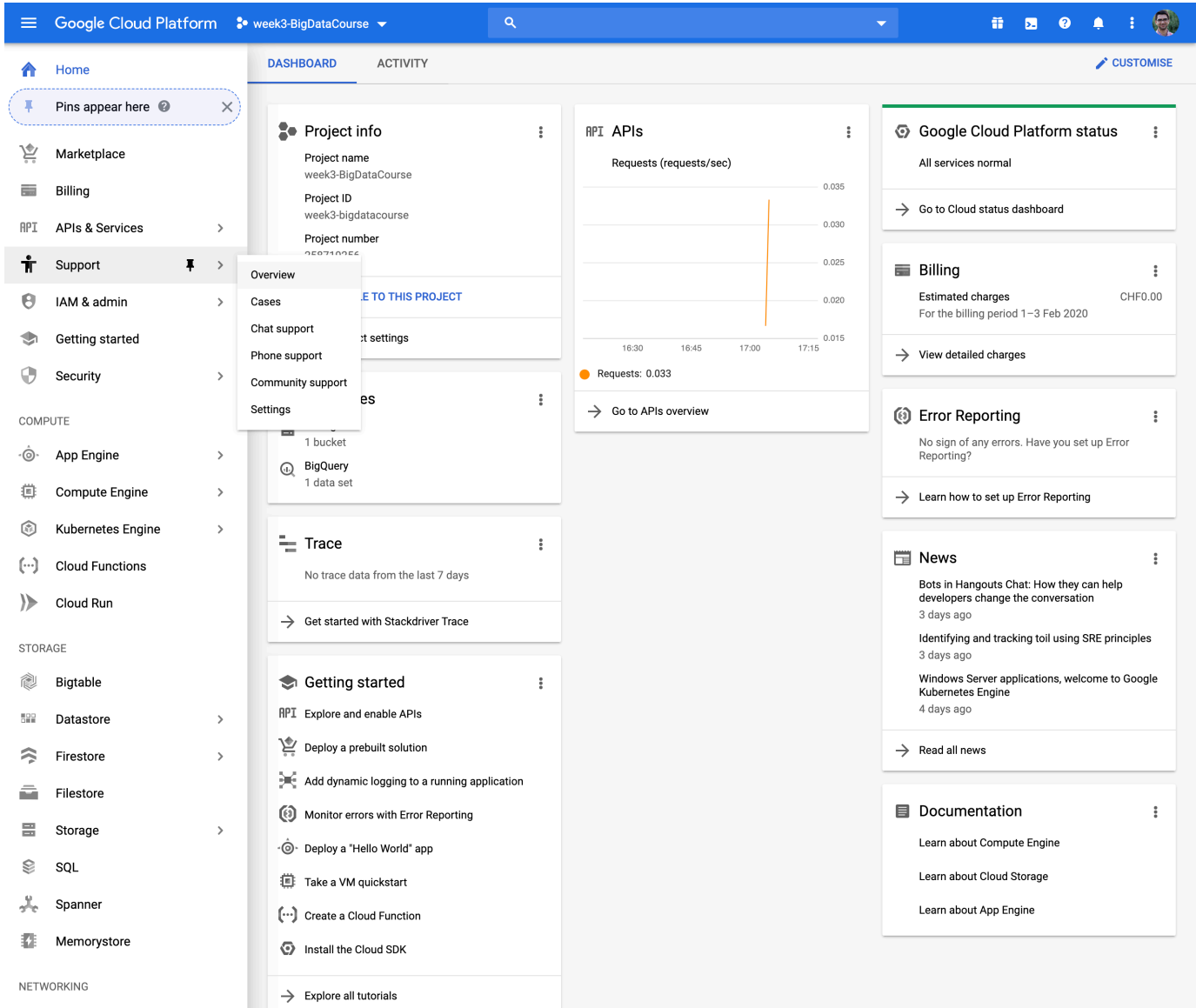
