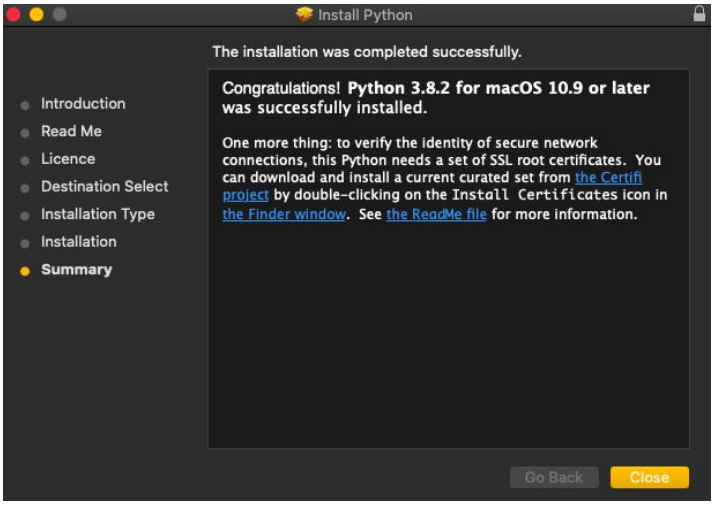
demo by: violeta mezeklieva

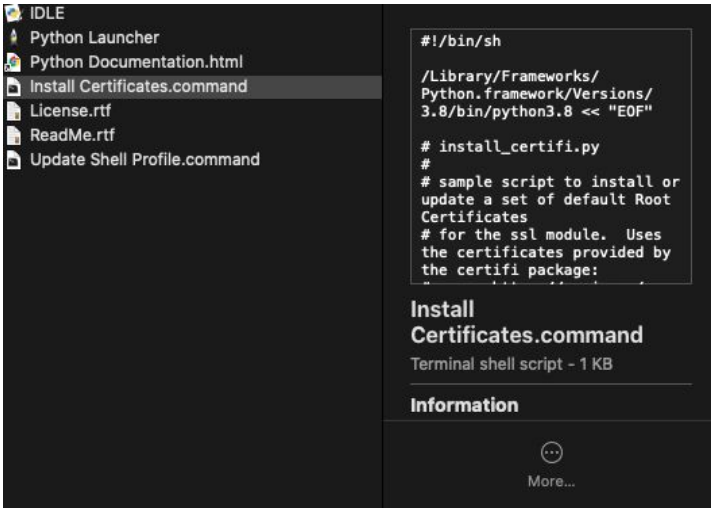
Installing Python

Instructions for a MAC OS user

Install the latest version of Python from this link: https://www.python.org/downloads/



Pay particular attention to the Installation Instructions - particularly the last step which asks you to download the Certificates as their needed to register your computer's identity as a having a secure network connection.



This is the screenshot inside the folder where Python was just installed.

Once you are done, close the installation window. Next step is to check if Python has been installed properly using the Terminal.

How to open Terminal?
Press *cmd* and the *spacebar*, type Terminal and press *enter.* The terminal will open, type *python 3 --version* and press enter. You should see the it returns Python 3.8.2

You are all set!

demo by: violeta mezeklieva

Installing Jupyter Notebooks

Instructions for a MAC OS user

Let's install this application using the Terminal. Don't be afraid. You will open a door into computer science you thought you might never encounter.

- Open the Terminal
- Type *pip3 install --upgrade pip*
- Press enter and wait for a message that confirms the installation.

- Type *pip3 install notebook*
- Press enter and wait for the installation to complete.

You are all set!

Translating terminology from the above:

What is *pip*?

It is a common command used to install and manage software packages.

What is a *package*?

All computer programs need a collection of resources to execute commands, these are called modules - a module being a library. These modules may include configuration data, documentation, pre-written code, etc. For example, when we installed Python3 above, the installation contained a set of modules that were downloaded as they are required to install Python 3 on your computer. Python can use many libraries to execute different operations - for reference, Python has 250+ libraries.

You can read more here - but you'll soon find yourself down a rabbit hole.

What are the common python libraries for data science? You can find more about them here.

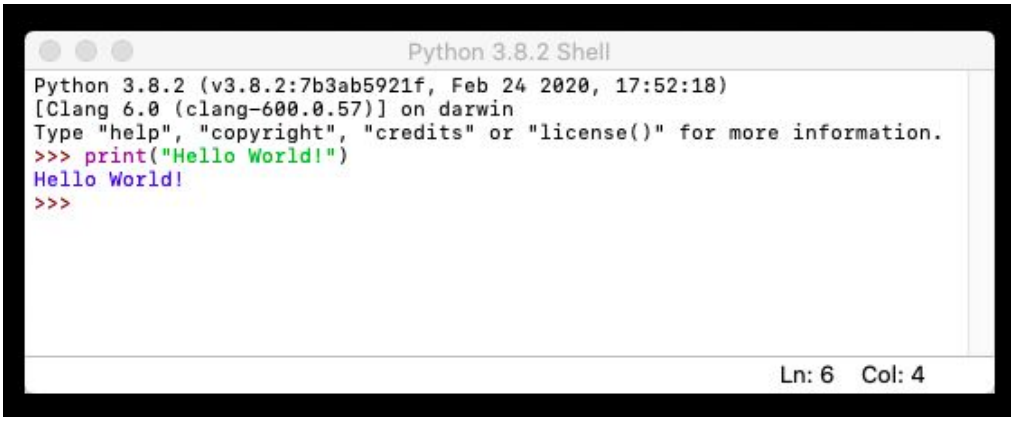
Where on my computer can I program in Python?

Python is a programming language that needs a specific environment in order to execute its commands. To put it simply, you would not use a Word Document if your intention is to calculate how much you earned this month; you would use an Excel Document instead. The same logic applies to Python.

