

---

---

# Cloud Datalab

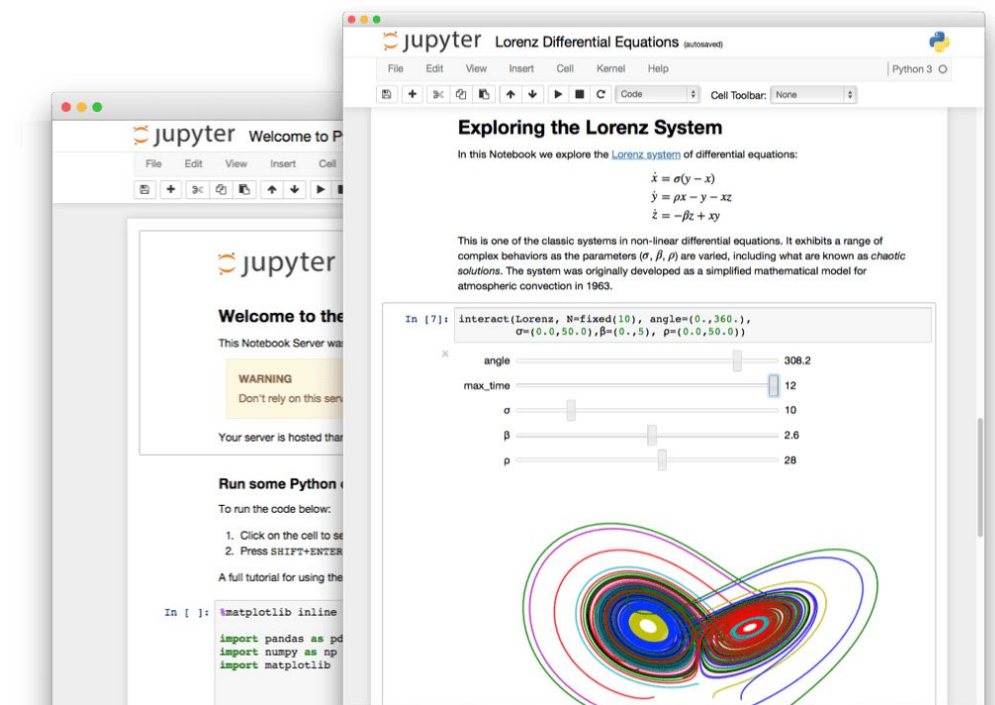
Jeff Liu  
2018/05/12

---

---

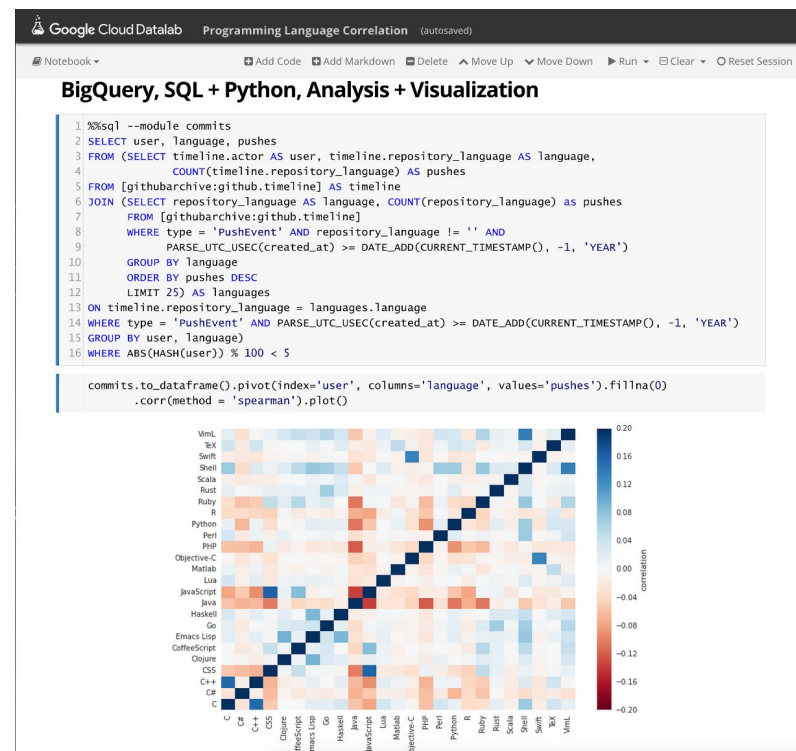
# Introduction

- 基於 Jupyter (IPython)
- 互動式介面, 即時輸出結果



# Introduction (Cont'd)

- 使用 Python
- 整合 Google 服務
  - Google BigQuery
  - Cloud Machine Learning Engine
  - Google Compute Engine
  - Google Cloud Storage



# Setup

1. 建立 Google Cloud Platform 帳號
2. 申請免費試用 (\$300)
3. 建立專案
4. 啟用 Compute Engine API and Source Repositories API
5. 開啟 Cloud Shell
6. 啟動 Datalab

→ Text Note: [goo.gl/vmJGJ4](https://goo.gl/vmJGJ4)

# 啟動APIs

This screenshot shows the Google Cloud Platform search interface. The search bar at the top contains the text 'api'. Below the search bar, a list of search results is displayed. The results include 'Credentials', 'Library', 'Dashboard', 'OAuth consent screen', 'Domain verification', and 'APIs & Services'. The 'APIs & Services' result is highlighted with a red box. The background shows the Google Cloud Platform dashboard with the 'Project info' section visible on the left.

Google Cloud Platform datalab

SEARCH: api

- API Credentials  
APIs & Services
- API Library  
APIs & Services
- API Dashboard  
APIs & Services
- API OAuth consent screen  
APIs & Services > Credentials
- API Domain verification  
APIs & Services
- API APIs & Services

Project info

Project name  
datalab

Project ID  
datalab-203808

Project number  
119772615026

This screenshot shows the 'APIs & Services' page in the Google Cloud Platform. The 'ENABLE APIS AND SERVICES' button is highlighted with a red box. Below the button, the 'Enabled APIs and services' section is visible, showing a list of enabled APIs and services. The background shows the Google Cloud Platform dashboard with the 'Project info' section visible on the left.