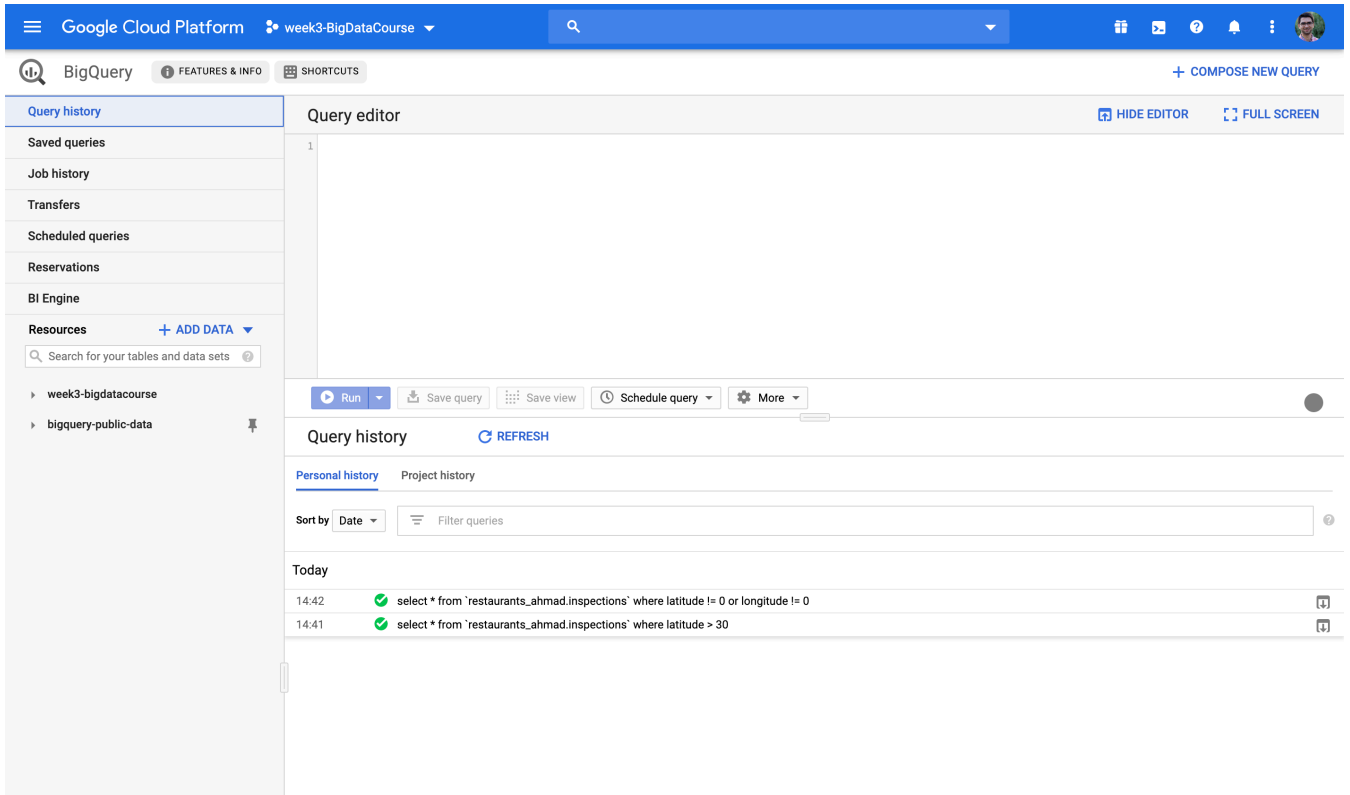
The web UI console

