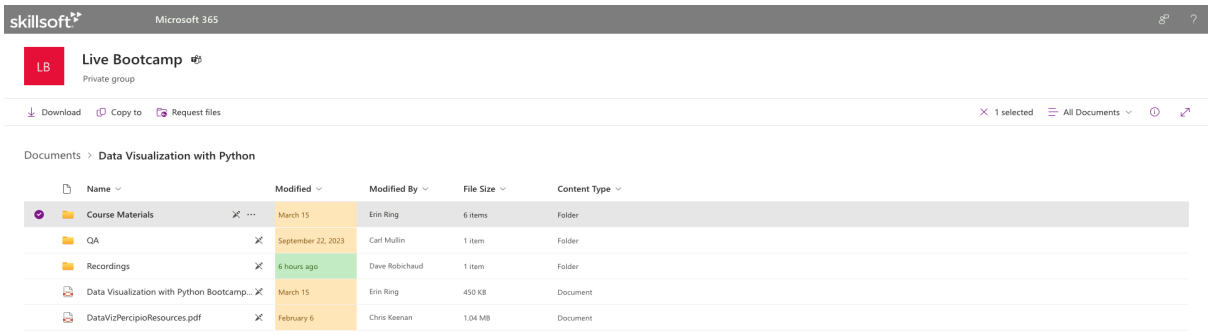
Google Colab Setup

Download files

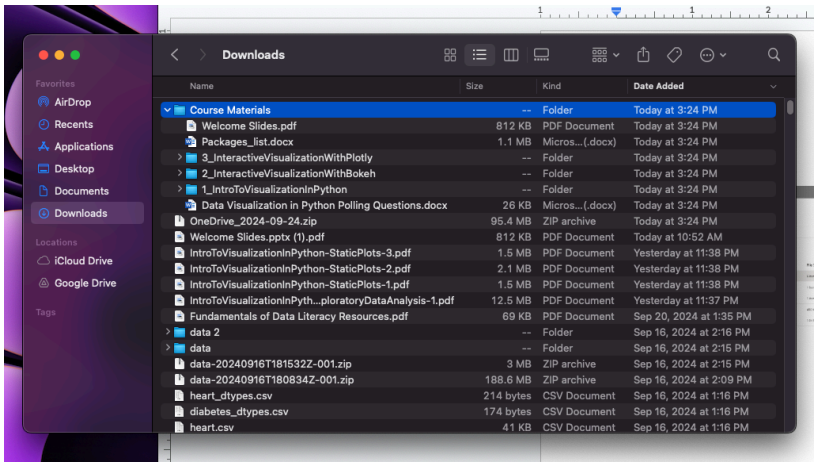
1. Go to Skillsoft SharePoint Site (or Github) to access files

https://skillsoft.sharepoint.com/sites/LiveBootcamp/Shared%20Documents/Forms/AllItems.aspx?id=%2Fsites%2FLiveBootcamp%2FShared%20Documents%2FData%20Visualization%20with%20Python



