How to read csv to dataframe in Google Colab

Asked 5 years, 7 months ago Modified 1 year, 6 months ago Viewed 182k times



I am trying to read a csv file which I stored locally on my machine. (Just for additional reference it is titanic data from Kaggle which is here.)

52

From this question and answers I learnt that you can import data using this code which works well from me.



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

()

Where I am lost is how to convert it to dataframe from here. The <u>sample google notebook page</u> listed in the answer above does not talk about it.

I am trying to convert the dictionary uploaded to dataframe using from_dict command but not able to make it work. There is some discussion on converting dict to dataframe here but the solutions are not applicable to me (I think).

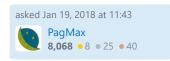
So summarizing, my question is:

How do I convert a csv file stored locally on my files to pandas dataframe on Google Colaboratory?



Share Improve this question Follow





Hey have a look at my answer here, using gspread stackoverflow.com/a/49397059/9017976 - MarshallMa Mar 21, 2018 at 2:25

7 Answers

Sorted by: Highest score (default)



step 1- Mount your Google Drive to Collaboratory

76

```
from google.colab import drive
drive.mount('/content/gdrive')
```



step 2- Now you will see your Google Drive files in the left pane (file explorer). Right click on the file that you need to import and select copy path. Then import as usual in pandas, using this copied path.

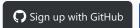


```
import pandas as pd
df=pd.read_csv('gdrive/My Drive/data.csv')
```

Done!













Pandas read_csv should do the trick. You'll want to wrap your uploaded bytes in an io.StringIO since read_csv expects a file-like object.

Here's a full example: https://colab.research.google.com/notebook#fileld=1JmwtF50mSghC-y3-BkvxLan0zYXqCJJf



The key snippet is:





```
import pandas as pd
import io
df = pd.read_csv(io.StringIO(uploaded['train.csv'].decode('utf-8')))
```

Share Improve this answer Follow

answered Jan 19, 2018 at 16:34



- 2 Thanks . @Bob Smith I have used your solution it works . can you also provide how i can use csv files which are already in the google drive ? - Mitu Vinci Mar 7, 2018 at 20:50
- 4 Mount your Google Drive and then read from there, e.g., /content/gdrive/My\ Drive/train.csv . Bob Smith Sep 15, 2018 at 17:15
- Just a small addendum on Bob's excellent comment from 3 years ago: right now it should be /content/drive/My\ Drive/trains.csv e.g. drop the g from gdrive. – Jose R Dec 7, 2021 at 23:23







Colab google: uploading csv from your PC I had the same problem with an excel file (*.xlsx), I solved the problem as the following and I think you could do the same with csv files: - If you have a file in your PC drive called (file.xlsx) then: 1-Upload it from your hard drive by using this simple code:



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

М

Press on (Choose Files) and upload it to your google drive.

2- Then:

```
import io
data = io.BytesIO(uploaded['file.XLSX'])
```

3- Finally, read your file:

```
import pandas as pd
f = pd.read_excel(data , sheet_name = '1min', header = 0, skiprows = 2)
#df.sheet_names
df.head()
```

4- Please, change parameters values to read your own file. I think this could be generalized to read other types of files! Enjoy it!

Share Improve this answer Follow

answered Mar 22, 2018 at 17:51





This worked for me:

```
from google.colab import auth
auth.authenticate_user()

from pydrive.drive import GoogleDrive
from pydrive.auth import GoogleAuth
from oauth2client.client import GoogleCredentials
gauth = GoogleAuth()
gauth.credentials = GoogleCredentials.get_application_default()
drive = GoogleDrive(gauth)

myfile = drive.CreateFile({'id': '!!!YOUR FILE ID!!!'})
myfile.GetContentFile('file.csv')
```

Replace !!!Your FILE ID!!! with the id of the file in google drive (this is the long alphanumeric string that appears when you click on "obtain link to share"). Then you can access file.csv with pandas' read_csv:

```
import pandas as pd
frm = pd.read_csv('file.csv', header=None)
```

Share Improve this answer Follow

answered May 31, 2018 at 20:43















In google colab, you only thing you have to know is the path of the csv file.



(1)

If you follow the steps that I have written below, your problem will be solved:

- 1. First of all, upload the CSV file on your google drive.
- 2. Then, open your google colab notebook and click on the 'Files' icon on the left side of the page.
- 3. Then, click on the 'Google Drive Folder' icon to mount your Google Drive.
- 4. Then, look for the csv file that you uploaded on your google drive (step 1), and copy its path.
- 5. Once you have the path, treat it as an ordinary path and use it in your code. It should look something like this:

```
df = pd.read_csv('/content/drive/MyDrive/File.csv')
```

Share Improve this answer Follow



answered Jun 28, 2021 at 20:08





this worked for me:

_

import pandas as pd import io



df=pd.read_csv(io.StringIO(uploaded['Filename.CSV'].decode('ISO-8859-1')))
df



Share Improve this answer Follow

answered Feb 21, 2022 at 14:57





Alternatively, you can use github to import files also. You can take this as an example: https://drive.google.com/file/d/1D6ViUx8 ledfBqcxHCrFPcqBvNZitwCs/view?usp=sharing



Also google does not persist the file for longer so you may have to run the github snippets time and again.



Share Improve this answer Follow

answered Feb 13, 2018 at 8:21



1