- After creating your account , you can see the console
- You can access several services from the panel on the left
- Try to play with this console a little bit!



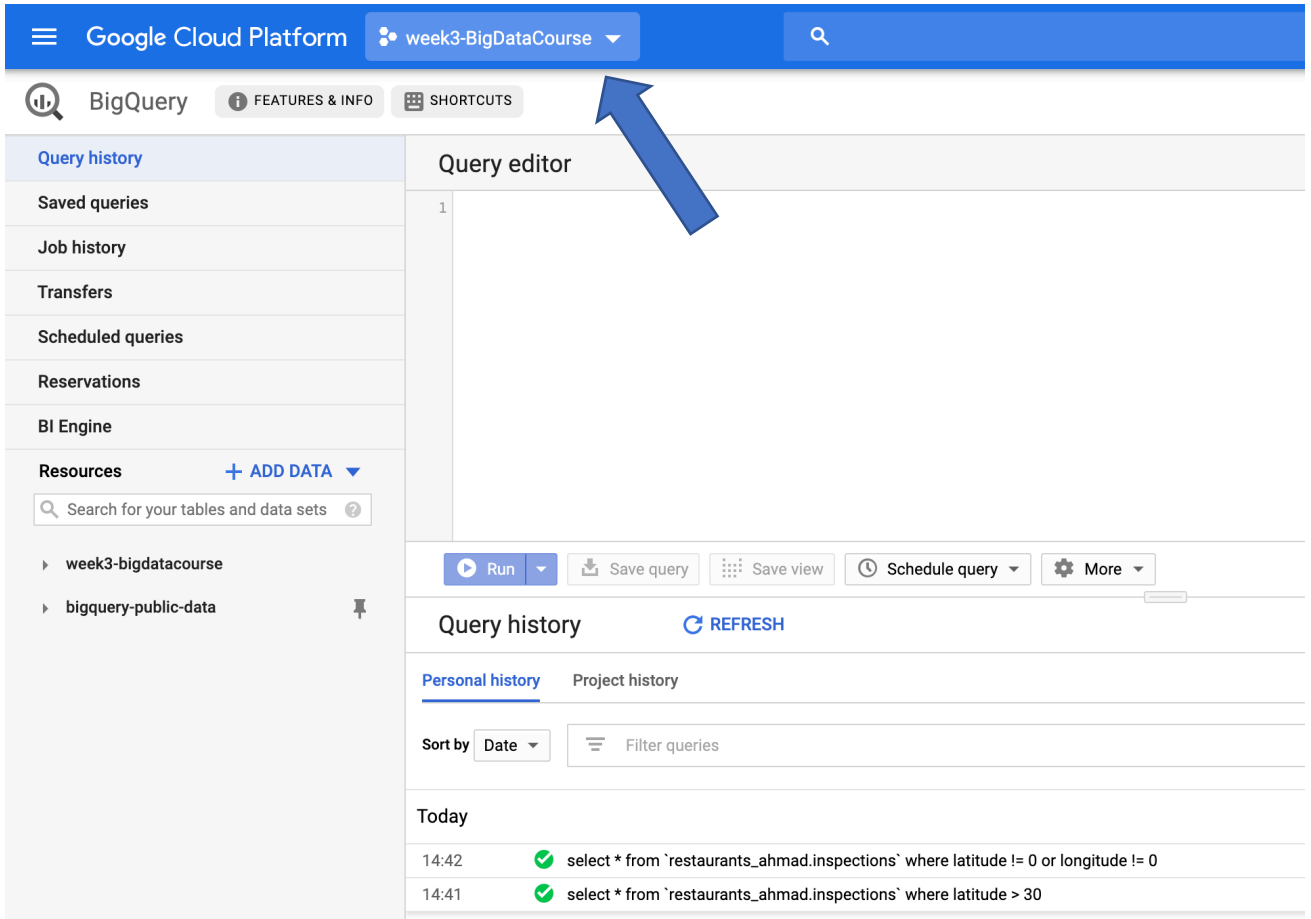
Big Query

- From the panel on the left click on BigQuery. You will see a page like this:



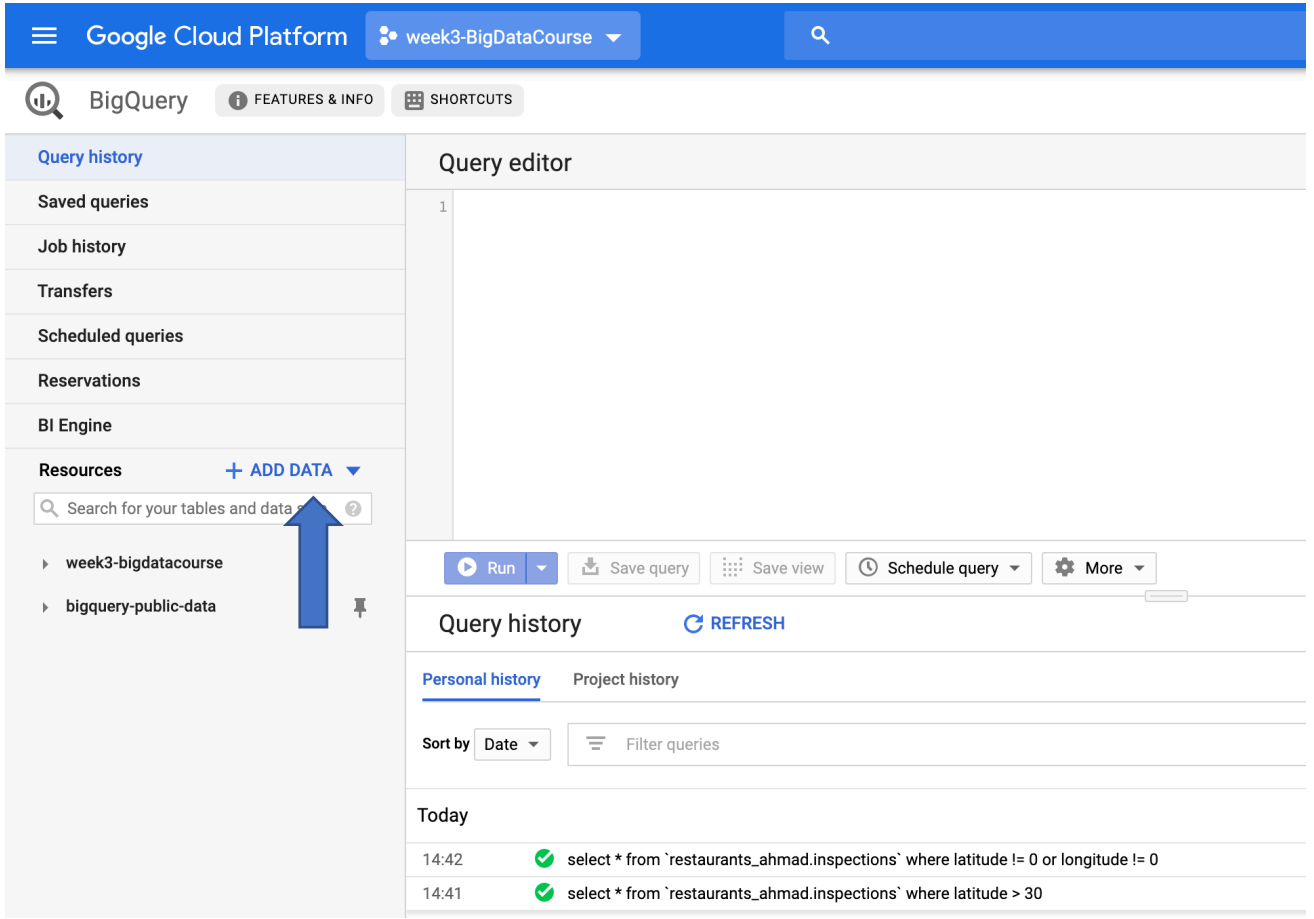
Create you own project in BigQuery

- On the top left you can see the project name which is set by default. But you can make your own project. Just click on it and then in the window that pops up select New Project.



