



Data Engineer Take Home

The following shouldn't take more than 3 hours to complete. **Please complete this exercise in Python.**

Setup

1. Ensure that you have a [Github](#) and [Google](#) account.
2. Activate your [free Google Cloud Platform](#) account, create a Google Cloud Platform project and go to BigQuery.
3. Click on Add Data followed by Explore public datasets. After clicking on any dataset, on the left, you should be able to view the list of datasets under a project called bigquery-public-data. For the purpose of this exercise, we will only be looking at one dataset, geo_international_ports.

The link should look something like this:

`https://console.cloud.google.com/bigquery?project=<INSERT_YOUR_PROJECT_HERE>&p=bigquery-public-data&d=geo_international_ports&t=world_port_index&page=table`

4. Create a *private* Github repository to store any code and your progress. Your repository should also indicate how any other user can use it.

Questions

For each of the following questions, create a script that creates a BigQuery table in your personal Google Cloud Project.

1. What are the 5 nearest ports to Singapore's **JURONG ISLAND** port?
(country = 'SG', port_name = 'JURONG ISLAND')
Your answer should include the columns port_name and distance_in_meters only.
2. Which country has the largest number of ports with a cargo_wharf? Your answer should include the columns country and port_count only.
3. You receive a distress call from the middle of the North Atlantic Ocean. The person on the line gave you a coordinates of lat: 32.610982, long: -38.706256 and asked for the nearest port with provisions, water, fuel_oil and diesel. Your answer should include the columns country, port_name, port_latitude and port_longitude only.

Submission

Submit your repository, which should have **git initialized** (ie. including `.git` folder), in a zip file and share your BigQuery dataset which contains the output tables to `data_members@foodpanda.com` by giving the email group the *BigQuery User* and *BigQuery Data Viewer* role. Please also share the private repository with the user `aaronmak`.