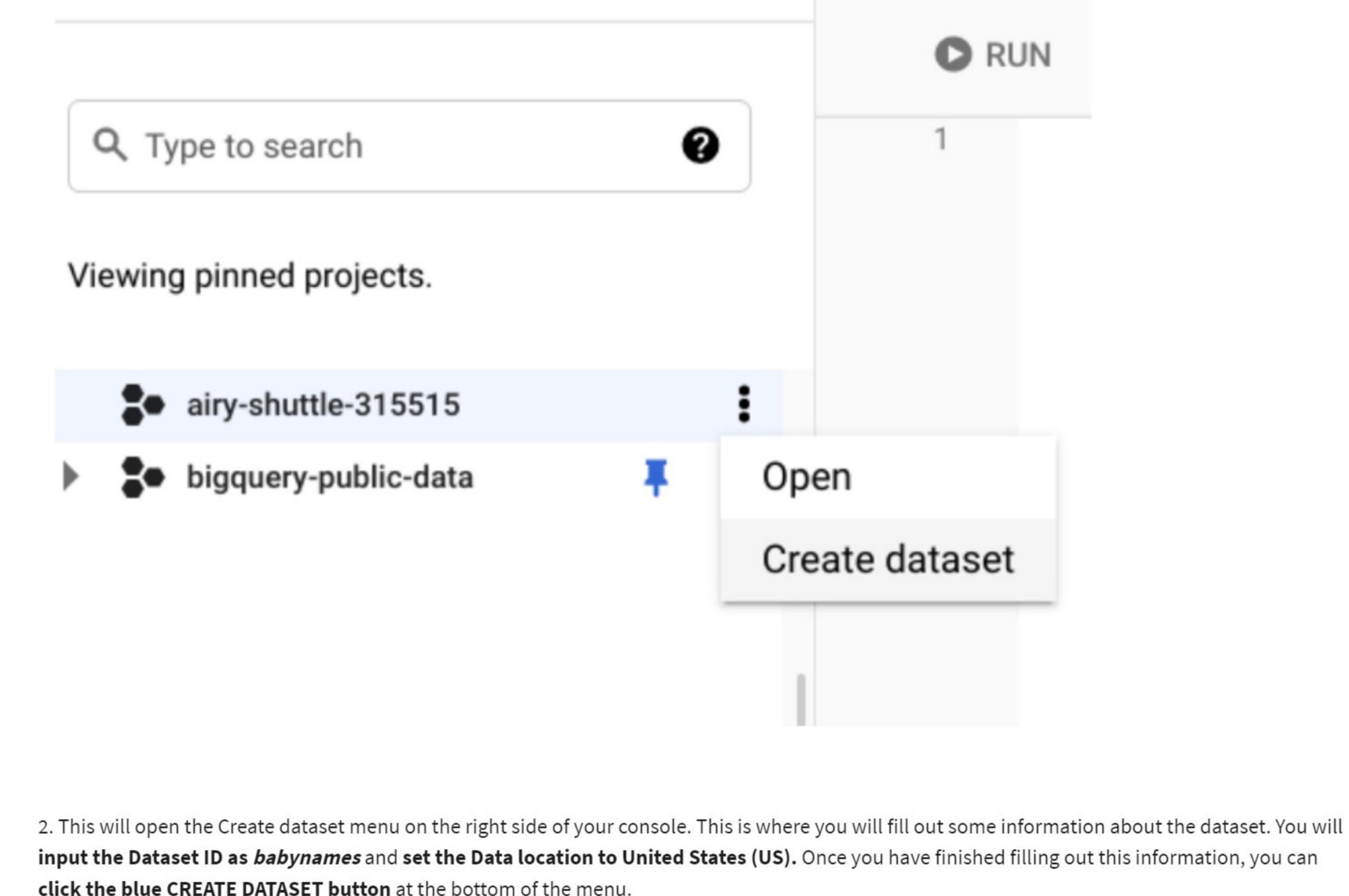
Remember where you saved this folder so you can reference it later. Step 2: Create a dataset Before you can upload your txt file and create a table to query, you will need to create a dataset to upload your data into and store your tables.

can find text files labelled by the year they contain. Open yob2014.txt to preview the data. You will notice that it's a .csv file with three columns.

