How to Import and Export Datasets in Google Colab

Download and Upload files in Colab from Local system and Google Drive

Thought Partner for Data Jul 25, 2020 · 3 min read ★

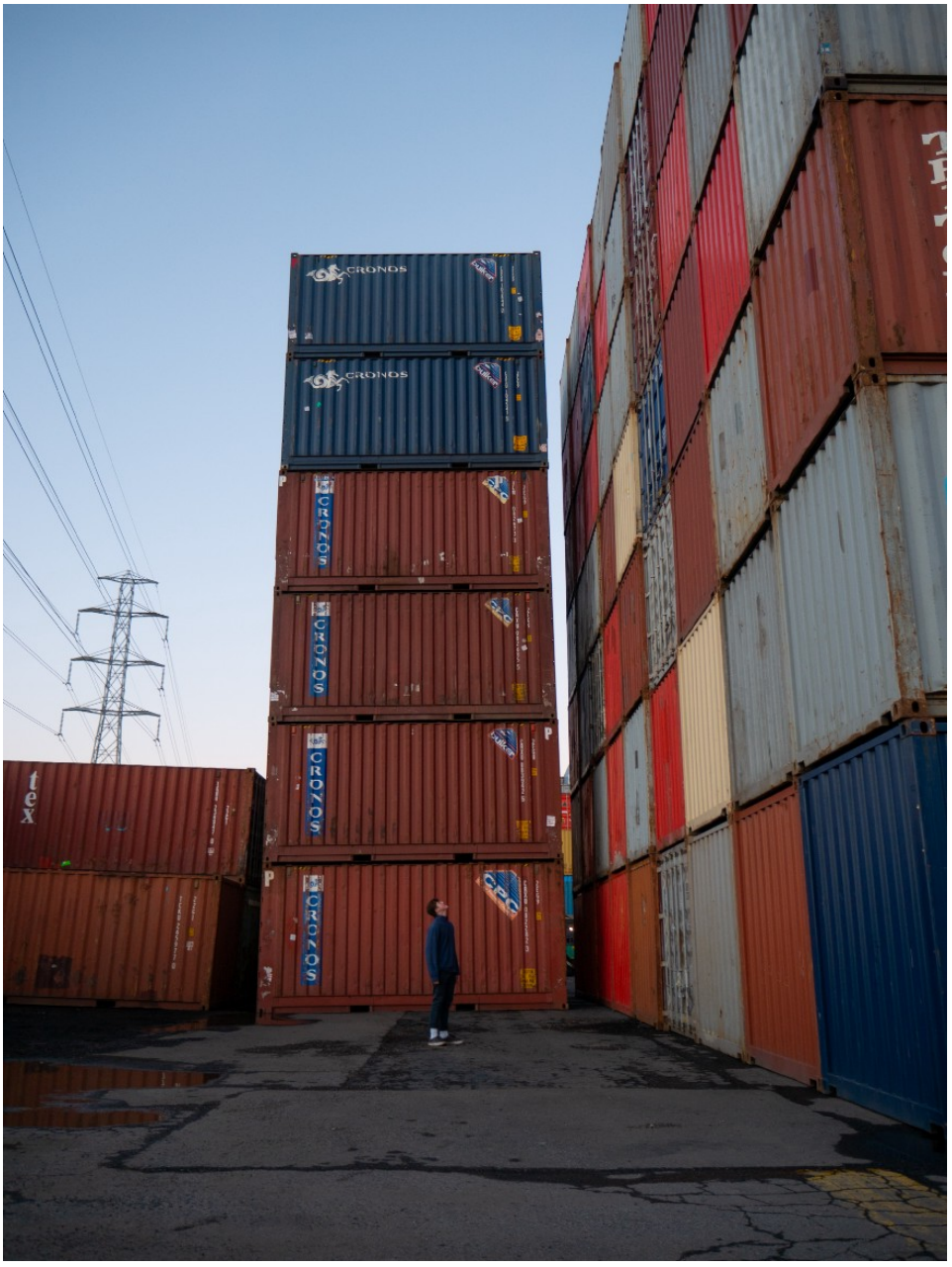


Photo by Pat Whelen on Unsplash

Google Colaboratory, known as Colab, is a free Jupyter Notebook environment with many pre-installed libraries like Tensorflow, Pytorch, Keras, OpenCV, and many more. It is one of the cloud services that support GPU and TPU for free. Importing a dataset and training models on the data in the Colab facilitate coding experience. We can apply different ways to import and download data in Colab. In this tutorial, I will discuss my experience in:

1. Importing data from Google Drive
2. Importing and downloading data in the local system

Mounting Google Drive

We can access files in drive using mounting Google Drive. Mounting Drive into the Colab meaning that setting up the google drive account as a virtual drive so that we can access the resources of the drive just like a local hard drive.