Google Cloud Platform datalab

API Dashboard

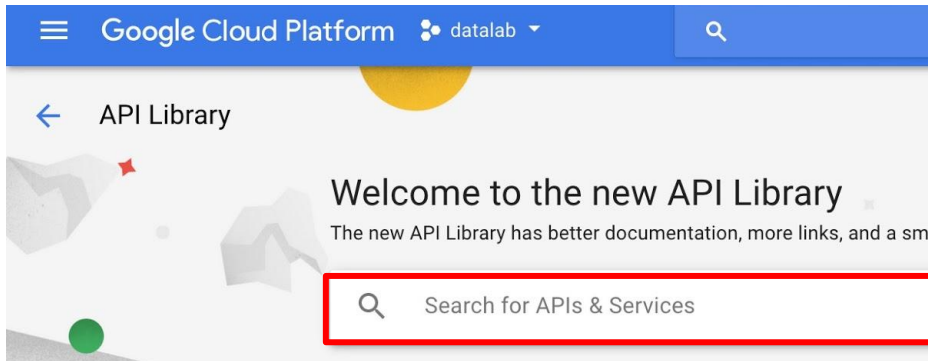
+ ENABLE APIS AND SERVICES

Enabled APIs and services  
Some APIs and services are enabled automatically

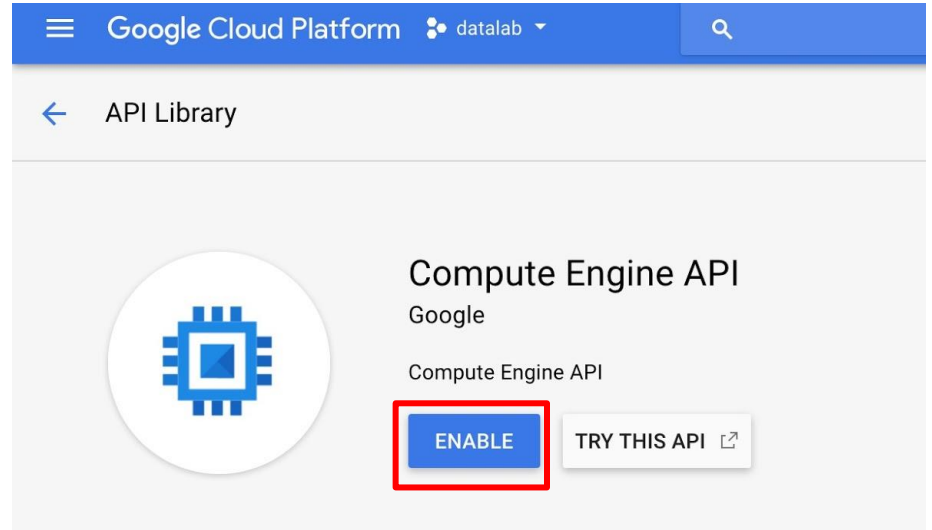
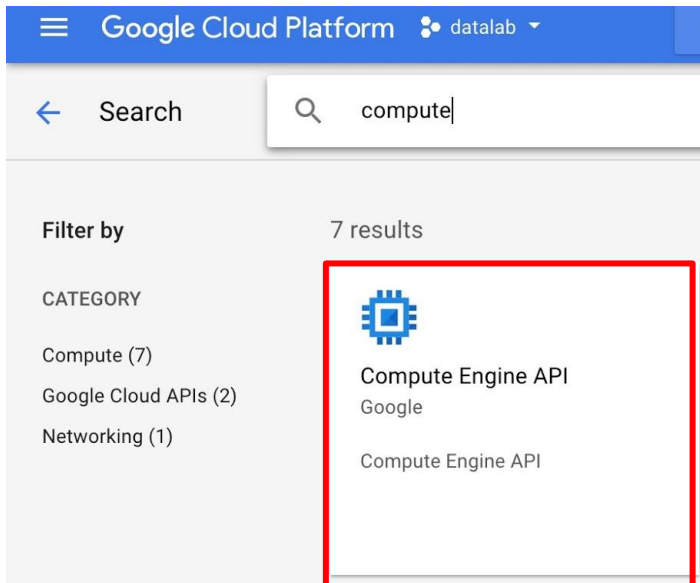
Traffic Errors Median latency

1 hour 6 hours 12 hours 1 day 2 days 4 days 7 days 14 days 30 days

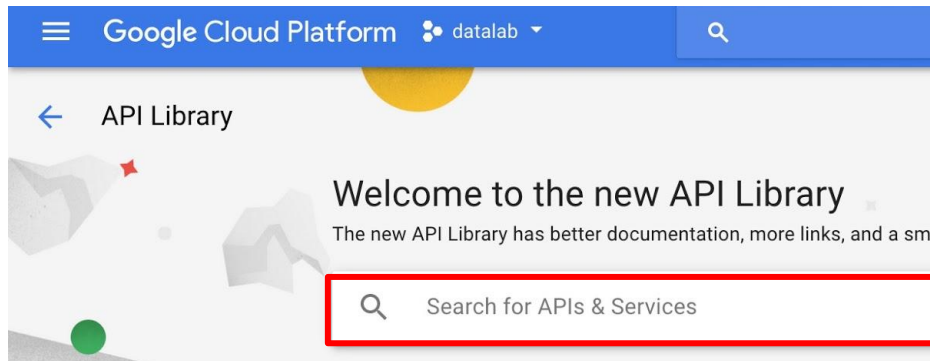
# 開啟Compute Engine API



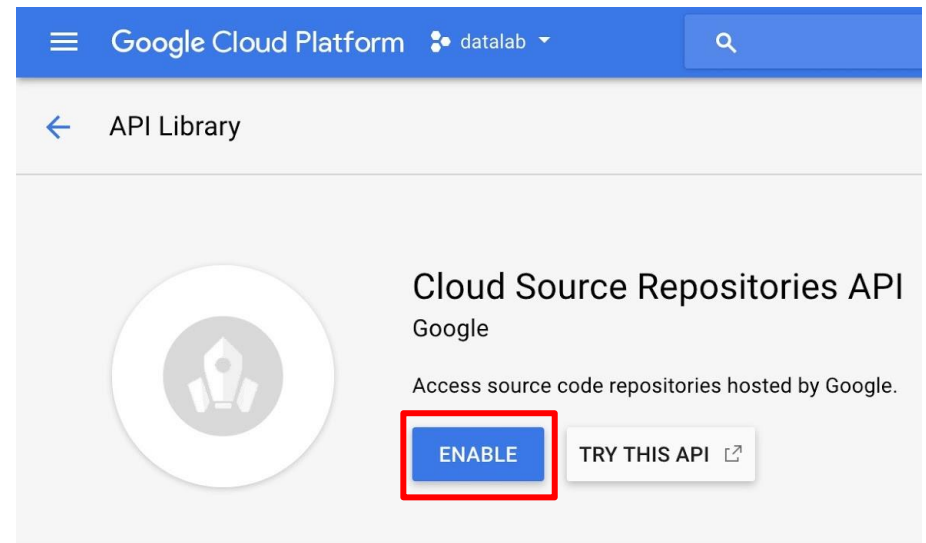
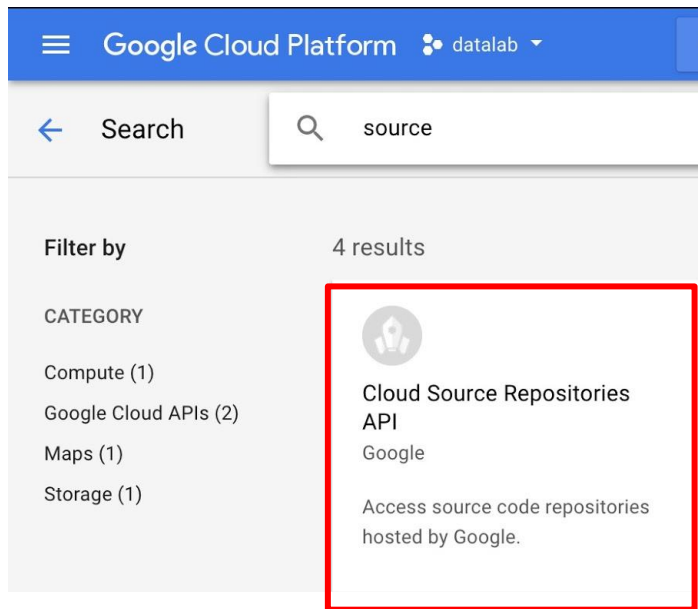
搜尋 compute