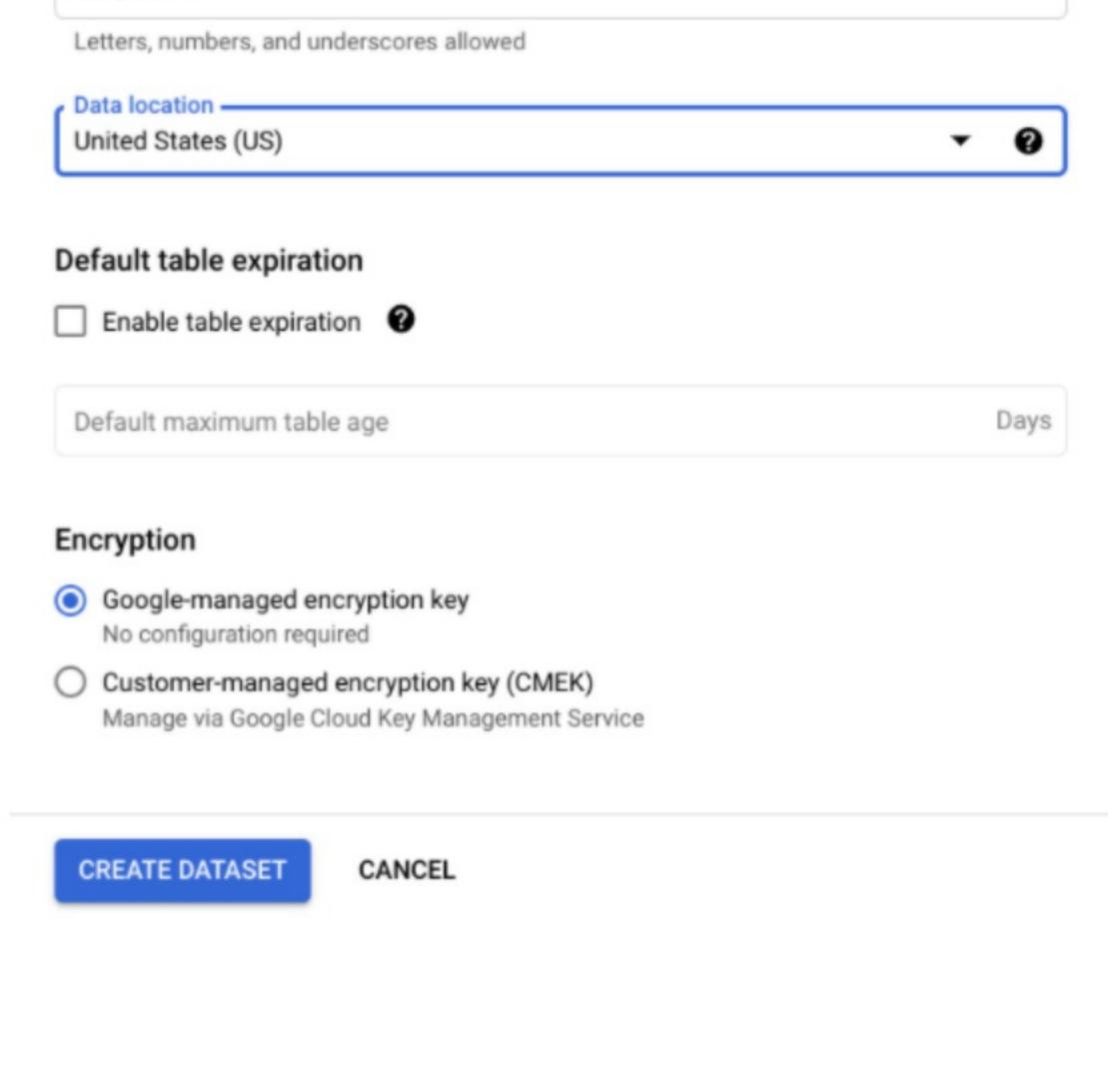
1. Go to the Explorer pane in your workspace and click the three dots next to your pinned project to open a menu. From here, select Create

dataset.

