

Start Lab

00:45:00

Explore and create ML datasets

45 minutes Free ★★★★★ [Rate Lab](#)

- Overview
- Setup
- Deployment Manager
- Enable APIs
- Launch AI Platform Notebooks
- Explore and create datasets
- End your lab

Overview

Duration is 1 min

Use this lab to explore data corresponding to taxi rides in New York City to build a Machine Learning model in support of a fare-estimation tool.

What you learn

In this lab, you will:

- Access and explore a public BigQuery dataset on NYC Taxi Cab rides
- Visualize your dataset using the Seaborn library
- Inspect and clean-up the dataset for future ML model training
- Create a benchmark to judge future ML model performance off of

Setup

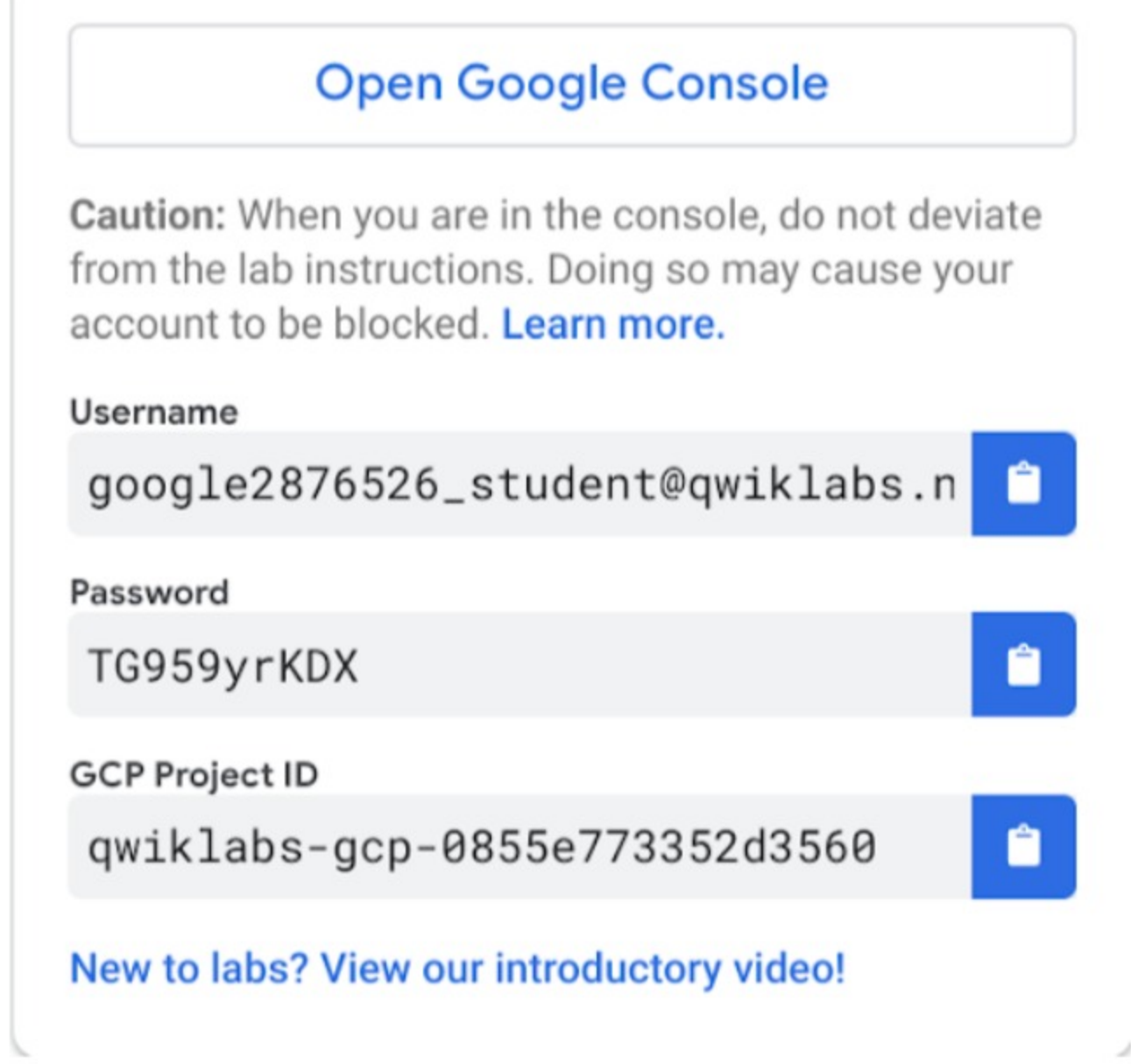
For each lab, you get a new Google Cloud project and set of resources for a fixed time at no cost.