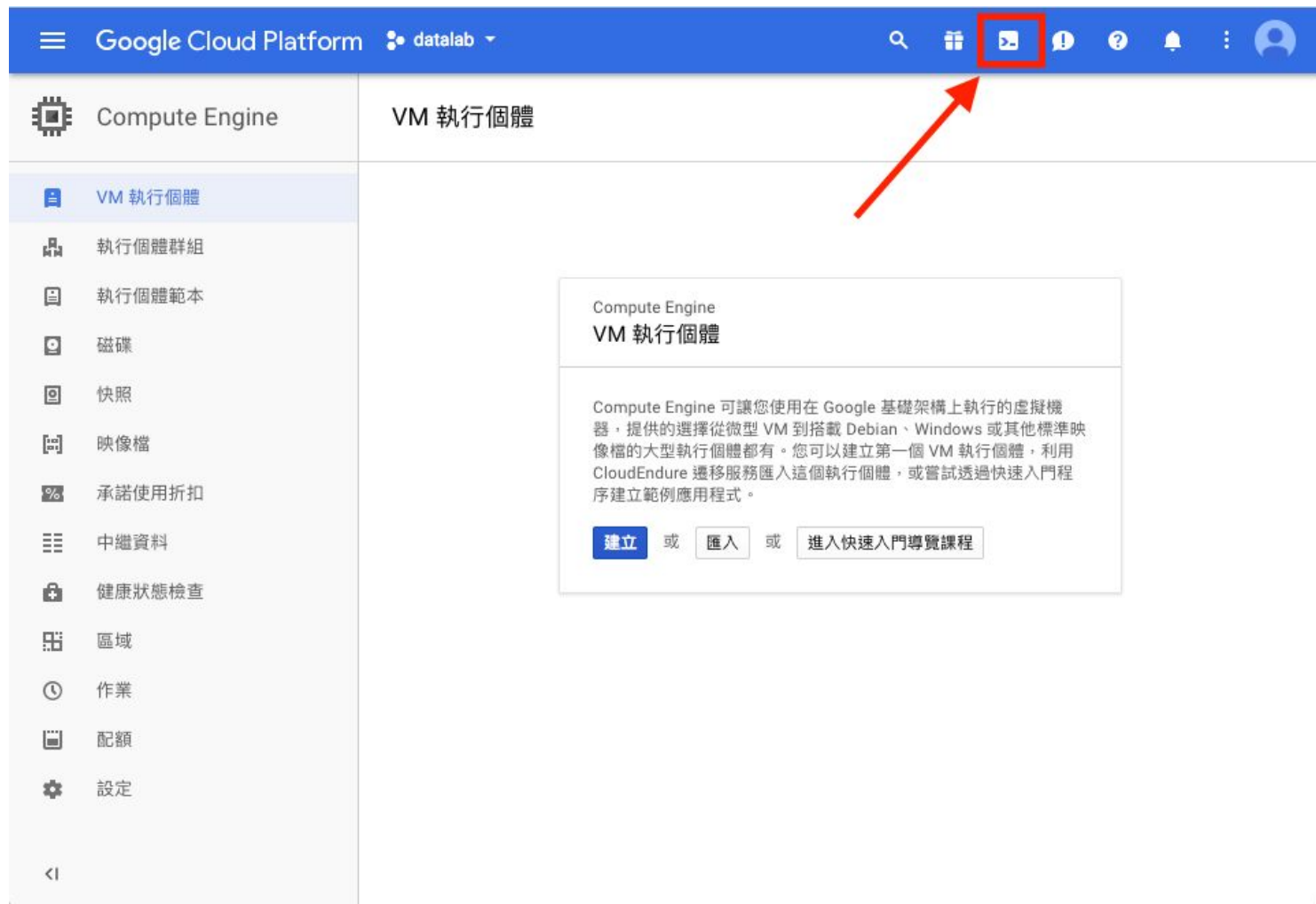
# 開啟 Source Repositories API



搜尋 source



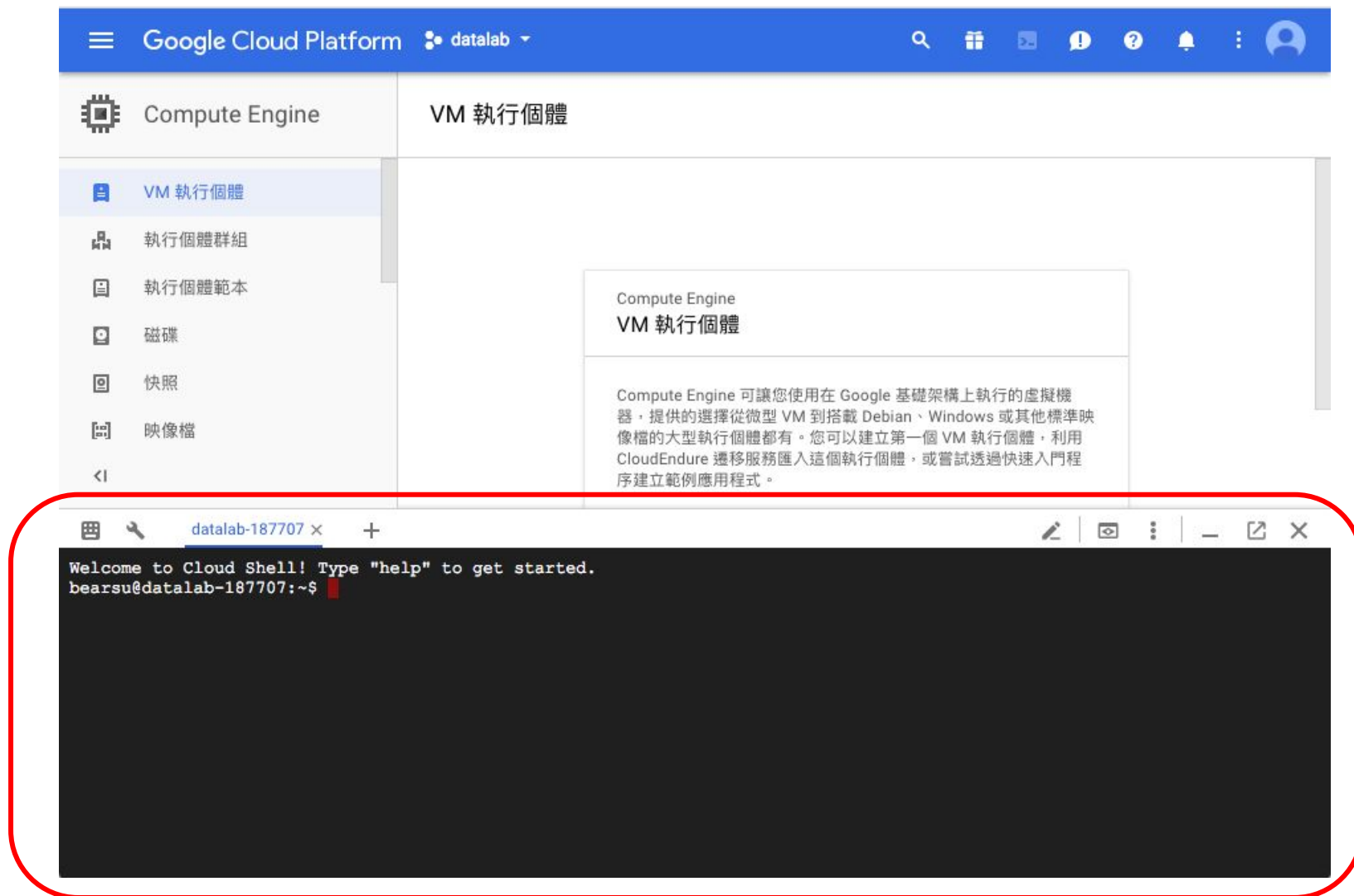
# 開啟 Cloud Shell



The screenshot shows the Google Cloud Platform (GCP) console interface. The top navigation bar is blue and contains the 'Google Cloud Platform' logo, a 'datalab' dropdown menu, a search icon, a 'Cloud Shell' icon (highlighted with a red box and a red arrow), a help icon, a notifications bell, and a user profile icon. The left sidebar is a light gray and contains a list of navigation items: 'Compute Engine' (with a gear icon), 'VM 執行個體' (with a server icon), '執行個體群組' (with a group icon), '執行個體範本' (with a document icon), '磁碟' (with a disk icon), '快照' (with a camera icon), '映像檔' (with a film strip icon), '承諾使用折扣' (with a percentage icon), '中繼資料' (with a list icon), '健康狀態檢查' (with a medical icon), '區域' (with a map icon), '作業' (with a clock icon), '配額' (with a document icon), and '設定' (with a gear icon). The main content area is white and displays the 'VM 執行個體' page. It features a title 'Compute Engine VM 執行個體' and a paragraph of text: 'Compute Engine 可讓您使用在 Google 基礎架構上執行的虛擬機器，提供的選擇從微型 VM 到搭載 Debian、Windows 或其他標準映像檔的大型執行個體都有。