2. Download Course Materials

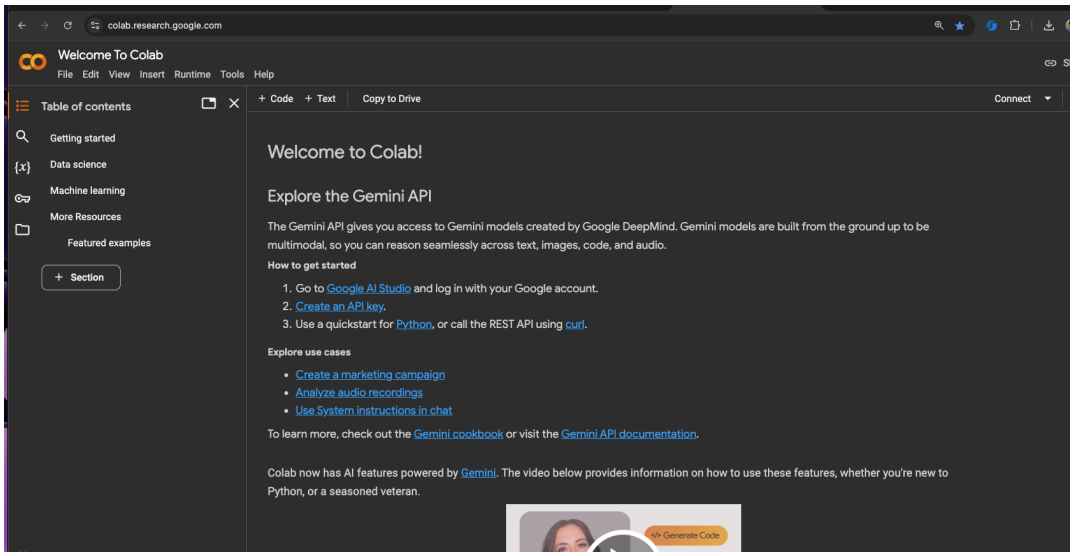
Unzip file if it doesn't automatically do this on download.



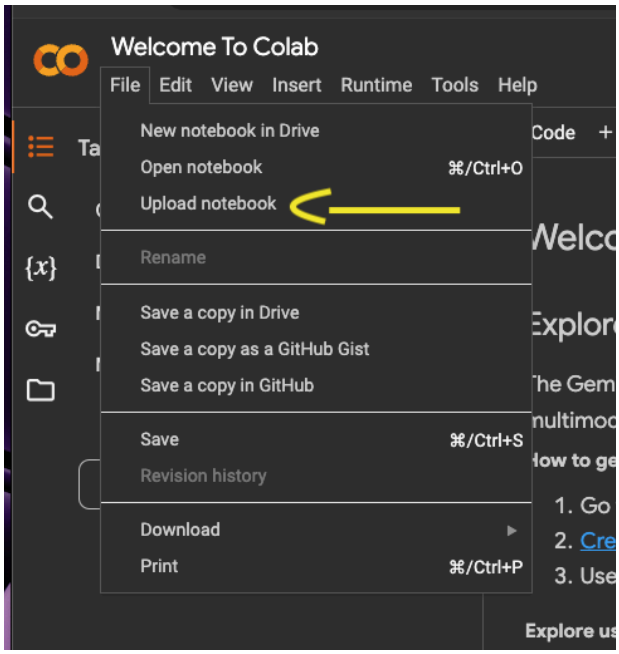
Move files if you like.