EDITOR Explorer + ADD DATA



Dataset ID * babynames Letters, numbers, and underscores allowed



⊞BABYNA... ▼ X

⊕ EDITOR ▼ X airy-shuttle-315515:babynames

Create table

Create table from:

Select file: 0

yob2014.txt

Source

Upload

Table name

names_2014

Step 3: Create table

Create dataset

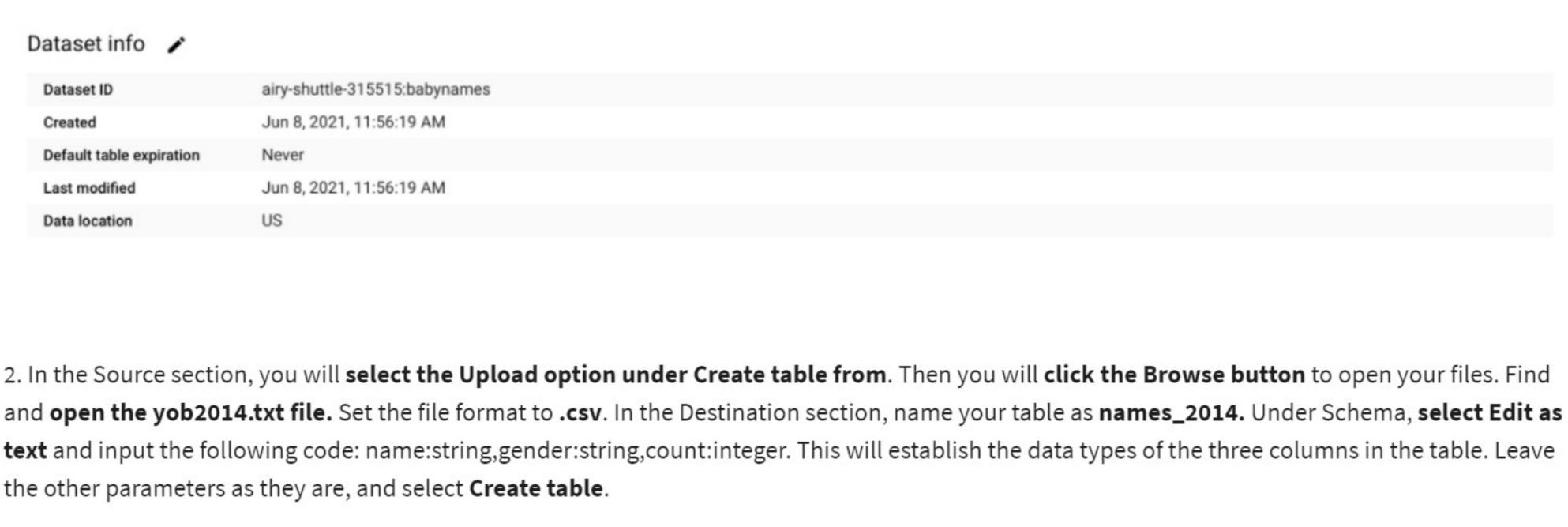
Labels 🖍 Description /

CREATE TABLE

Now that you have a custom dataset stored in your project space, this is where you will add your table.

None None

1. From the babynames dataset, click the CREATE TABLE button. This will open another menu on the right side of your console.