您可以建立第一個 VM 執行個體，利用 CloudEndure 遷移服務匯入這個執行個體，或嘗試透過快速入門程序建立範例應用程式。' Below the text are three buttons: '建立' (Build), '或 匯入' (or Import), and '或 進入快速入門導覽課程' (or Enter the quickstart tour).



# Cloud Shell 的介面



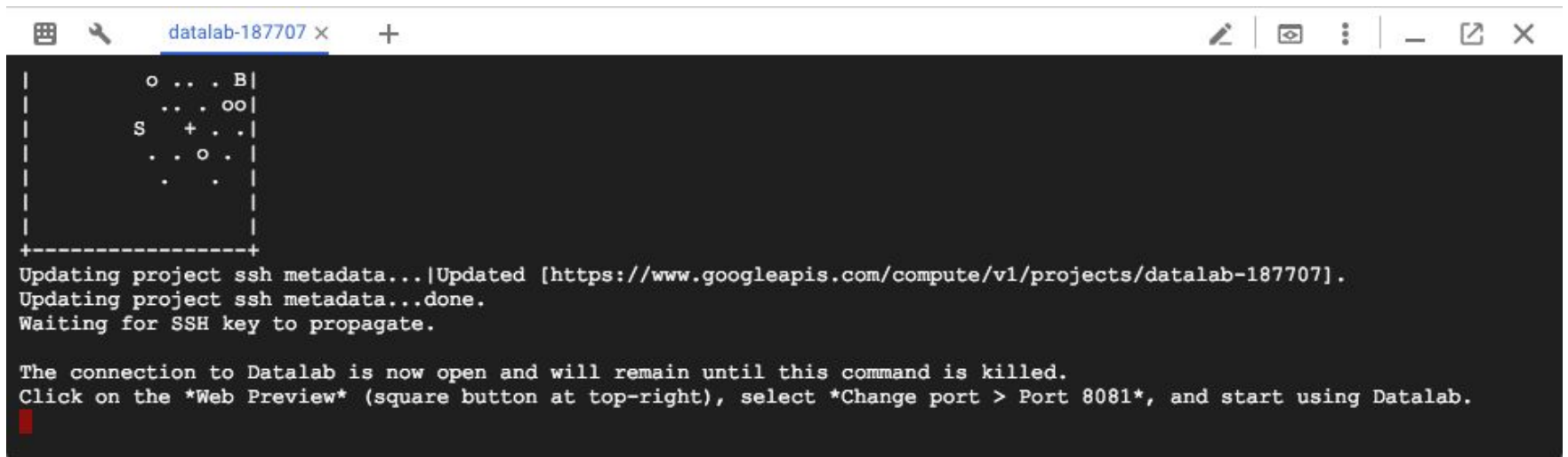
# 啟動 Datalab

輸入指令：

```
datalab create lab-machine \  
    --machine-type n1-standard-4 \  
    --zone asia-east1-b
```

```
Welcome to Cloud Shell! Type "help" to get started.  
bearsu@datalab-187707:~$ datalab create lab-machine \  
>   --machine-type n1-standard-4 \  
>   --zone asia-east1-b  
Creating the instance lab-machine  
Created [https://www.googleapis.com/compute/v1/projects/datalab-187707/zones/asia-east1-b/instances/lab-machine].  
Connecting to lab-machine.  
This will create an SSH tunnel and may prompt you to create an rsa key pair. To manage these keys, see https://cloud.google.com/compute/docs/instances/adding-removing-ssh-keys  
Waiting for Datalab to be reachable at http://localhost:8081/  
This tool needs to create the directory [/home/bearsu/.ssh] before  
being able to generate SSH keys.  
  
Do you want to continue (Y/n)? y
```