Public Datasets in Google cloud

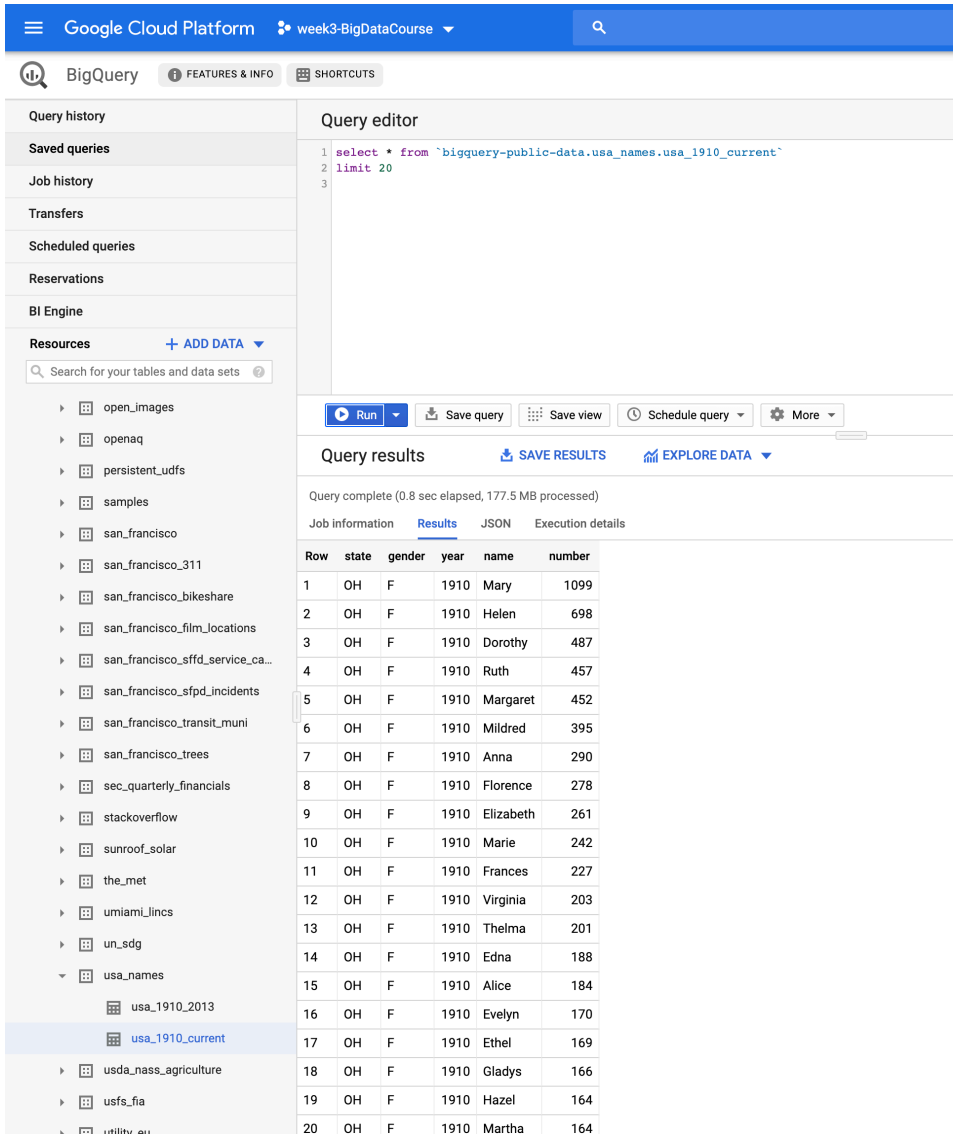
- There are several public datasets available within Google cloud which you can use them
- To do so, just click on the ADD DATA button in the Resources tab on the left and then explore the datasets available.
- For instance you can search for “usa_names” or just “names” and explore the dataset of names selected for newborns in USA in different years



Run sample queries using the query editor

- Run some sample queries on the `usa_1910_current` table using the query editor
- Pay attention to the way the table name is specified in the example query in the screenshot