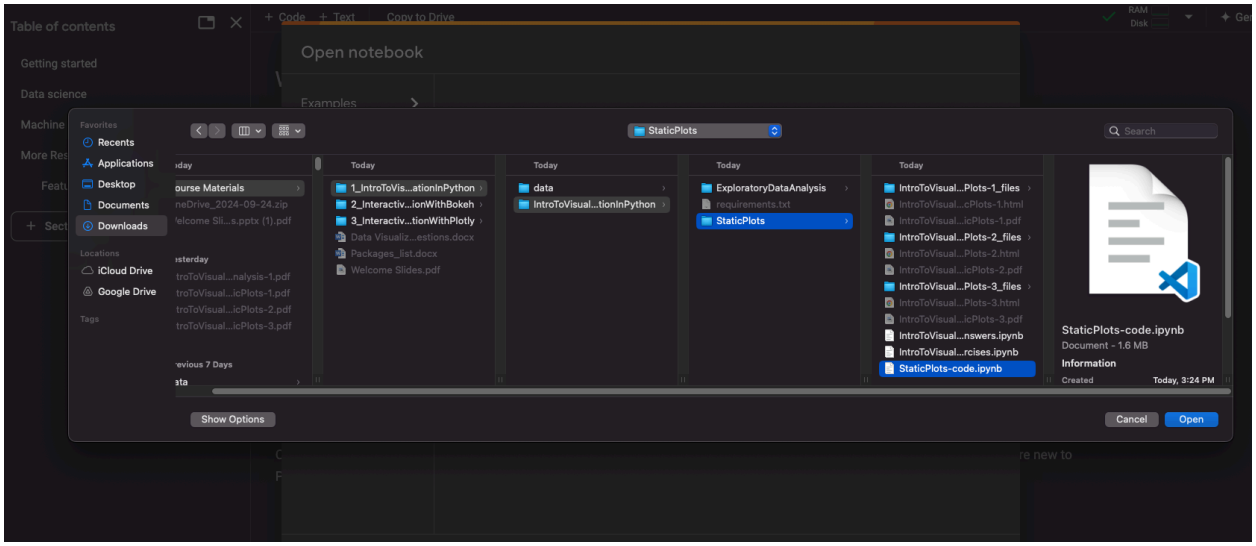
Go to Google Colab

https://colab.research.google.com/

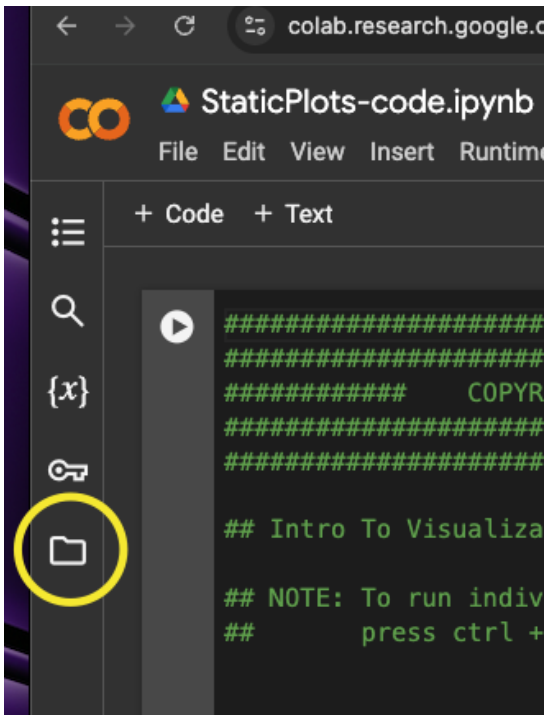


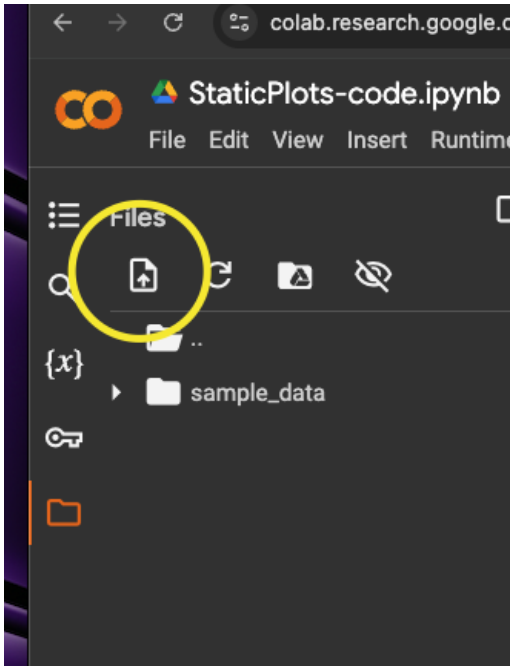
1. Upload Code notebook

