

Power BI, Google Sheets and BigQuery – Do they work together?



By Sangeetha Mahesh

Hi Again! I'm back with another post on BigQuery, Power BI and Google Sheets. Today, we'll be looking at how to connect Power BI and Google Sheets.

Ideally, I'd want to create a DirectQuery connection between Power BI and Google Sheets, because the data in my Google spreadsheet changes frequently, and we need the latest data to be reflected in the Power BI report. Unfortunately, Power BI doesn't currently support a DirectQuery connection to Google Sheets.

This is where BigQuery will help us out. We'll try to use BigQuery as an intermediary to connect Power BI to Google Sheets and access the data we need in the way we need it. We'll need to explore creating a DirectQuery connection between Power BI and an external table on BigQuery.

What is a Direct Query Connection?

Power BI offers different ways to connect to its data sources – Import, DirectQuery and Live connection. A DirectQuery Connection is when the data isn't stored in the Power BI Server but in the data source itself. Every time you refresh the report, it queries the data from the source and gets it.

Power BI recommends that it's best to use the 'Import' connection whenever possible. However, if we have data changes frequently and reports that must reflect the latest data, then DirectQuery fits

best. Performance might take a small hit, but it becomes essential when dealing with huge datasets which can't be imported to Power BI.

Data sources which allow for direct connections are Google BigQuery, SQL Server, Azure Data Lake etc., You can find more information about Power BI supported [data sources here](#).

What is an External table?

Google BigQuery has two ways in which it stores its data:

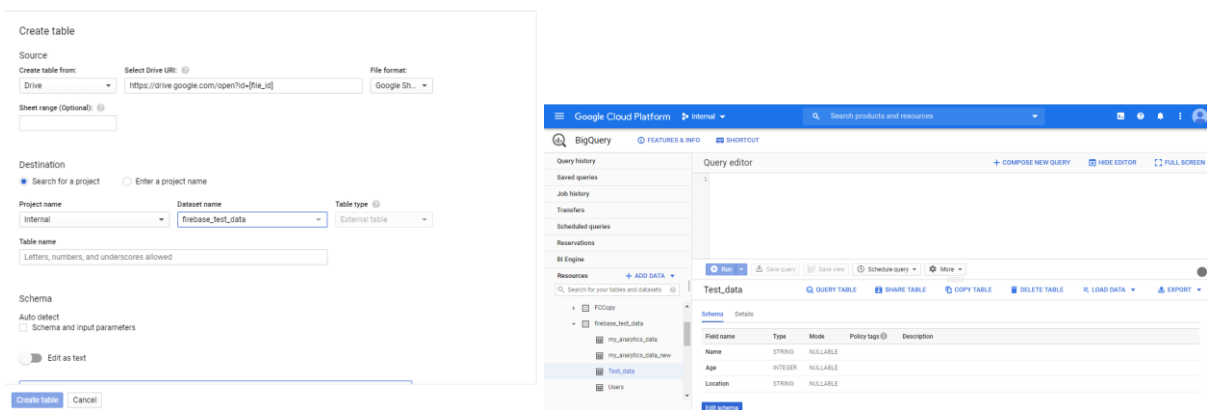
1. Native Table – this is where the data gets imported and is loaded into BigQuery or a table is created in BigQuery and data is inserted. This consumes storage on BQ
2. External Table – this is where the data doesn't get imported but stays in the source. This table is linked to the data source and it gets the data every time the associated BigQuery table is queried. This doesn't consume storage on BQ

Any data which is stored on Google Drive can only be accessed as an external table on BigQuery and this goes for Google sheets as well.

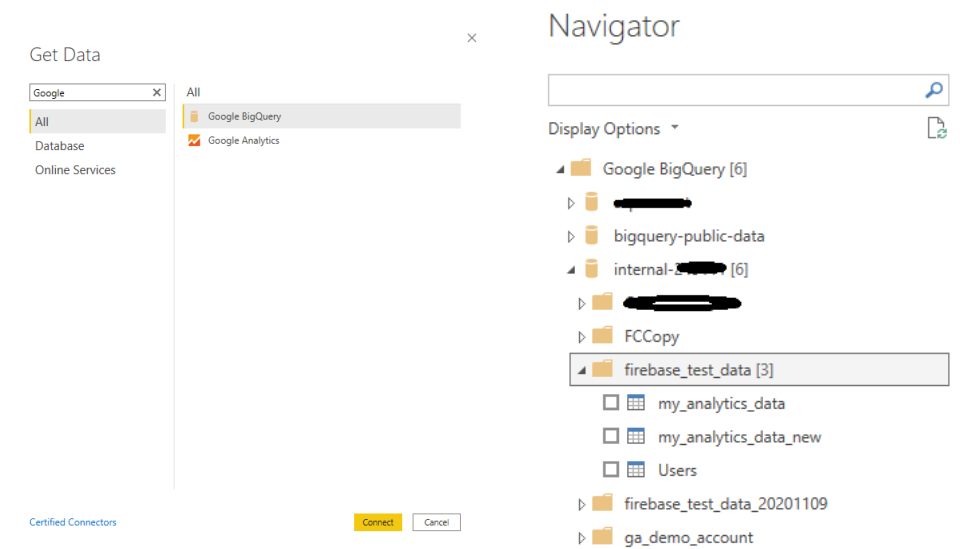
Time for some Action!

