

COMP 4613

Machine Learning

Daniel L. Silver

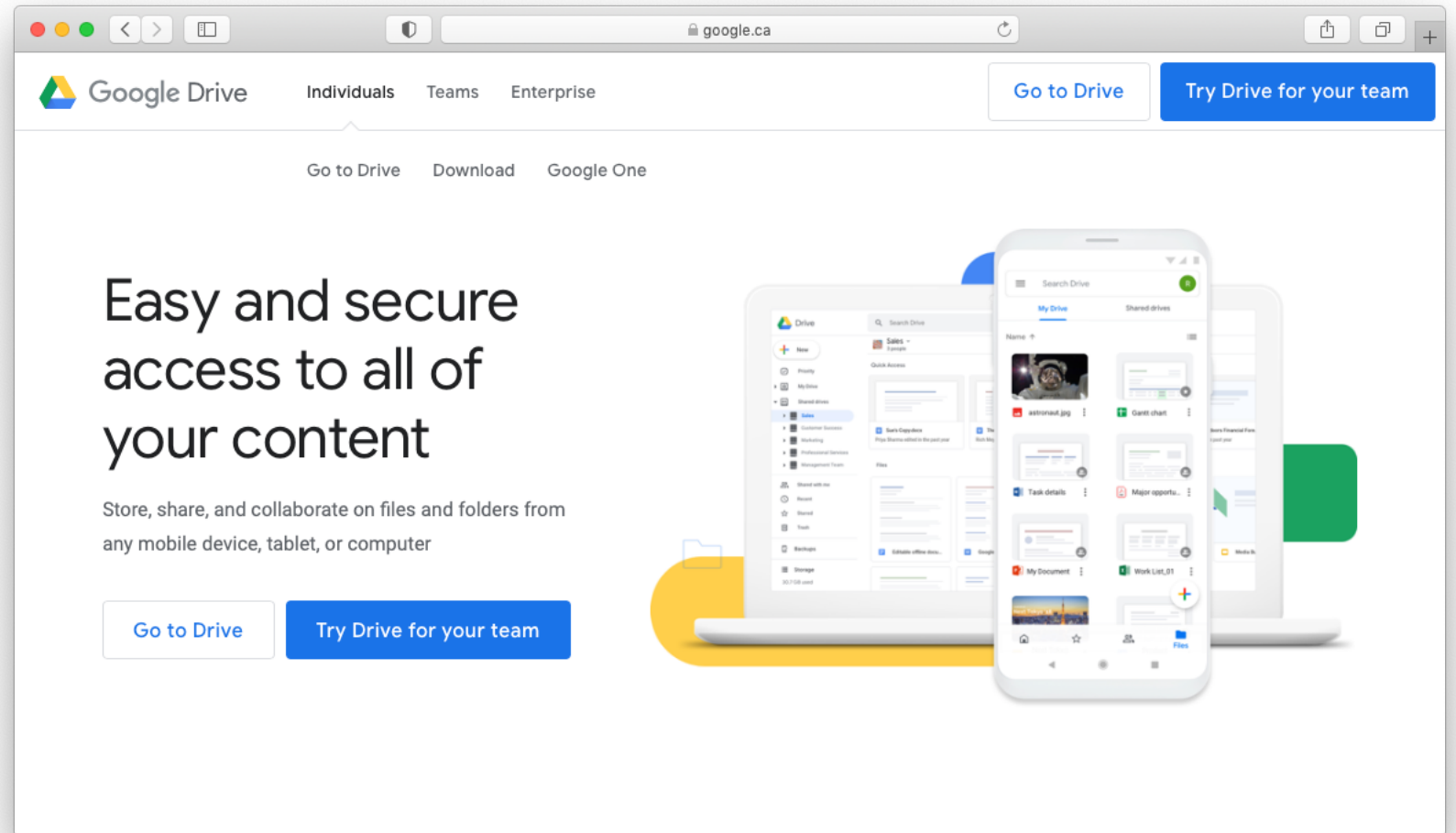
Colab and the Python Data Science Stack

Google Drive

Store, share, and collaborate on files and folders from any mobile device, tablet, or computer

Connects seamlessly to Google Colaboratory

<https://www.google.ca/drive/>



Jupyter Notebooks

Project Jupyter exists to develop open-source software, open-standards, and services for interactive computing across dozens of programming languages.

Allows you to easily combine code and documentation in one file.