# Datalab 啟動成功



A terminal window titled "datalab-187707" with standard window controls. The terminal output shows a ASCII art logo for Datalab, followed by status messages: "Updating project ssh metadata...[Updated [https://www.googleapis.com/compute/v1/projects/datalab-187707]].", "Updating project ssh metadata...done.", and "Waiting for SSH key to propagate.". A final message states: "The connection to Datalab is now open and will remain until this command is killed. Click on the \*Web Preview\* (square button at top-right), select \*Change port > Port 8081\*, and start using Datalab." A small red cursor is visible at the bottom left of the terminal area.

```
|      o . . . B|  
|      . . . oo|  
|  S    + . . |  
|      . . o . |  
|      . . .   |  
+-----+  
Updating project ssh metadata...[Updated [https://www.googleapis.com/compute/v1/projects/datalab-187707]].  
Updating project ssh metadata...done.  
Waiting for SSH key to propagate.  
  
The connection to Datalab is now open and will remain until this command is killed.  
Click on the *Web Preview* (square button at top-right), select *Change port > Port 8081*, and start using Datalab.  
█
```

# 依訊息開啟 Datalab 網頁介面

The screenshot displays the Google Cloud Platform (GCP) console interface for Datalab. The top navigation bar shows 'Google Cloud Platform' and 'datalab'. The left sidebar lists various resources: Compute Engine, VM 執行個體 (selected), 執行個體群組, 執行個體範本, 磁碟, 快照, and 映像檔. The main content area is titled 'VM 執行個體' and shows a table of VM instances. A single instance named 'lab-machine' is listed in the 'asia-east1-b' region with internal IP 10.140.0.2 and external IP 35.185.131.172. Below the table, a terminal window for 'datalab-187707' is open, showing the process of updating SSH metadata and establishing a connection to Datalab. A red arrow points from the terminal output to a context menu that appears over the terminal, listing port options: 通訊埠 8080, 通訊埠 8081 (highlighted with a red box), 通訊埠 8082, 通訊埠 8083, and 通訊埠 8084. The menu also includes options for '功能預覽 8080', '變更通訊埠', and '關於網頁預覽'.

Google Cloud Platform datalab

Compute Engine

VM 執行個體

執行個體群組

執行個體範本

磁碟

快照

映像檔

VM 執行個體

篩選 VM 執行個體

名稱	區域	建議	內部 IP	外部 IP	連線
lab-machine	asia-east1-b		10.140.0.2	35.185.131.172	SSH

datalab-187707

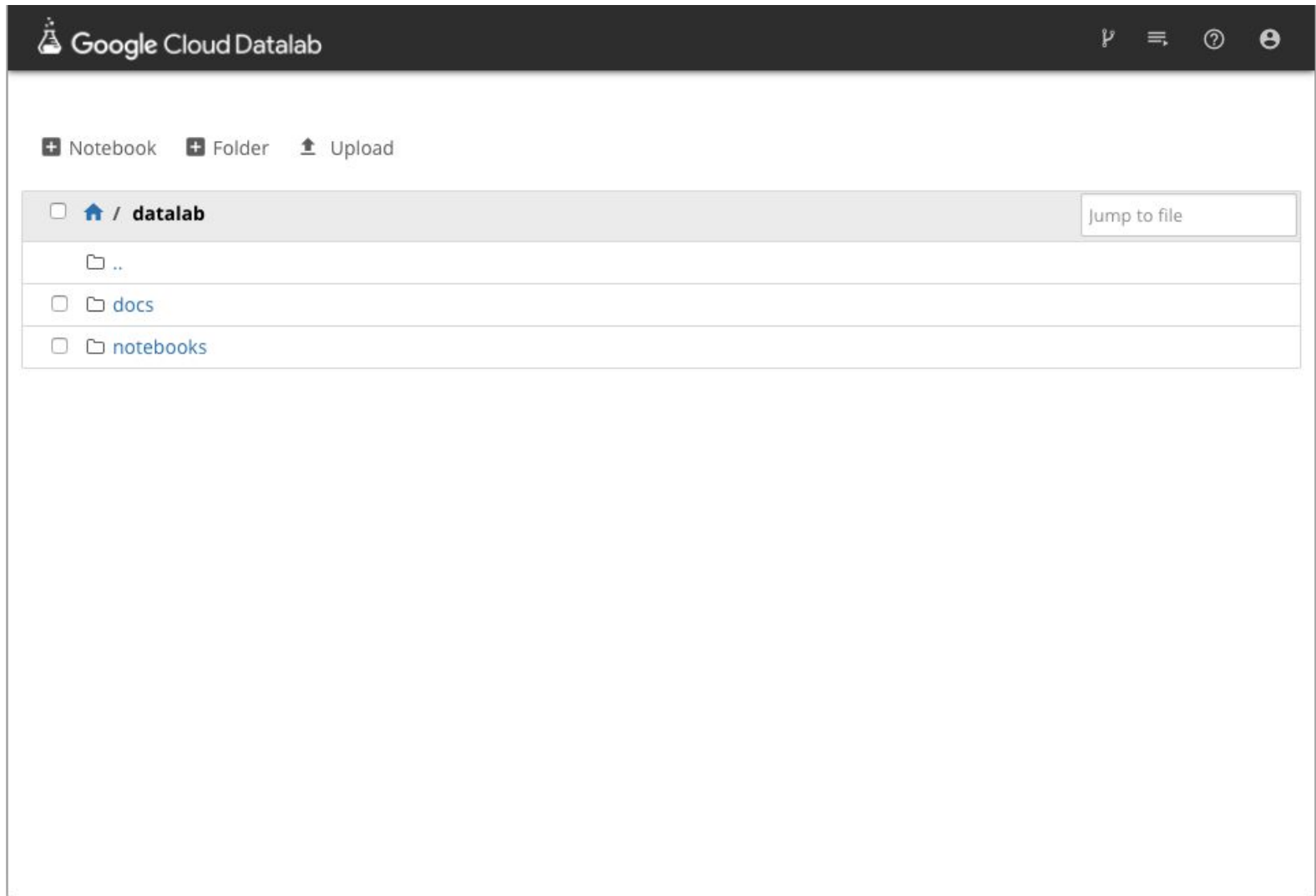
```
|
o . . . B|
. . . oo|
S + . .
. . o .
. . .
+-----+
Updating project ssh metadata...[Updated [https://www.googleapis.com/compute/v1/project/datalab-187707].
Updating project ssh metadata...done.
Waiting for SSH key to propagate.

The connection to Datalab is now open and will remain until this command is killed.
Click on the *Web Preview* (square button at top-right), select *Change port > Port 8081*, and start using Datalab.
```