I have a Google spreadsheet named 'Test_data' on my Google Drive. I'll load this spreadsheet to BigQuery using the web console.

I'll presume that you have a BigQuery account, a project and a dataset to use for this exercise. Pick a dataset and click on the 'Create Table' option. You'll then be presented with a screen like below. Fill in the details and your BQ external table is ready.



Now it's time to test the DirectQuery connection in Power BI. So, open Power BI and click on the 'Get Data' option. Select the 'Google Big Query' and click 'Connect'. In the resulting window, navigate to the dataset containing the 'Test_data' table.



Well, that's strange! I had 4 tables in my dataset but I can only view 3 of them. This is because the Power BI Google BigQuery Connector doesn't support External BigQuery tables. It only supports the Native BigQuery tables. Since Test_data is an external table, it doesn't get listed here.

This information is not mentioned in the [Power BI documentation for Google BigQuery](#) either. So, we'll have to look at alternative ways to get that data.

Is there a workaround?

Yes. We'll have to convert our External table into a Native table on BigQuery so that Power BI can access it. Let's look at how to do get this done.

Creating a new BigQuery table

This can be done on the BQ web console with a DDL statement. Here it goes

```
CREATE TABLE `PROJECT_NAME.DATASET_NAME.TEST_DATA_NEW`
AS SELECT * FROM `PROJECT_NAME.DATASET_NAME.TEST_DATA`
```

The above statement creates a new table named 'Test_data_new' and copies the data from 'Test_data' table onto it. Now we have a Native table in BigQuery containing the google spreadsheet data.

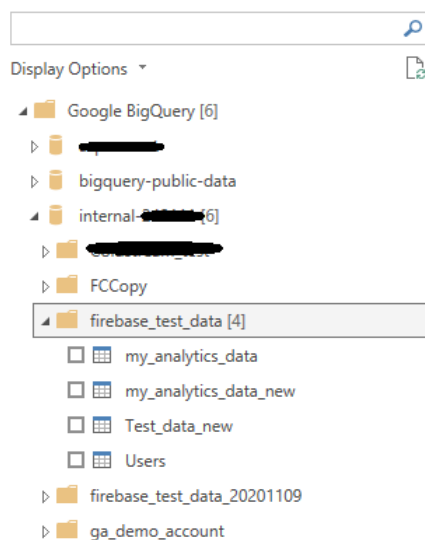
It's also possible to write a [scheduled query](#) and update the Native table on a recurring basis once it has been created. Note that the minimum time that BigQuery allows you to refresh a table is at 15 minutes. So, with this solution, your report gets updated data every 15 minutes. You could use the below query to get started.

```
INSERT `PROJECT_NAME.DATASET_NAME.TEST_DATA_NEW` (NAME, AGE, LOCATION)
SELECT * FROM `PROJECT_NAME.DATASET_NAME.TEST_DATA`
```

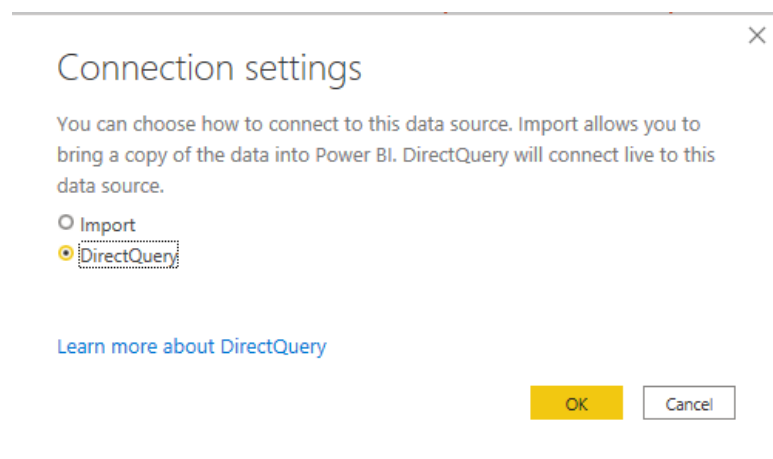
Connecting to Power BI

Let's retrace our steps. Open Power BI and click on the 'Get Data' option. Search for Google BigQuery and select it. Once it opens up, navigate to your new table (Test_data_new) and click on the Load button.

Navigator



You'll then be presented with a Connecting settings window, where you need to select the DirectQuery option and click OK.



There you go! I can finally access my Google spreadsheets data in Power BI and the report gets updated data every 15 minutes. Now, that is a perfect solution to this problem. Please feel free to ask any questions by contacting me through the contact box!