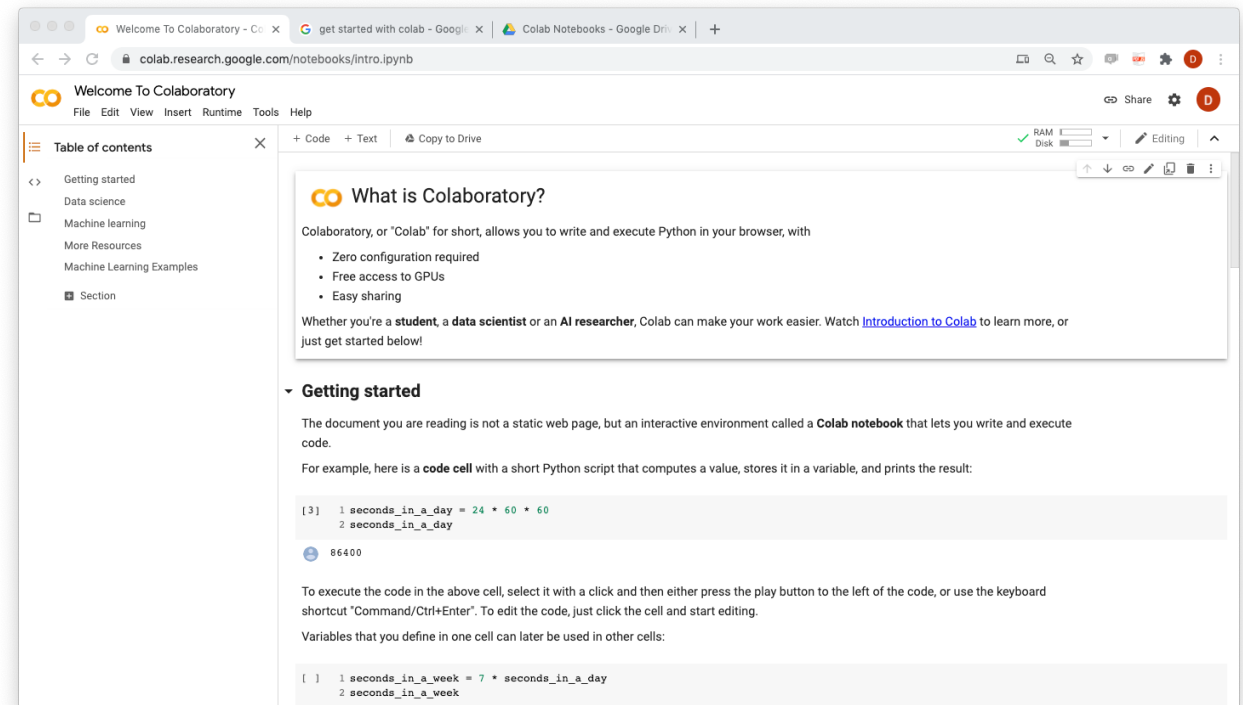
Used by Google Colaboratory.

<https://jupyter.org/>



What is Colaboratory? aka. Colab?

- A free Jupyter notebook environment
- Requires no setup and runs entirely in the cloud
- You can write and execute code
- Save and share your analyses
- Access powerful computing resources
- All for free from your browser
- [Intro to Colaboratory](#)
- [Colab Tutorials](#)



Python Data Science Stack



Python Data Science Stack

Python training/review:

- <https://www.learnpython.org/>

Overview of most used data science libraries:

- <https://stackabuse.com/the-best-data-science-libraries-in-python/>

Full Stack Python website:

- <https://www.fullstackpython.com/table-of-contents.html>
- <https://www.fullstackpython.com/data.html>

Anaconda Distribution for your own computer:

- <https://www.anaconda.com/>

Python Machine Learning with Scikit-learn

- Main scikit-learn website
 - <http://scikit-learn.org/stable/>
- Installing scikit-learn
 - <http://scikit-learn.org/stable/install.html>
- Overview of ML and scikit-learn
 - <http://scikit-learn.org/stable/tutorial/basic/tutorial.html>
- Colab Tutorial
 - https://www.tutorialspoint.com/google_colab/index.htm
- Introduction to machine learning in Python with scikit-learn and Jupyter notebooks (video)
 - <https://www.dataschool.io/machine-learning-with-scikit-learn/>
- Associated Jupyter Notebooks
 - <https://github.com/justmarkham/scikit-learn-videos>
 - just click the green “Clone or download” button
- scikit-learn user guide - Release 0.23.0
 - <http://scikit-learn.org/stable/downloads/scikit-learn-docs.pdf>

Tutorial / Assignment

- Please get started on the
- Assignment – Colab Python Data Science Stack
- It has you complete two tutorials:
 - Jupyter Notebook for Beginners: A Tutorial
 - <https://www.dataquest.io/blog/jupyter-notebook-tutorial/>
 - Note that you are to run this in Google Colab via your browser and not directly on your laptop. So you do not need to install Anaconda and Jupyter Notebooks.
 - 10 Minutes to pandas
 - <http://pandas.pydata.org/pandas-docs/stable/10min.html>
 - Complete the tutorial down to and including the Concat section. Also include the section on Plotting and Getting Data in/out.
- You will need assistance from
 - [Loading data: Drive, Sheets, and Google Cloud Storage](#)