功能預覽 8080  
變更通訊埠  
關於網頁預覽

通訊埠 8080  
通訊埠 8081  
通訊埠 8082  
通訊埠 8083  
通訊埠 8084

# Datalab 介面



# Datalab 基本指令

重新連上：

```
datalab connect lab-machine
```

停止：

```
datalab stop lab-machine
```

刪除：

```
datalab delete lab-machine
```

刪除(保留硬碟)：

```
datalab delete --keep-disk lab-machine
```

# 建立 Notebook

 Google Cloud Datalab

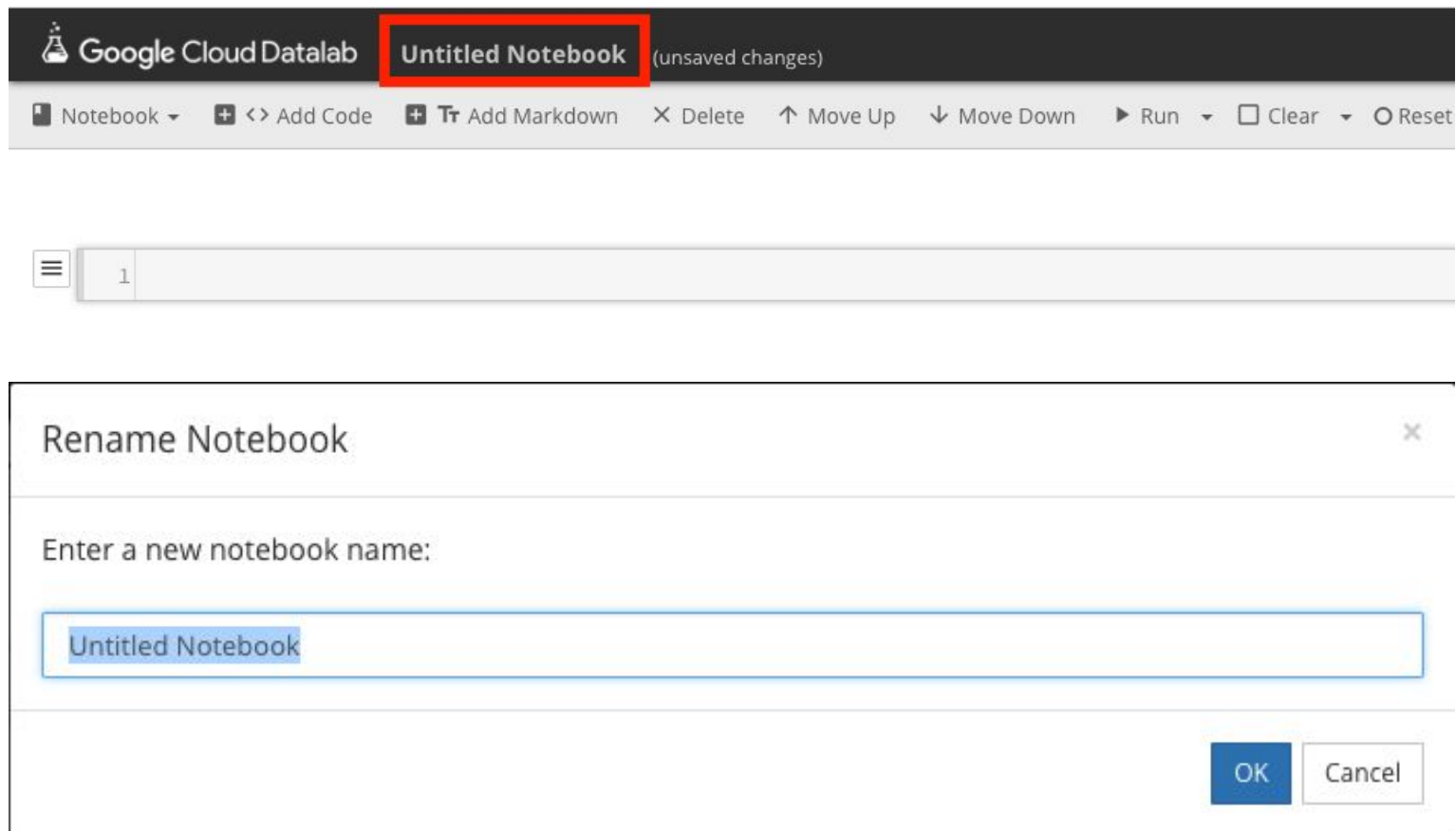
 Notebook  Folder  Upload

☐ [home](#) / [datalab](#) / **notebooks**

☐ [..](#)

☐ [README.md](#)

# 編輯名稱

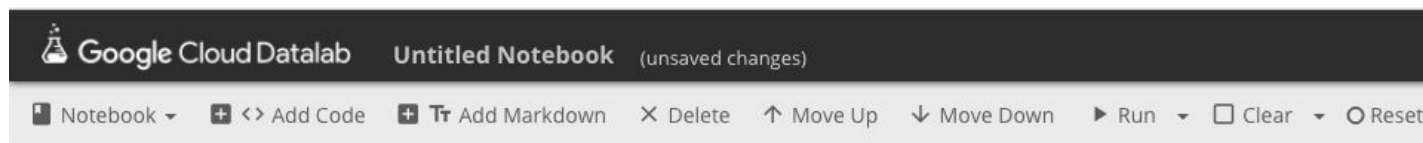


The screenshot displays the Google Cloud Datalab interface. At the top, a dark header bar contains the Google Cloud Datalab logo and the notebook title 'Untitled Notebook' (unsaved changes), which is highlighted with a red rectangle. Below the header, a toolbar includes buttons for 'Notebook', 'Add Code', 'Add Markdown', 'Delete', 'Move Up', 'Move Down', 'Run', 'Clear', and 'Reset'. The main workspace shows a single code cell labeled '1'. A 'Rename Notebook' dialog box is open in the foreground, featuring a title bar with a close button, a label 'Enter a new notebook name:', and a text input field containing 'Untitled Notebook'. The dialog has 'OK' and 'Cancel' buttons at the bottom right.



# 基本語法 - Python

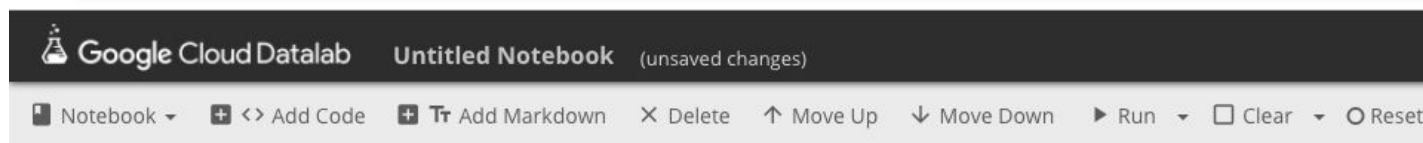
- 輸入 python 程式碼，點擊「Run」或是「Ctrl + Enter」執行程式碼區塊。



Google Cloud Datalab Untitled Notebook (unsaved changes)

Notebook + <> Add Code + Tr Add Markdown X Delete ↑ Move Up ↓ Move Down ▶ Run □ Clear ○ Reset