Step 1 To connect Google Drive (GDrive) with Colab, execute the following two lines of code in Colab:

```
from google.colab import drive

drive.mount("/content/gdrive")
```

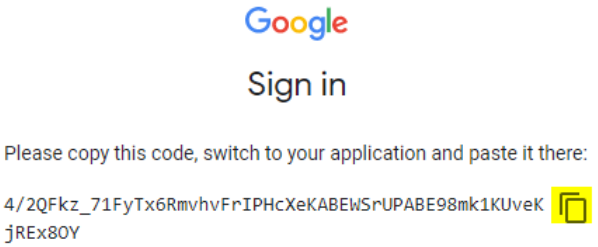
Running the shell will return a URL link and ask for an authorization code:

```
1 from google.colab import drive
2 drive.mount("/content/gdrive")
```

```
Go to this URL in a browser: https://accounts.google.com/o/oauth2/auth?cli

Enter your authorization code:
```

Step 2 Follow to the mentioned link, sign in Google account, and copy the authorization code by clicking at highlighted spot:



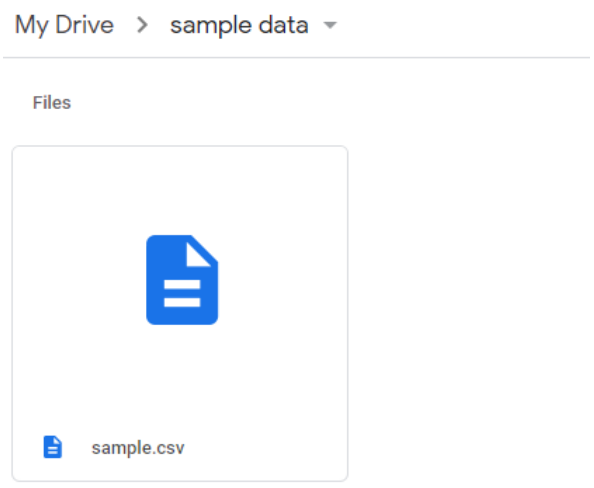
Step 3 Paste the authorization code in the shell and finally, Google Drive will be mounted at */content/gdrive*. Note that, files in the drive are under the folder */content/gdrive/My Drive/*. Now, we can import files in GDrive using a library like Pandas.

```
1 from google.colab import drive
2 drive.mount("/content/gdrive")
```

```
Go to this URL in a browser: https

Enter your authorization code:
··········
Mounted at /content/gdrive
```

Step 4 For instance, we have a dataset (sample.csv) in */My Drive/sample data* folder in GDrive.



Executing the following two lines of code will import the data into Colab:

```
import pandas as pd

pd.read_csv('/content/gdrive/My Drive/sample data/sample.csv')
```

Importing Data from Local System

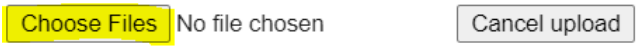
Step1 Run the following two lines of code to import data from the local system.

```
from google.colab import files

uploaded = files.upload()
```

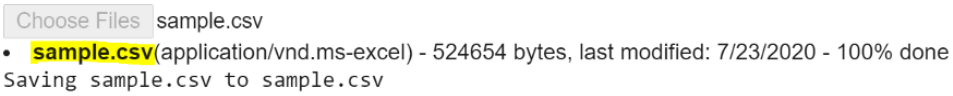
Executing the shell will invoke a browse button:

```
1 from google.colab import files
2 uploaded = files.upload()
```



Step 2 Browsing directories in the local system, we can upload data into Colab:

```
1 from google.colab import files
2 uploaded = files.upload()
```



Finally, we can read the data using a library like Pandas:

```
pd.read_csv("sample.csv")
```

Downloading Data into Google Drive and Local System

We can easily download data into local directories by executing the following two lines of codes given the dataset is already in CSV format:

```
from google.colab import files

files.download('sample.csv')
```

A pandas dataframe can be downloaded executing the following code. Given that the name of the dataframe is “sample”, then to save the file in Google Drive using one line code:

```
sample.to_csv('sample.csv')
```

Finally, to download in the local drive execute the following single line code:

```
files.download('sample.csv')
```

As an additional source, you can also read this blog to understand more about how to deal with files in Colab.

Happy Reading !!!

Google Colab | Mounting Google Drive | Data Science | Machine Learning | Python