PROGRAMMING WITH PYTHON

GO.GWU.EDU/LIBWORKSHOPS

Today's Instructor

- Dan Kerchner kerchner@gwu.edu

Materials: go.gwu.edu/pyw

Today's Plan

~2.5 hours:

Basic Python Language Concepts

1 hour

Lunch Break

~2.5 hours

Data Analysis with Pandas

About today...

- Ask questions!
- If you're stuck:
  - Ask us
  - Help each other out!
- If something is confusing in the workshop, it probably needs improvement; let us know.
- Stay as long as you like

Objectives

- Gain familiarity with one environment for using Python (Google Colab), and awareness of others
- Learn Python language basics
- Load in a data set as a Pandas DataFrame
- Explore and transform ("wrangle") the DataFrame
- Create data visualizations
- Learn how to look things up, how to interpret errors
- Gain confidence to try things we didn't learn today!

Why Python?

- Free
- General purpose
- Easy to learn
- Readable*
- Community-developed / Open Source
- Widely used and documented
- Good built-in and contributed libraries

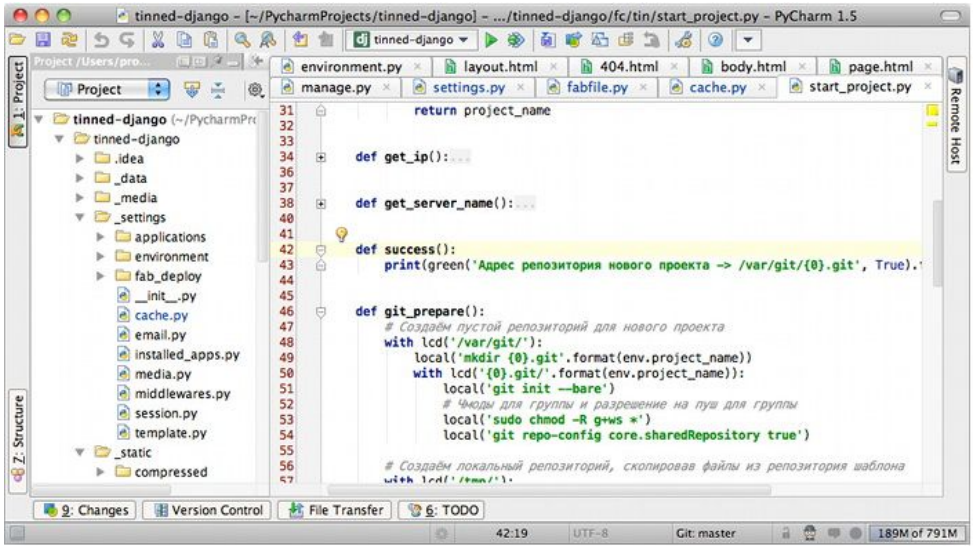
Different ways to use Python

- Command line/REPL

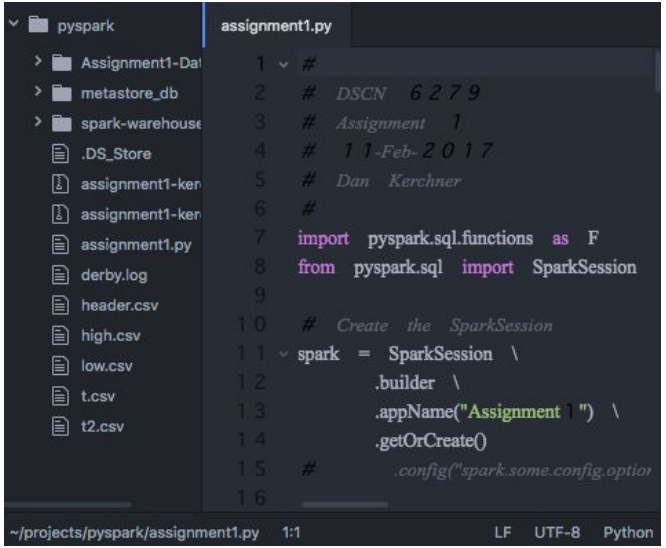
```
Last login: Mon Mar 20 22:09:33 on ttys001
GLSS-M17LFFT:~ kerchner$ python
Python 2.7.10 (default, Oct 23 2015, 19:19:21)
[GCC 4.2.1 Compatible Apple LLVM 7.0.0 (clang-700.0.59.5)] on darwin
Type "help", "copyright", "credits" or "license" for more information.
>>>
>>> opinion = "This workshop is awful!"
>>> opinion == True
False
>>>
```