1. Make sure you signed into Qwiklabs using an **incognito window**.
2. Note the lab's access time (for example, **02:00:00**) and make sure you can finish in that time block.

There is no pause feature. You can restart if needed, but you have to start at the beginning.

3. When ready, click **START LAB**.

4. Note your lab credentials. You will use them to sign in to the Google Cloud Console.



5. Click **Open Google Console**.
6. Click **Use another account** and copy/paste credentials for **this** lab into the prompts.

If you use other credentials, you'll get errors or **incur charges**.

7. Accept the terms and skip the recovery resource page.

Do not click **End Lab** unless you are finished with the lab or want to restart it. This clears your work and removes the project.

Deployment Manager

This lab is using a deployment manager script to create the Cloud AI Platform instance you will need for this exercise.

The notebook instance will contain the github repository you need to complete this assignment. It should take 2 - 3 minutes for the instance to be ready.

Please wait before launching the Jupyter notebook, otherwise the script may be interrupted and the repository may not be cloned.

Enable APIs

On the **Navigation menu** () , click **APIs & services**.

Scroll down and confirm that your APIs are enabled.

If an API is missing, click **ENABLE APIS AND SERVICES** at the top, search for the API by name, and enable it for your project.

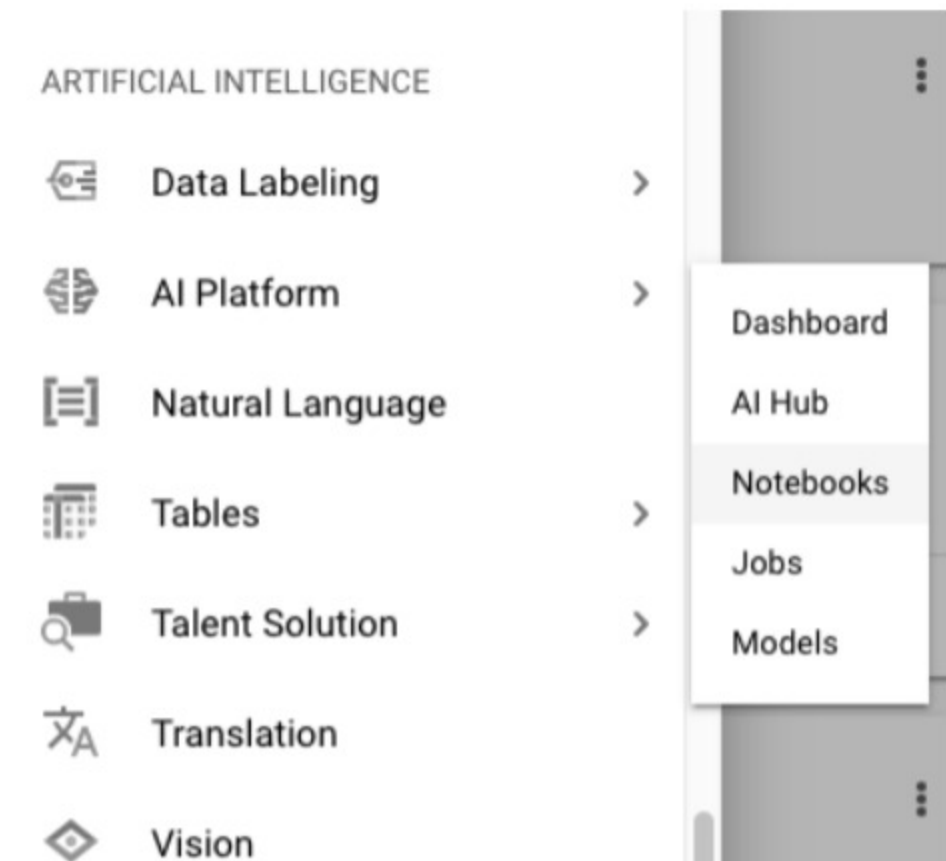
- [BigQuery](#)

Launch AI Platform Notebooks

To launch AI Platform Notebooks:

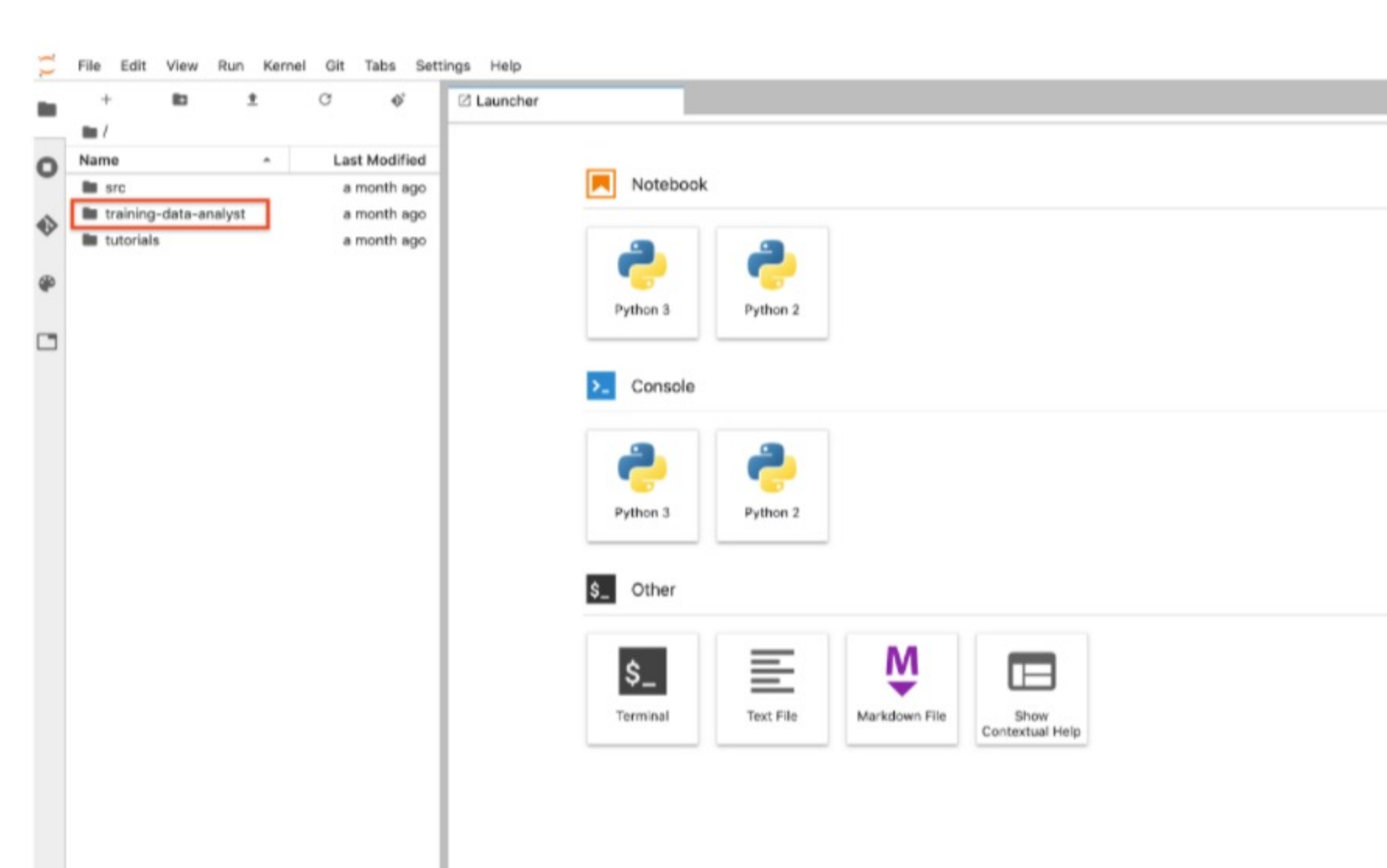
Step 1

Click on the Navigation Menu. Navigate to AI Platforms, then to Notebooks.



Step 2

Click **Open JupyterLab**. A JupyterLab window will open in a new tab. The github repository that you will use should have been already cloned by the Deployment Manager script. The repository is named `training-data-analyst`.



Explore and create datasets

Duration is 30 min

Step 1

In the notebook interface, navigate to **training-data-analyst > courses > machine_learning > deeplive2 > launching_into_ml > solutions** and open **explore_data.ipynb**.

Step 2

In the notebook interface, click on **Edit > Clear All Outputs** (click on Edit, then in the drop-down menu, select Clear All Outputs).

End your lab

When you have completed your lab, click **End Lab**. Qwiklabs removes the resources you've used and cleans the account for you.

You will be given an opportunity to rate the lab experience. Select the applicable number of stars, type a comment, and then click **Submit**.

The number of stars indicates the following:

- 1 star = Very dissatisfied
- 2 stars = Dissatisfied
- 3 stars = Neutral
- 4 stars = Satisfied
- 5 stars = Very satisfied

You can close the dialog box if you don't want to provide feedback.

For feedback, suggestions, or corrections, please use the **Support** tab.

Last Tested Date: 04-30-2020

Last Updated Date: 12-18-2019

©2020 Google LLC All rights reserved. Google and the Google logo are trademarks of Google LLC. All other company and product names may be trademarks of the respective companies with which they are associated.