



How to Get Started with Data Analysis In Python

Anna Grace Dickerson

July 17, 2025

Clone the 2 Github repos we'll be working with or just copy the files

1. *FreeCodeCamp-Pandas-Real-Life-Example*. Github link:
<https://github.com/ine-rmotr-curriculum/FreeCodeCamp-Pandas-Real-Life-Example>
2. *LoadingDatasetsToPython*. Github link:
[A screenshot of a GitHub repository page for 'FreeCodeCamp-Pandas-Real-Life-Example'. The repository is public and has 72 watchers. The 'master' branch is selected, showing 2 branches and 0 tags. The file list includes 'data', '.gitignore', 'Exercises_1.ipynb', 'Exercises_2.ipynb', 'Lecture_1.ipynb', and 'Lecture_2.ipynb', all from the 'Initial commit'. The 'Code' button is highlighted in yellow, and the 'Clone' dropdown menu is open, showing the 'HTTPS' URL: 'https://github.com/ine-rmotr-curriculum/FreeCodeCamp-Pandas-Real-Life-Example.git'. The 'Copy url to clipboard' button is also highlighted.](https://github.com/AnnaGraceDickerson>LoadingDatasetsToPython</div><div data-bbox=)

```
Repos — -zsh — 80x24

Last login: Fri Jul 11 21:23:20 on ttys000
(base) agrac@d-172-18-66-28 Repos % git clone https://github.com/ine-rmotr-curriculum/FreeCodeCamp-Pandas-Real-Life-Example.git
Cloning into 'FreeCodeCamp-Pandas-Real-Life-Example'...
remote: Enumerating objects: 17, done.
remote: Counting objects: 100% (13/13), done.
remote: Compressing objects: 100% (11/11), done.
remote: Total 17 (delta 4), reused 2 (delta 2), pack-reused 4 (from 1)
Receiving objects: 100% (17/17), 3.12 MiB | 7.52 MiB/s, done.
Resolving deltas: 100% (4/4), done.
(base) agrac@d-172-18-66-28 Repos %
```