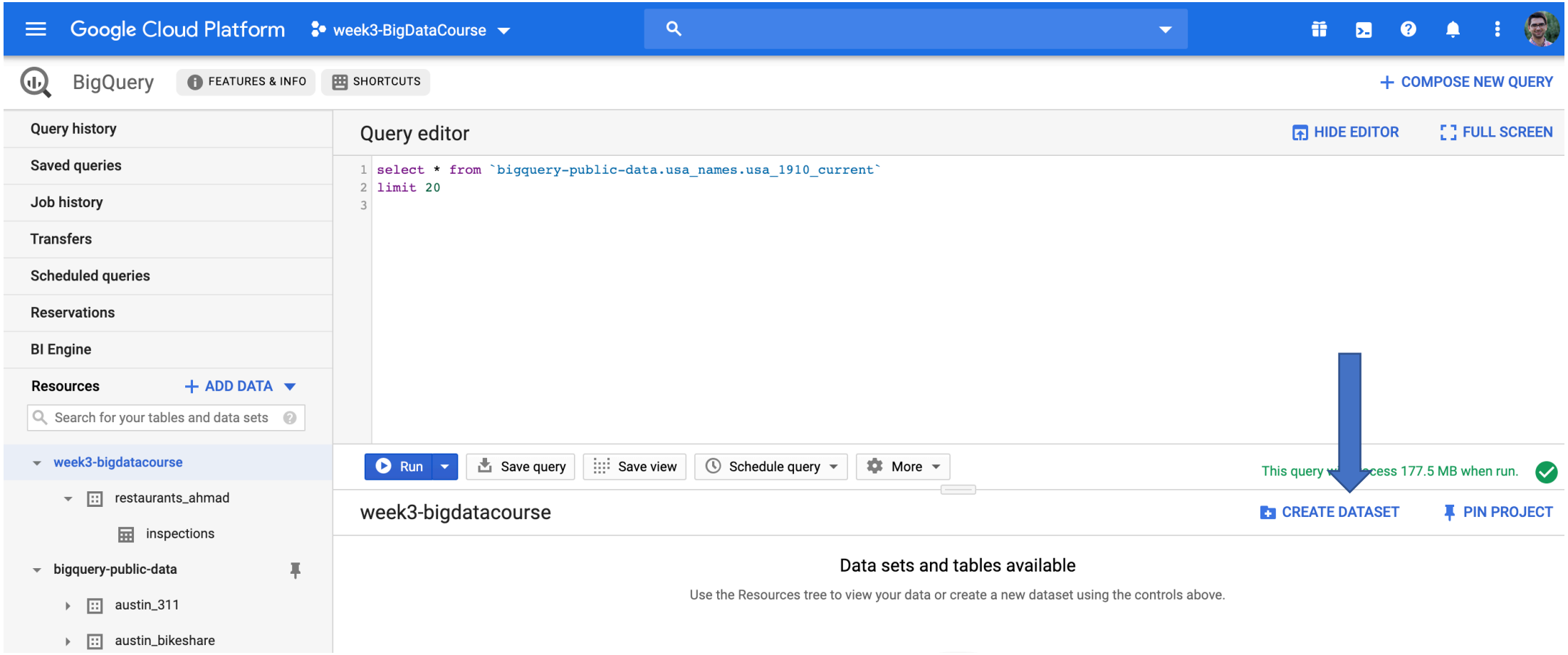
Project_name.dataset.table.name



Create your own dataset in BigQuery

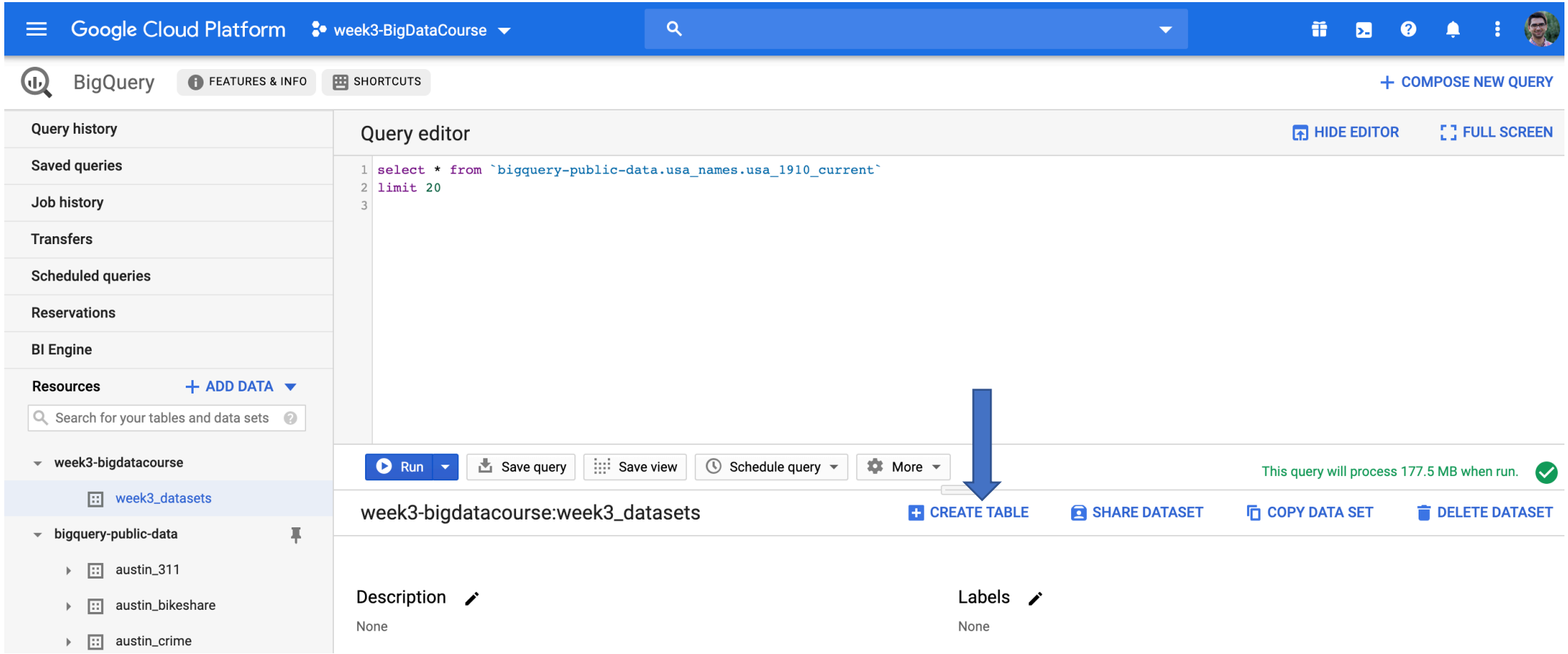
- It is possible to upload data to google cloud and create your own data-set in BigQuey
- For that, first click on your project name on the right panel
- Then click in the CREAT DATASET button
- In the window that pops up, choose a proper name for your dataset
- It is better to choose locations in EU to have less latency when querying your data
- Notice that a dataset can contain several tables (eg, csv files that you upload)
- Check out the screen shot in the next slide

Create Dataset



Create table

Click on the dataset you just created. Then click on CREAT TABLE to upload a csv file.



Create table: World cup data

- You can upload different formats like csv, json, etc.
- The schema could be automatically inferred by BigQuery, however if there are problems then you would probably need to specify the schema
- As an example try to upload the Teams.csv and Players.csv from previous week and create two tables under the dataset you made
- Try to run some queries on these tables

Query editor

```
1 select * from `week3-bigdatacourse.week3_datasets.Teams`
2 where ranking < 30
3 order by ranking asc
4
```

Access BigQuery tables from Python

In Colab:

- You need to use the google.cloud package (already available in colab)
- You would also need to give your credentials. It is easy to do this in colab. Run the following code:

```
from google.colab import auth
auth.authenticate_user()
print('Authenticated')
```

- You will be redirected to a page (and possibly you need to sign in to your google account). There is a key there, just copy it and paste it in colab in the box that appeared by running the above code.
- You're all set to connect to BigQuery client

Access BigQuery tables from Python

In your local machine:

- Follow the instruction here to install the client library and setting up the authentication.
- Edit the `bash_profile` file in mac (or `.bash_rc` file in linux) and add the line below to it. Instead of '[PATH]', enter the path to the json file you just downloaded (which contain your credentials)

```
export GOOGLE_APPLICATION_CREDENTIALS="[PATH]"
```

- Save your changes and do `source .bash_profile` to apply the changes.
- Make sure to run your python script or jupyter notebook from the same shell that you create the above environment variable