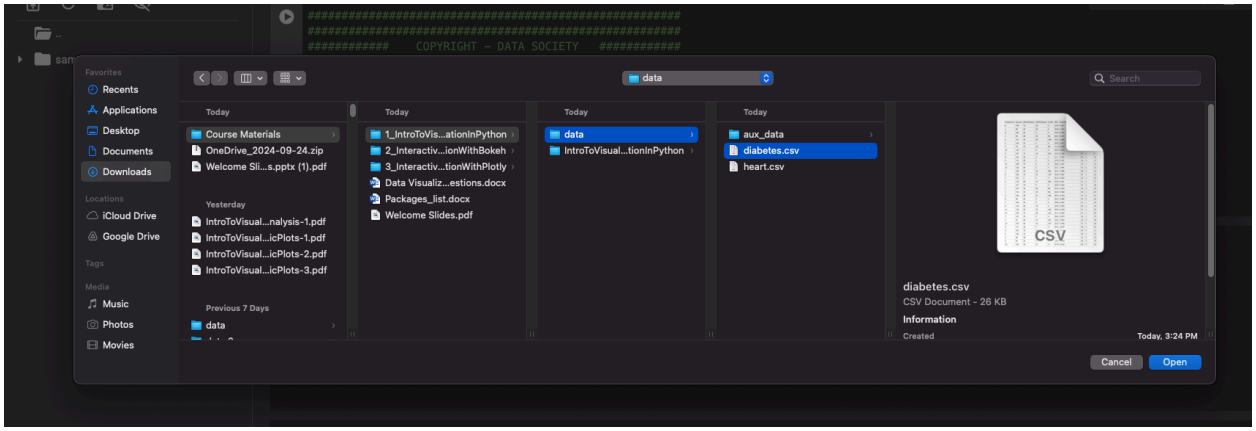




2. Upload Dataset

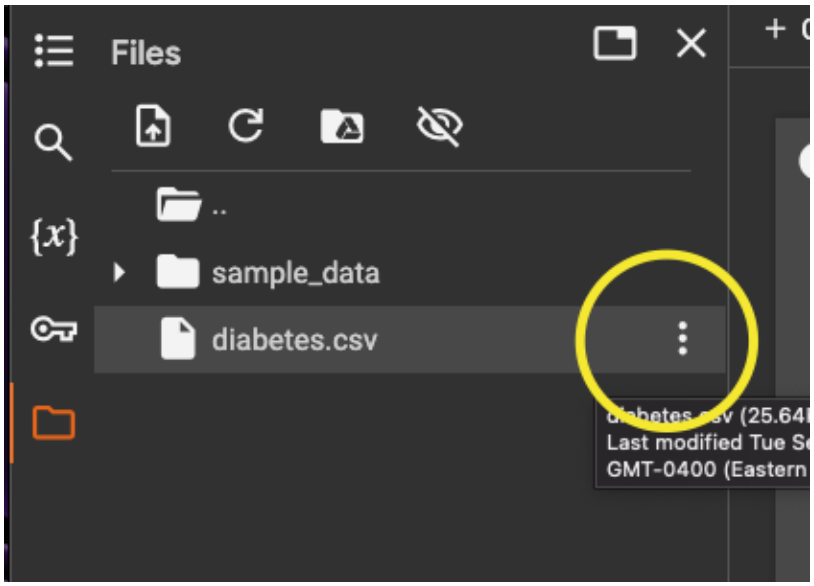



colab.research.google.c
StaticPlots-code.ipynb
File Edit View Insert Runtim
Files
..
sample_data