If Installed, Launch Jupyter Notebook

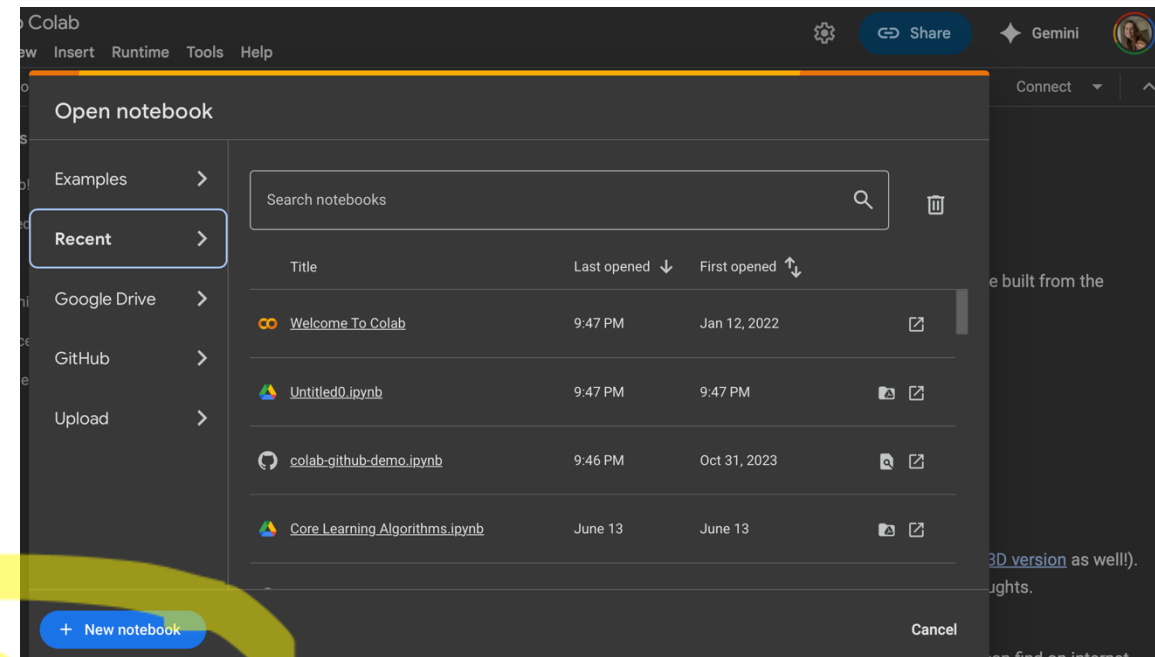
- Type 'jupyter lab' into the terminal/powershell/command line

```
((base) agrac@d-172-18-66-28 Repos % jupyter lab
[I 2025-07-11 21:43:50.222 ServerApp] jupyter_lsp | extension was successfully linked.
[I 2025-07-11 21:43:50.224 ServerApp] jupyter_server_terminals | extension was successfully linked.
[I 2025-07-11 21:43:50.226 ServerApp] jupyterlab | extension was successfully linked.
[I 2025-07-11 21:43:50.227 ServerApp] notebook | extension was successfully linked.
[I 2025-07-11 21:43:50.371 ServerApp] notebook_shim | extension was successfully linked.
[I 2025-07-11 21:43:50.426 ServerApp] notebook_shim | extension was successfully loaded.
[I 2025-07-11 21:43:50.427 ServerApp] jupyter_lsp | extension was successfully loaded.
[I 2025-07-11 21:43:50.428 ServerApp] jupyter_server_terminals | extension was successfully loaded.
[I 2025-07-11 21:43:50.430 LabApp] JupyterLab extension loaded from /opt/homebrew/Cellar/jupyterlab
3.4.1/libexec/lib/python3.13/site-packages/jupyterlab
[I 2025-07-11 21:43:50.430 LabApp] JupyterLab application directory is /opt/homebrew/Cellar/jupyterlab
/4.3.4_1/libexec/share/jupyter/lab
[I 2025-07-11 21:43:50.430 LabApp] Extension Manager is 'pypi'.