SHARE DATASET

+ COMPOSE NEW QUERY

TOTAL

COPY DATASET

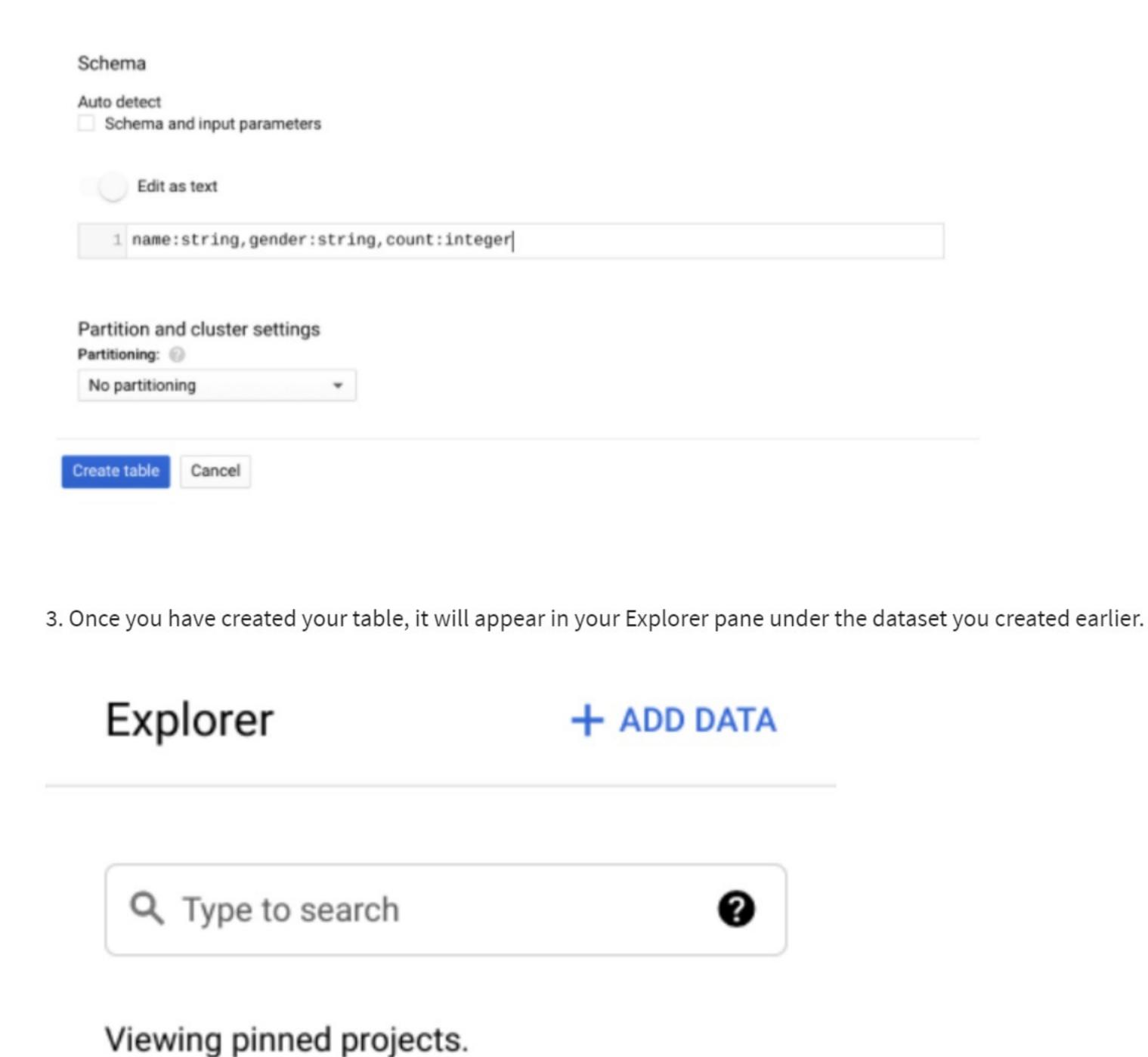
≔ AUTHORIZE ROUTINES

Destination Search for a project Enter a project name test babynames Native table