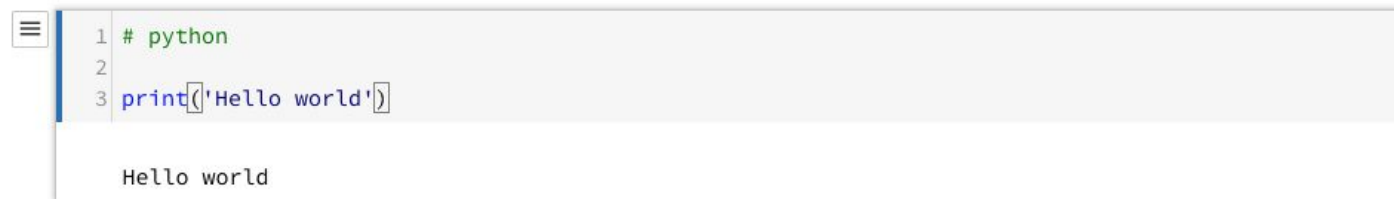
```
1 # python
2
3 print('Hello world')
```



Google Cloud Datalab Untitled Notebook (unsaved changes)

Notebook + <> Add Code + Tr Add Markdown X Delete ↑ Move Up ↓ Move Down ▶ Run □ Clear ○ Reset

```
1 # python
2
3 print('Hello world')
```

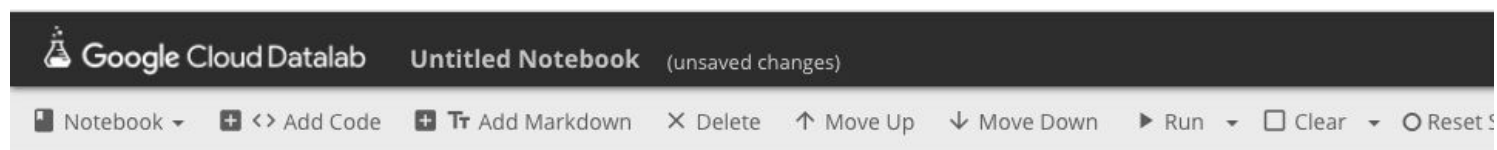


1 # python
2
3 print('Hello world')

Hello world

# 基本語法 - shell

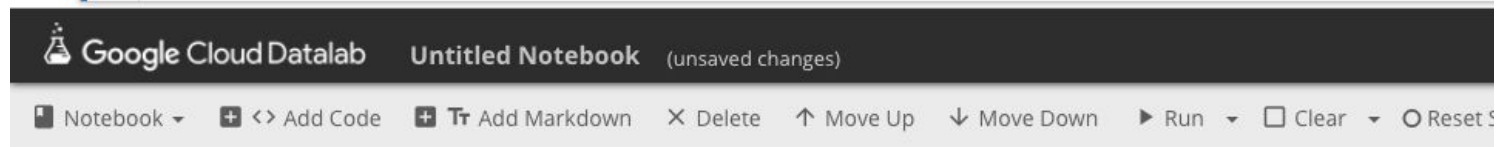
- shell 指令前面要加上「！」



Google Cloud Datalab Untitled Notebook (unsaved changes)

Notebook < + <> Add Code + Tr Add Markdown X Delete ↑ Move Up ↓ Move Down ▶ Run ▢ Clear ○ Reset S

```
1 # shell
2
3 !pwd
```



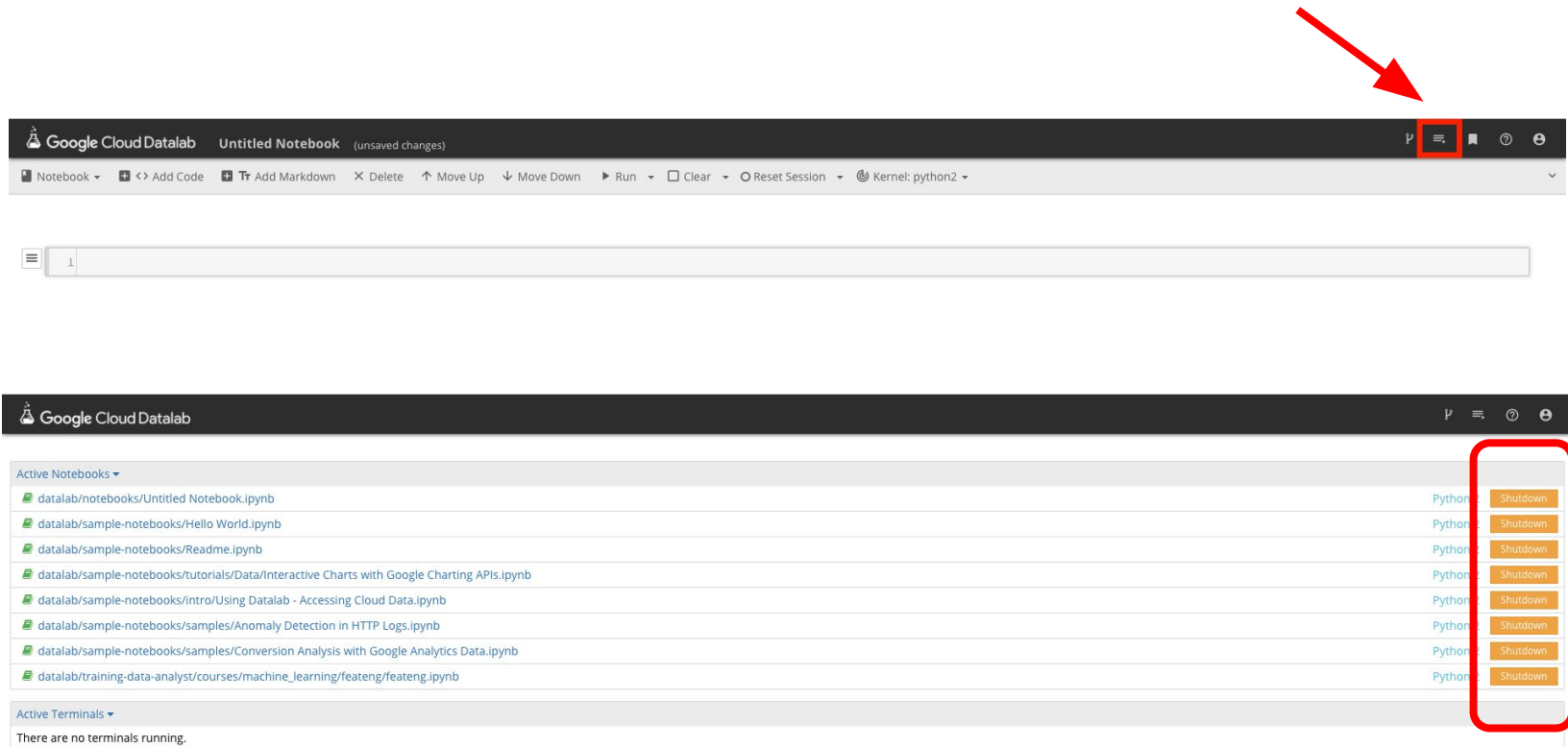
Google Cloud Datalab Untitled Notebook (unsaved changes)

Notebook < + <> Add Code + Tr Add Markdown X Delete ↑ Move Up ↓ Move Down ▶ Run ▢ Clear ○ Reset S

```
1 # shell
2
3 !pwd
```

/content/datalab/notebooks

# 停止 Sessions



The image shows two screenshots of the Google Cloud Datalab interface. The top screenshot shows the top bar of the notebook editor with a red box around the menu icon (three horizontal lines) and a red arrow pointing to it. The bottom screenshot shows the 'Active Notebooks' section, which is a table listing running notebooks. A red box highlights the 'Shutdown' button for each notebook in the list.