[I 2025-07-11 21:43:50.451 ServerApp] jupyterlab | extension was successfully loaded.
[I 2025-07-11 21:43:50.452 ServerApp] notebook | extension was successfully loaded.
[I 2025-07-11 21:43:50.453 ServerApp] The port 8888 is already in use, trying another port.
[I 2025-07-11 21:43:50.454 ServerApp] Serving notebooks from local directory: /Users/agrac/Users/agrac/Repos
/Repos
[I 2025-07-11 21:43:50.454 ServerApp] Jupyter Server 2.15.0 is running at:
[I 2025-07-11 21:43:50.454 ServerApp] http://localhost:8889/lab?token=caea7dace285646f972cb81e3f090f106dbc5bff2fbaeec9
6dbc5bff2fbaeec9
[I 2025-07-11 21:43:50.454 ServerApp] http://127.0.0.1:8889/lab?token=caea7dace285646f972cb81e3f090f106dbc5bff2fbaeec9
0f106dbc5bff2fbaeec9
[I 2025-07-11 21:43:50.454 ServerApp] Use Control-C to stop this server and shut down all kernels (
ce to skip confirmation).
[C 2025-07-11 21:43:50.456 ServerApp]

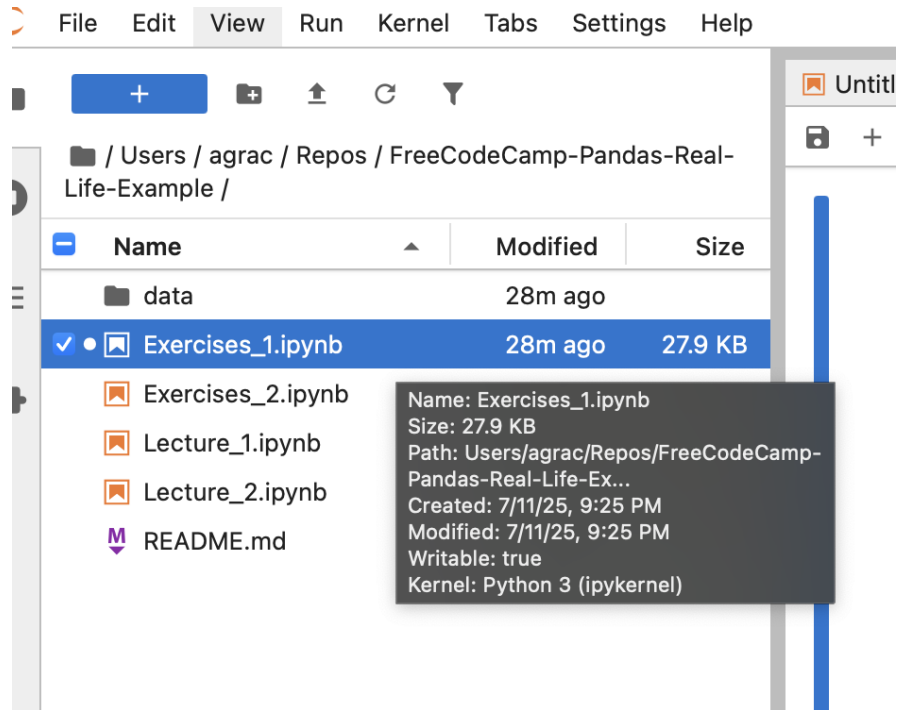
To access the server, open this file in a browser:
file:///Users/agrac/Library/Jupyter/runtime/jpserver-33557-open.html
Or copy and paste one of these URLs:
http://localhost:8889/lab?token=caea7dace285646f972cb81e3f090f106dbc5bff2fbaeec9
http://127.0.0.1:8889/lab?token=caea7dace285646f972cb81e3f090f106dbc5bff2fbaeec9
[I 2025-07-11 21:43:50.789 ServerApp] Skipped non-installed server(s): bash-language-server, docker
```