Different ways to use Python

- Integrated Development Environment (IDE) – Spyder, pyCharm, pyDev, Sublime, …

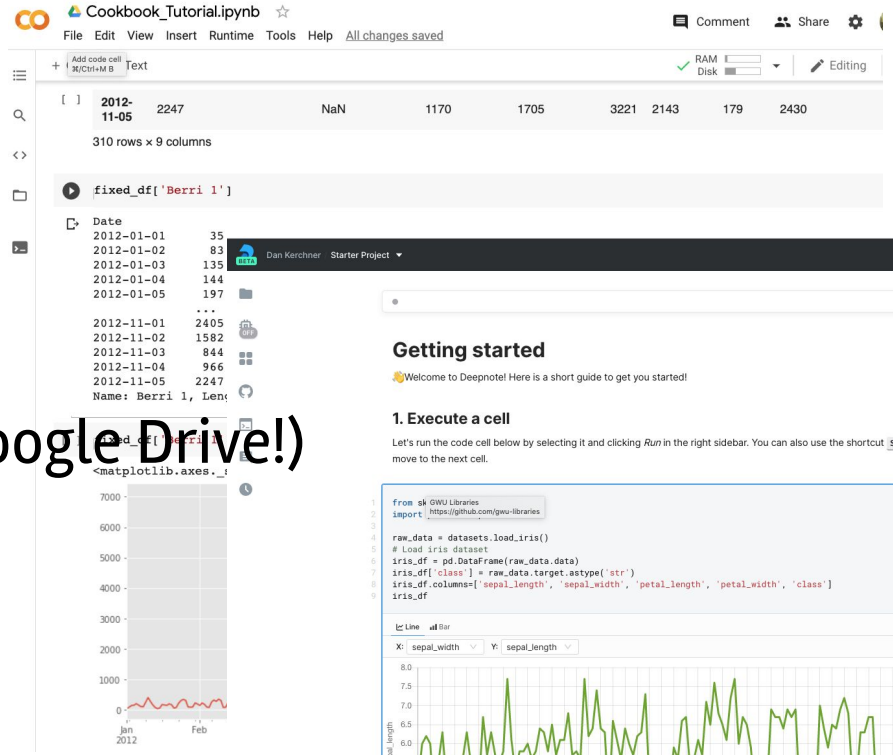


- File editor (e.g. Atom, vim) + command line tools (pip, virtualenv, …)



Different ways to use Python (continued)

- "Notebooks":
  - Jupyter notebooks
  - Google Colab
    (available in your Google Drive!)
  - Kaggle notebooks
  - Deepnote
  - Binder



Even more ways to use Python

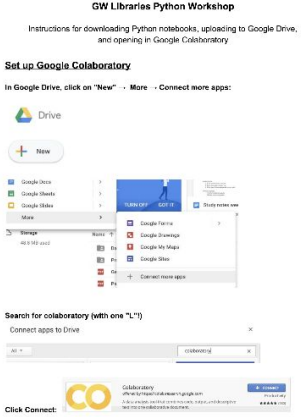
Anaconda = Python (and R) plus:

- **Jupyter notebooks**
- lots of libraries
  - data processing
  - analytics
  - scientific computing
  - including: **Pandas**

Setup

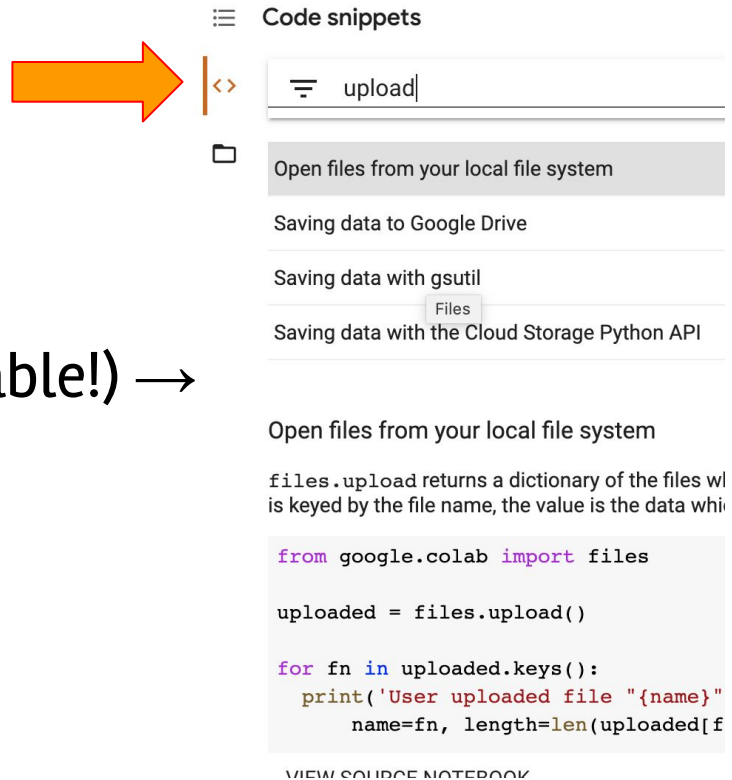
- Google Colaboratory
  https://colab.research.google.com

- Backup plan: https://jupyter.lai.gwu.edu



Uploading Data (and doing other things) in Google Colab

Use "Code Snippets" (searchable!) →



Some recommendations

- Write assuming your code will be read (incl. by Future You)
- Version your code GitHub
- Learn to be "Pythonic" in your style
- Isolate your projects from each other – use venv
- Stuck? Try an Internet search
- Which version of Python are you using?
- Find good code examples and make them work
- Keep learning!

Some Python libraries/frameworks

Building web applications	Django Flask
Scientific/numerical	Numpy Scipy Pandas
Machine Learning	scikit-learn, tensorflow
Data Visualization (check out www.python-graph-gallery.com)	matplotlib bokeh ggplot (like ggplot2 in R) plotly (<- interactive) seaborn

To Learn More

- Kaggle: kaggle.com/learn
- learnpython.org
- Software Carpentry, Data Carpentry (not just Python)
- docs.python.org/3/tutorial (and docs.python.org/)
- GW Online: Get data off the ground with Python
- Upcoming Python workshops @ GW Libraries
- LinkedIn learning it.gwu.edu/linkedin-learning courses
  - 253 Python, 6 Pandas
- More on Pandas:
  - Pandas cookbook: github.com/jvns/pandas-cookbook
  - pandas.pydata.org/pandas-docs/stable/10min.html
  - pandas.pydata.org/pandas-docs/stable/tutorials.html
  - pandas.pydata.org/pandas-docs/stable/cookbook.html
  - www.datacarpentry.org/python-ecology-lesson/

Contact us:

Coding Consultations (with Dan & colleagues):

go.gwu.edu/coding

Stats Appointments (with Stats grad students):

academiccommons.gwu.edu/data-consulting

→Penji→Data Consulting

Workshop Materials: go.gwu.edu/pyw

kerchner@gwu.edu