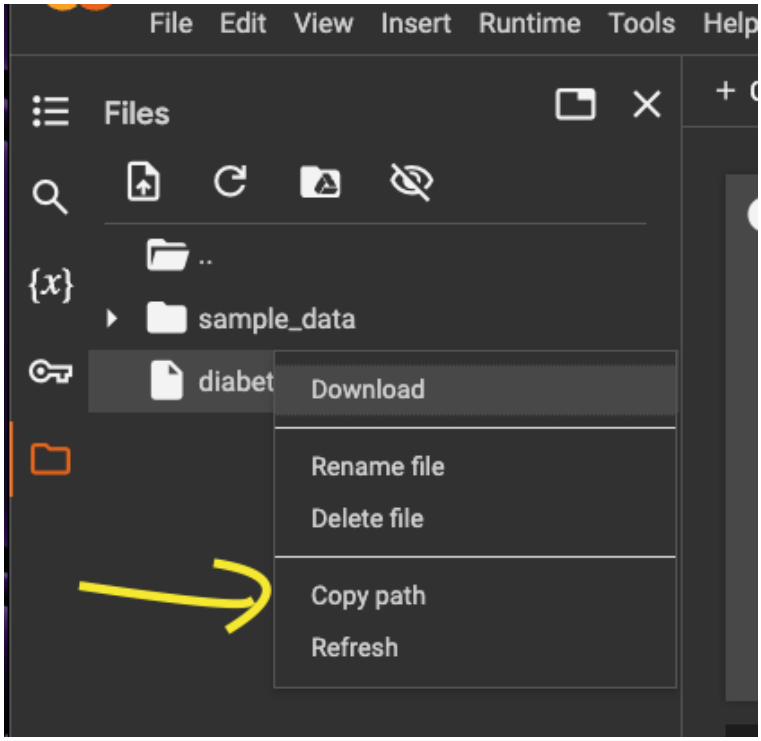

COPYRIGHT - DATA SOCIETY
Favorites
Recents
Applications
Desktop
Documents
Downloads
Locations
iCloud Drive
Google Drive
Tags
Media
Music
Photos
Movies
data
Search
Today
Course Materials
OneDrive_2024-09-24.zip
Welcome Sli...s.pptx (1).pdf
Yesterday
IntroToVisual...nalysis-1.pdf
IntroToVisual...icPlots-1.pdf
IntroToVisual...icPlots-2.pdf
IntroToVisual...icPlots-3.pdf
Previous 7 Days
data
1_IntroToVis...ationInPython
2_Interactiv...ionWithBokeh
3_Interactiv...tionWithPlotly
Data Visualiz...estions.docx
Packages_list.docx
Welcome Slides.pdf
data
IntroToVisual...tionInPython
aux_data
diabetes.csv
heart.csv
CSV
diabetes.csv
CSV Document - 26 KB
Information
Created
Today, 3:24 PM
Cancel
Open