Else, Open Google Colab Through a Google Account

- Go to <https://colab.research.google.com>
- Sign in to your google account if not already
- Click +New Notebook



Open Exercise_1.ipynb Using the File Explorer

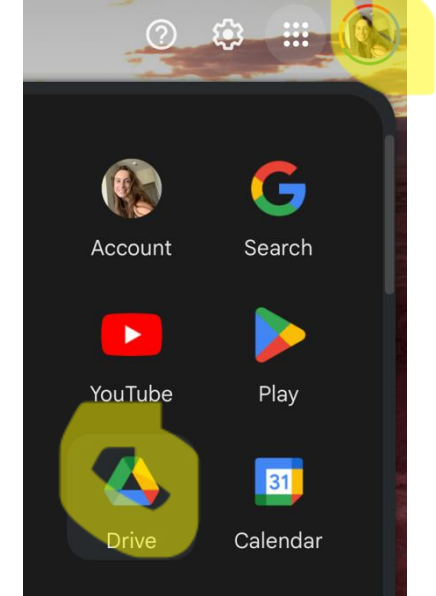
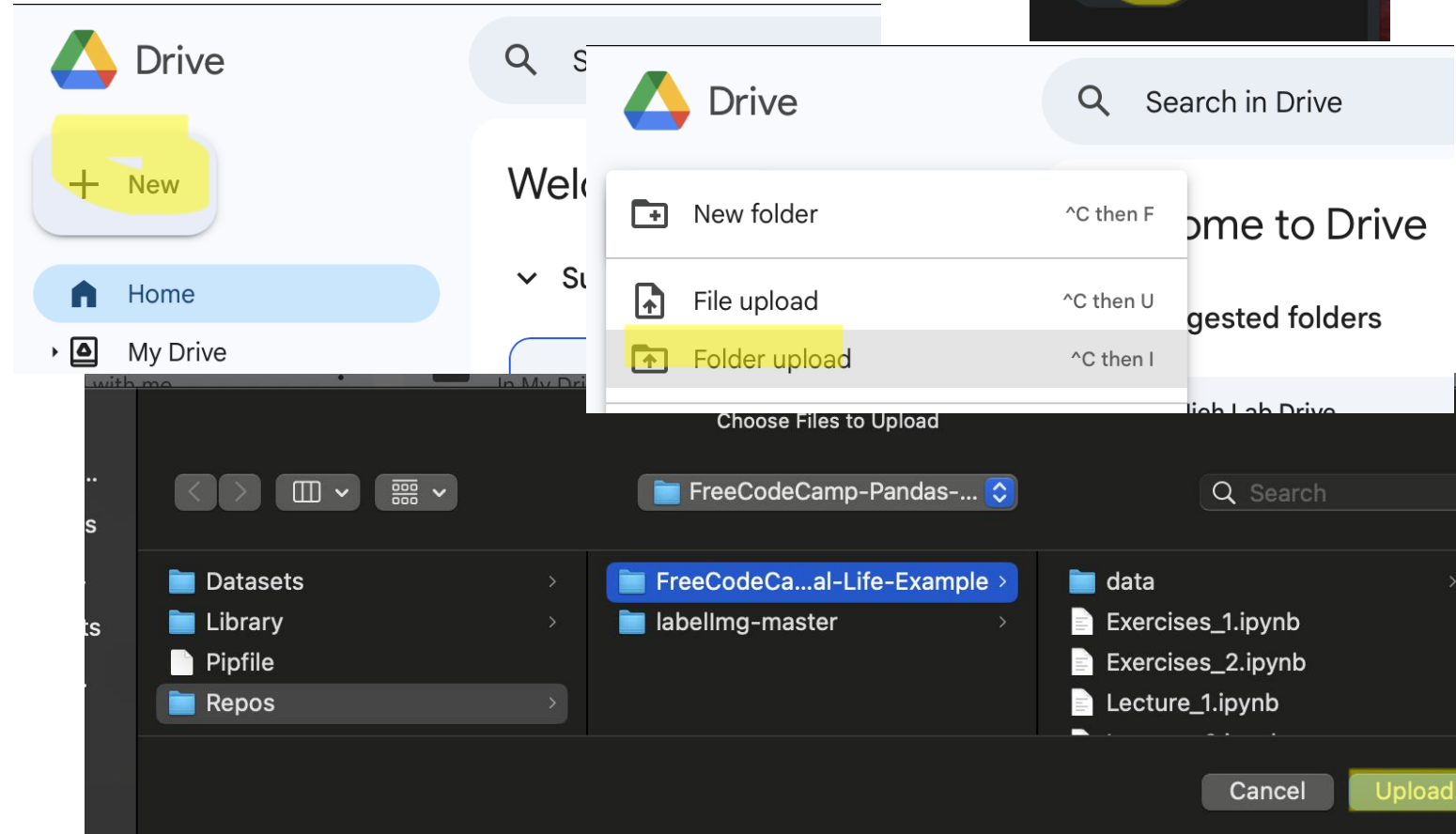