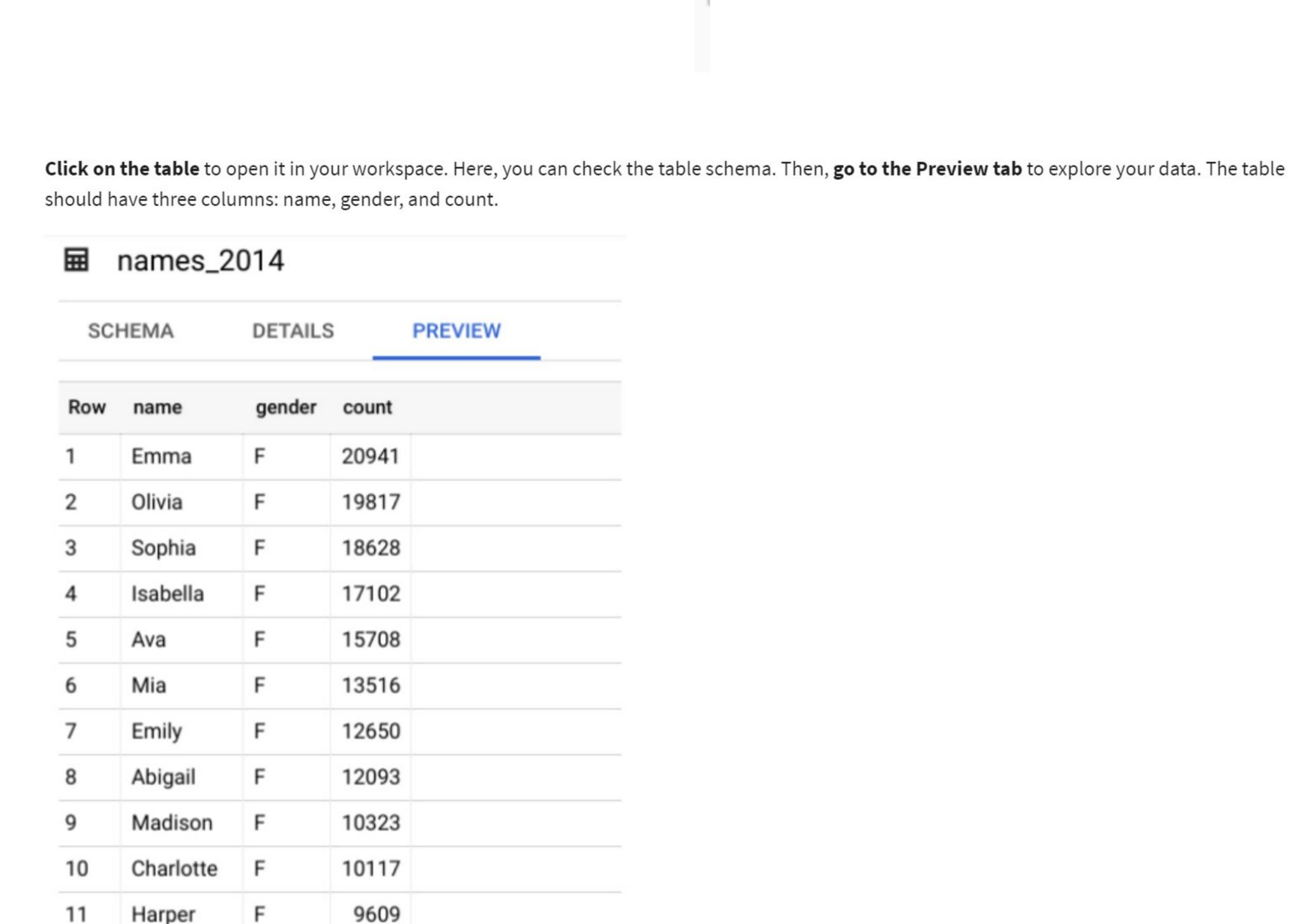
Browse

File format:

CSV



bigquery-public names_2014



airy-shuttle-315515

babynames

names_2014

Avery

`babynames.names 2014`

F

Harper

Sofia

11

12

13

WHERE

LIMIT

ORDER BY

gender = 'M'

count DESC

Query your custom table

Now that your table is set up, you're ready to start writing queries and answering questions about this data. For example, let's say you were interested

Click **COMPOSE NEW QUERY** to start a new query for this table. Then copy and paste this code: SELECT name, count FROM

in the top five baby names for boys in the United States in 2014.

9599

9573

results. Then, you're sorting how you want your results to appear with ORDER BY. Because you are ordering by the count in descending order, you will get names and the corresponding count from largest to smallest. And finally, LIMIT tells SQL to only return the top five most popular names and their counts.

Once you have input this in your console, select **RUN** to get your query results.

you could upload each of the files from 2015 to 2019 to find the top baby names for those years.

Jacob

Noah

Up for a challenge?

Confirmation and reflection After running the query on your new table, what was the third most popular baby name for boys in 2014?

If you are comfortable creating your own custom tables, try uploading more files from the baby names dataset into tables you can query. For example,

This query SELECTs the name and count columns from the names_2014 table. Using the WHERE clause, you are filtering for a specific gender for your

Mason William Correct

response to each of the following questions: Why is being able to use data from different sources useful as a data analyst?

Why is being able to use data from different sources useful as a data analyst? How can you use BigQuery custom tables and datasets in your future analysis projects?

How can you use BigQuery custom tables and datasets in your future analysis projects?

queries for more data sources, which will be a key skill as a data analyst.

Correct Congratulations! In this activity, you created a new dataset within your project, uploaded a .csv file to create a new table, and ran a SQL query. A good response would include that being able to evaluate and use different data sources allows you access more data.

2. In this activity, you explored public data in BigQuery and used it to create a custom table. In the text box below, write 2-3 sentences (40-60 words) in

To find that Mason was the third most popular baby name for boys in 2014, you queried your custom table and checked the results. Going

forward, you'll be able to upload your own data sources into BigQuery for future analysis projects. This will allow you to practice writing SQL

As a data analyst, being able to evaluate data sources and use the appropriate tool to analyze them is important. For instance, you were able to use SQL to analyze a dataset that was previously stored on your computer as a .csv file.

1/1 point