3. Load dataset into notebook

A. Find dataset file path



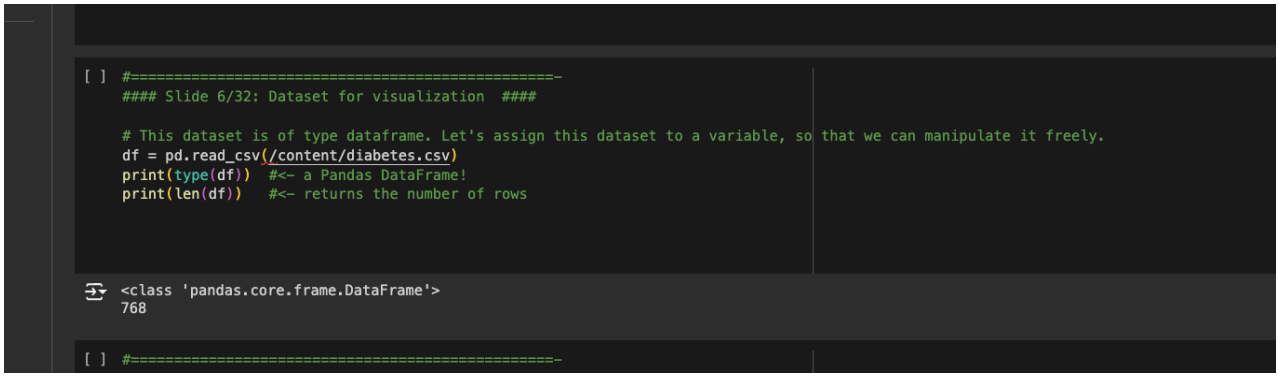
B. Copy file path



/content/diabetes.csv

C. Paste file path into the df variable

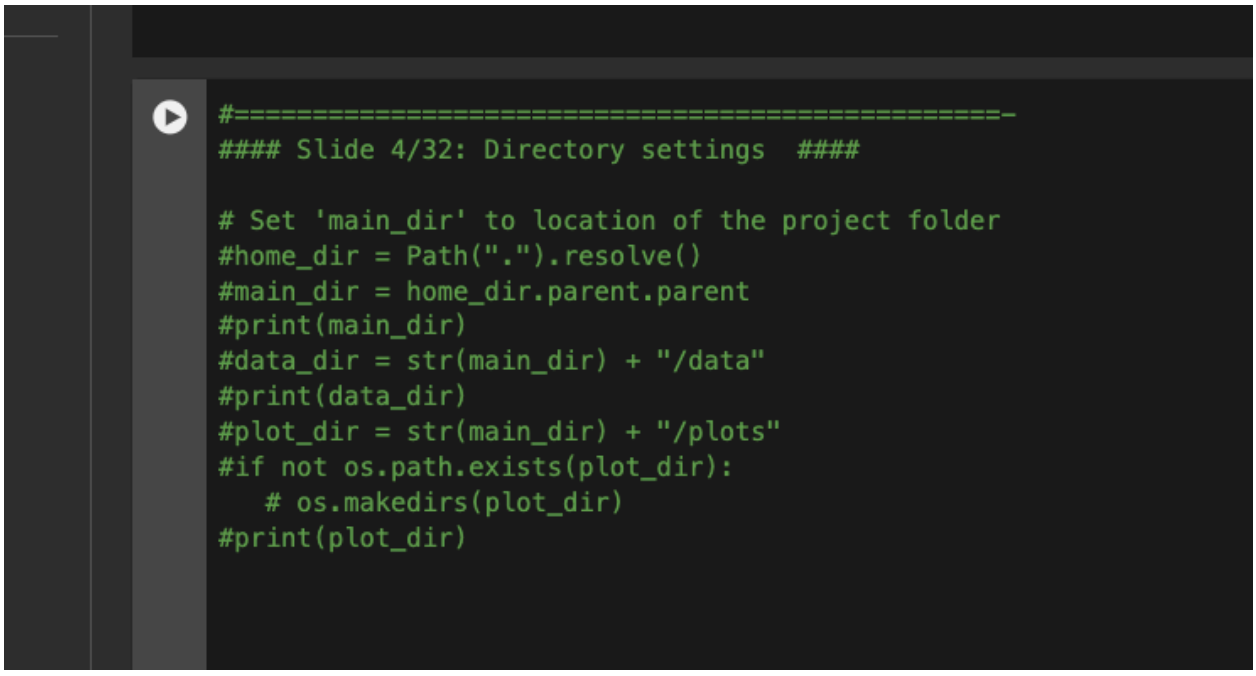
In between the quotes inside the pd.read_csv function



```
[ ] #=============================================-
    #### Slide 6/32: Dataset for visualization  ####

    # This dataset is of type dataframe. Let's assign this dataset to a variable, so that we can manipulate it freely.
    df = pd.read_csv(/content/diabetes.csv)
    print(type(df))  #<- a Pandas DataFrame!
    print(len(df))   #<- returns the number of rows

    <class 'pandas.core.frame.DataFrame'>
    768

[ ] #=============================================-
```

```
df = pd.read_csv("/content//diabetes.csv")
```

D. Comment out existing directory information (We don't need these in Colab)



```
#=================================================-
#### Slide 4/32: Directory settings  ####

# Set 'main_dir' to location of the project folder
#home_dir = Path(".").resolve()
#main_dir = home_dir.parent.parent
#print(main_dir)
#data_dir = str(main_dir) + "/data"
#print(data_dir)
#plot_dir = str(main_dir) + "/plots"
#if not os.path.exists(plot_dir):
   # os.makedirs(plot_dir)
#print(plot_dir)
```