Open Exercise_1.ipynb by

1. Uploading the Folder to Your Google Drive
2. Mounting Your Google Drive

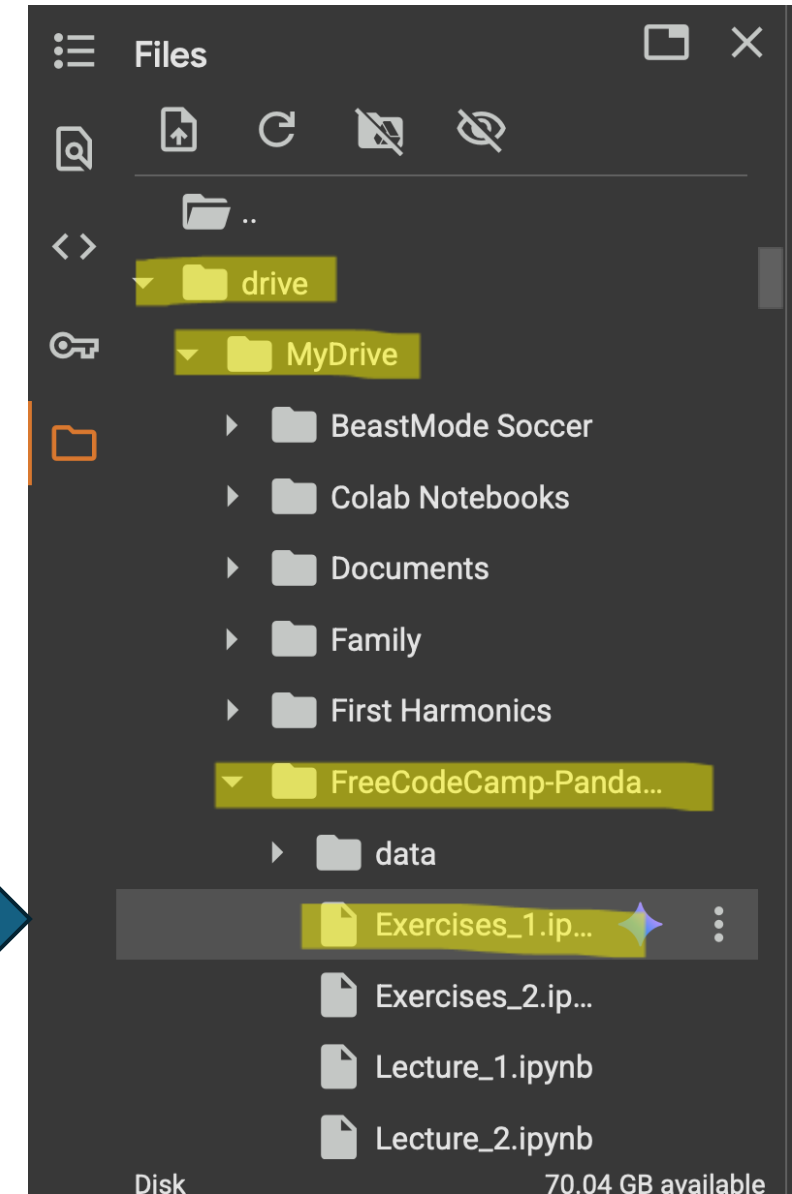
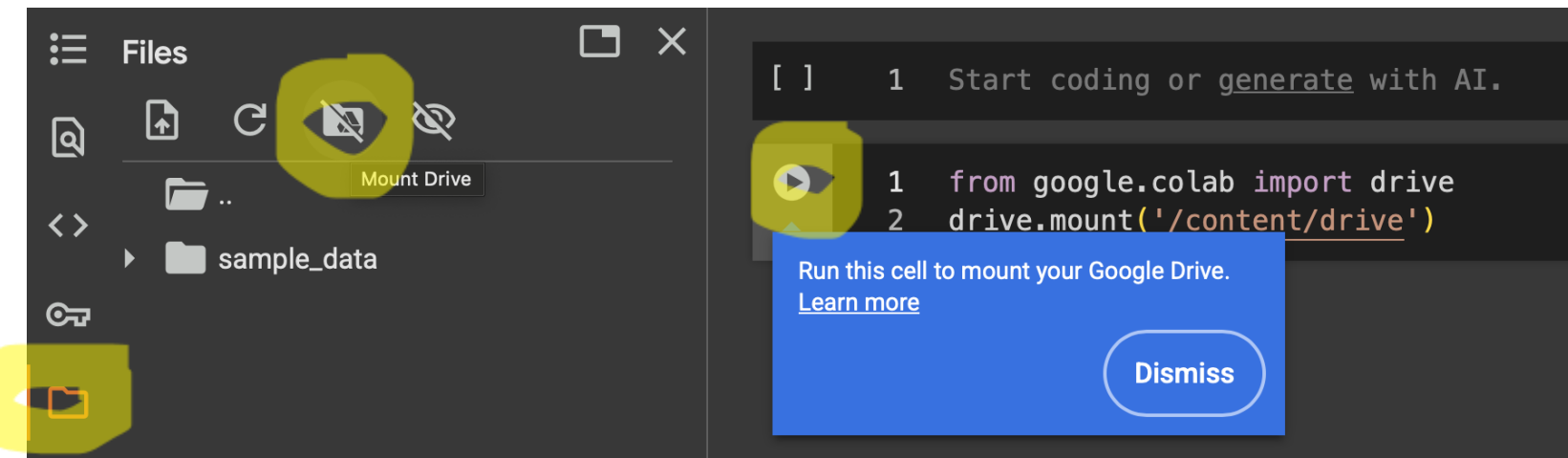
Step 1:

- Go to your google drive
- Click +New > Folder Upload
- Find and Upload the [FreeCodeCamp](#) folder & the [LoadingDatasetsToPython](#) you just downloaded or cloned from the repo



Another step for Google Colab

- Go back to Google Colab
- Click the folder image on the menu bar on the right
- Click the Google Drive folder image in the menu bar that pops out
- Click Connect to google drive and allow all access
- Run the code cell it generates to mount your google drive
- This will make a new folder named 'drive' appear
- Open this folder, navigate to the FreeCodeCamp folder you just uploaded and open Exercise_1.ipynb



Now We Code

Learn what you can do with the pandas package which can be used on any table. You can read in csv's, excel's, json's, sql tables, R tables, etc.

Learn Data Analysis Functions

- Lecture 1 (can go back to FreeCodeCamp if you want to hear what they say)
- Exercise 1
- Lecture 2
- Exercise 2

Learn How to Read in a Variety of Data Types

- Open [ExamplesOfLoadingInDifferentTypesOfData.ipynb](#) in the other LoadingDatasetsToPython repo you cloned/downloaded
- Run the cells to upload the different types of data. Try out some of the data analysis functions you learned in the Exercise/Lecture notebooks on this data.

Start Applying Your New Skills to Your Own Projects

- Finally, using the functions you've learned, can you read in one of your own data files and play around with it?

Bonus: Keep going through FreeCodeCamp's *Data Analysis* modules on your own to keep expanding and practicing your skills. Maybe even try out some other courses that seem relevant – perhaps *Data Visualization* or *Machine Learning with Python*.