In the folder where Python was installed, you will see an application called IDLE (Integrated Development and Learning Environment) this is the default environment.

Let's test it to see how it works!

Navigate to the folder where Python was downloaded, and open IDLE. A white screen - "Shell" in computer science terms - will open. Type *print("Hello World")* and press enter.

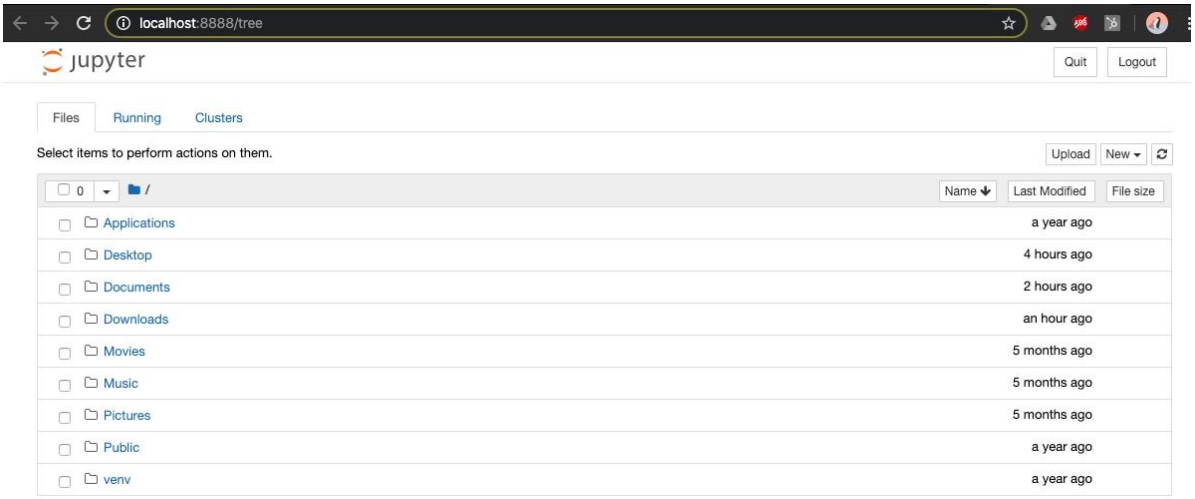


What are Jupyter Notebooks for?

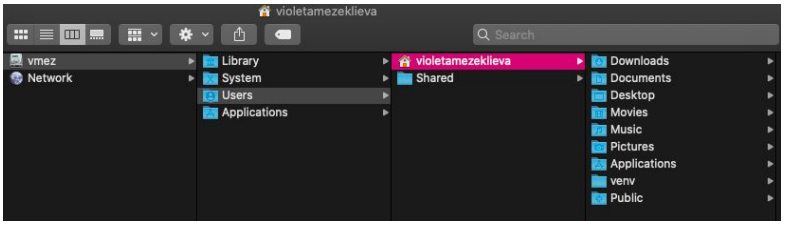
The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Read this article published in Nature to hear why is Jupyter so popular. There are alternative Apps that are useful to share data projects, for example Google Colaboratory - you can read a review here - and any others. At the end of the day, they work similar to each other and you'll have to learn to adapt. The benefits of Jupyter Projects is that it is an open-source tool whereas Google Colab is proprietary to Google, and you can only share your notebook with people who have an account with this company.

Let's see what Jupyter Notebooks looks like!

Open your Terminal and type *jupyter notebook* press enter and wait for a few seconds. A new window will open on your browser - called the Notebook Dashboard - and looks similar to this screenshot.



If you open Finder, and navigate to your user account in your system, you will see that Jupyter mimics the folders and the arrangement you have set up here:



Click through the folders to see how Jupyter Notebooks and your computer are linked.
This step is particularly important because it will help you understand where you save your files, and will enable you to better understand how to manage the information on your computer.

If your plan is to analyse and visualise data, you will need to install some of the most common packages for data analysis. Let me guide you on how to do that:

- Navigate to a folder or create a new one where you want to keep your Jupyter notebooks.

- Select “new” on the top right corner, and select “python 3”. This automatically opens a window with your first Jupyter Notebook.
- You will need to know the exact path where Jupyter Notebook has been installed. Type inside the code below and run the code either by pressing the button that reads “run” or by pressing “shift+enter”
    - **import sys**
    - **sys.executable**
- Copy the output: it should return something like this /Library/Frameworks/Python.framework/Versions/3.8/bin/python3.8
- Go to Terminal and open a new window
- And now you will need to install pandas and seaborn.
- Type **-m pip install pandas** and press enter. Wait until pandas is installed (you will see there are depending packages that are installed as well such as numpy and matplotlib
- Type **-m pip install seaborn** and press enter. Wait until seaborn is installed
- Go back to the Jupyter Notebook and test if these packages have been installed by typing in the code cell:
    - **import pandas as pd**
    - **import seaborn as sns**
- And you are ready to go!

To leave Jupyter Notebooks, you can press *Quit* on the top right corner or *log-out* first and then close the window on your browser and then quit the Terminal.

If you log-out it will ask you to create a password. If you choose this route, you will be asked to enter this password every time you open Jupyter Notebook. To set your password, you will be asked to provide the token - which you will find in the last line of the Terminal after http://localhost:8888/?token= or http://127.0.0.1:8888/?token=. Copy that number, and then type the password of your choice. Make sure you remember it! Otherwise it’s a pain to reset it.

If you want to learn more about the Jupyter Project, go to their website.

Ready to start coding?

Kaggle has great resources for learning. Following this link to find exercises you can do either on IDLE or Jupyter Notebook right from the gecko.