Active Notebooks		
datalab/notebooks/Untitled Notebook.ipynb	Python	Shutdown
datalab/sample-notebooks/Hello World.ipynb	Python	Shutdown
datalab/sample-notebooks/Readme.ipynb	Python	Shutdown
datalab/sample-notebooks/tutorials/Data/Interactive Charts with Google Charting APIs.ipynb	Python	Shutdown
datalab/sample-notebooks/intro/Using Datalab - Accessing Cloud Data.ipynb	Python	Shutdown
datalab/sample-notebooks/samples/Anomaly Detection in HTTP Logs.ipynb	Python	Shutdown
datalab/sample-notebooks/samples/Conversion Analysis with Google Analytics Data.ipynb	Python	Shutdown
datalab/training-data-analyst/courses/machine_learning/feateng/feateng.ipynb	Python	Shutdown

Active Terminals

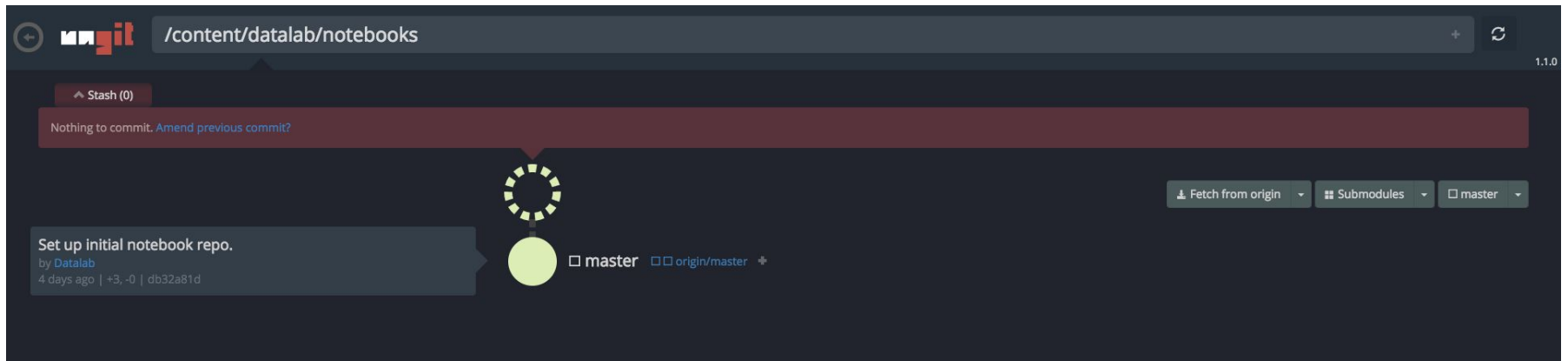
There are no terminals running.

# 透過 git 下載 Notebooks

- 點擊右上角分支圖示

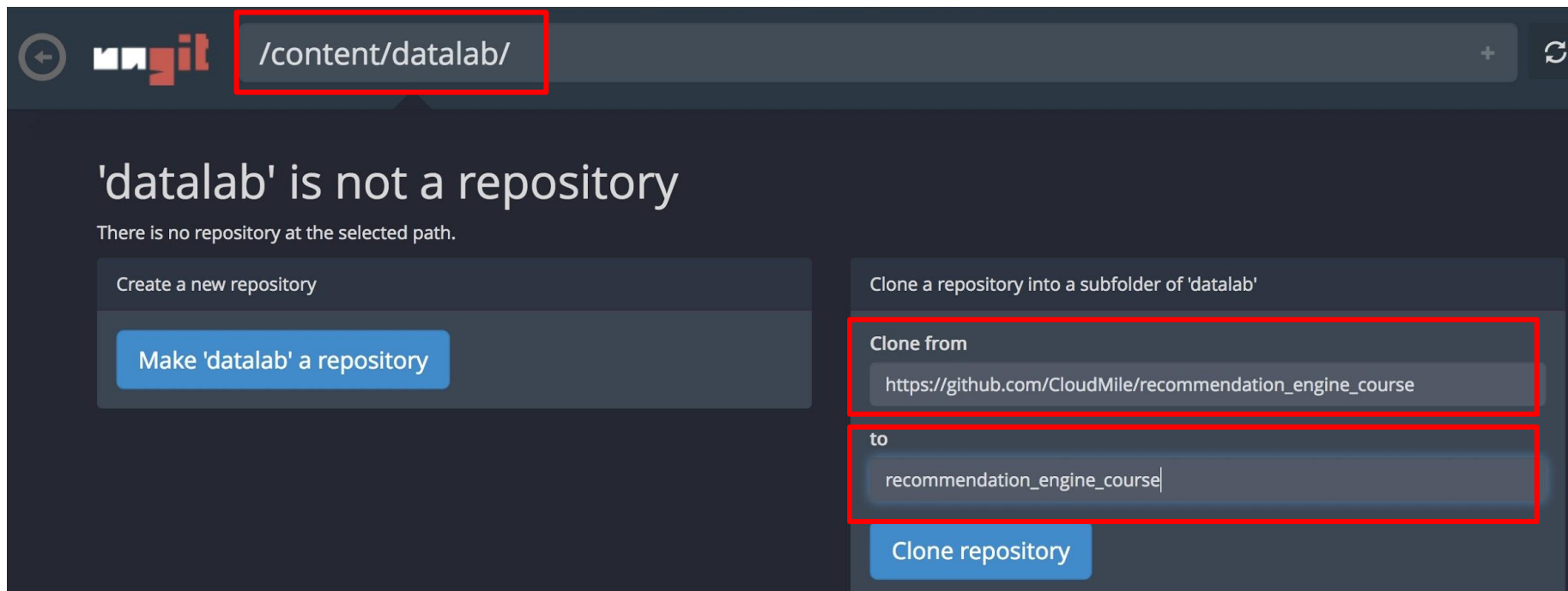


# 透過 git 下載 Notebooks



# 透過 git 下載 Notebooks

- 修改工作路徑後按下 Enter
- 於「Clone from」欄位填入：
  - [https://github.com/CloudMile/recommendation\\_engine\\_course](https://github.com/CloudMile/recommendation_engine_course)
- 於「to」欄位填入：recommendation\_engine\_course
- 點擊「Clone repository」



# 透過 git 下載 Notebooks

- 已新增檔案



# Reference

- <https://cloud.google.com/datalab/?hl=zh-tw>
- <https://cloud.google.com/datalab/docs/?hl=zh-tw>
- <https://github.com/google/datalab>
- <http://jupyter.org/>
- <https://codelabs.developers.google.com/> search `datalab`
- <https://cloud.google.com/datalab/docs/reference/command-line/commands>