Getting Started with Python

Biochemistry Boot Camp 2021
Session #12
Nick Fitzkee
nfitzkee@chemistry.msstate.edu

1

What is Python?

- Simple, interpreted programming language

- Guido van Rossum, 1991

- Interpreted, not compiled: A program (*Python interpreter*) parses python commands/code and executes the statements, one at a time

- Compiled languages (like C and FORTRAN) use a *compiler* to convert code into machine language, which is run natively

2

2

Why (or Why Not) Python?

Pros	Cons
• Very simple syntax, easy to read	• Not as fast as C/C++/FORTRAN
• Installed on most Linux, Mac	• Windows support not as good
• Huge library of standard tools	• Some technical tasks are more challenging (e.g., direct memory access)
– Standard functions (file I/O, strings, lists)	• Still requires rigor of learning programming
– Math/Science/Bio	
• Many complex tasks are simplified (memory management)	
• Relatively easy to get working code	

3

3

Python 2 vs. Python 3

- Different _versions_ of Python interpreter exist (e.g. 2.7.18, 3.9.5)
  - New versions fix bugs, add features, extend the language

- Many syntax features from 2.x.x work in 3.x.x, but some do not
  - "under the hood" is very different

- Why? Simplify and streamline code, make syntax more consistent

- Python 2 officially discontinued in 2020, but a significant code base still exists

4

4

Do I Already Have Python?

- Starting python at the command line: Simply type "python"
  - May need to type "python3" (Mac, some Linux systems)

```
[nfitzkee@blargh: ~] python
Python 2.7.18 (default, Mar  8 2021, 13:02:45)
[GCC 9.3.0] on linux2
Type "help", "copyright", "credits" or "license" for more information.
>>>
>>>
[nfitzkee@blargh: ~]
[nfitzkee@blargh: ~] python3
Python 3.8.5 (default, May 27 2021, 13:30:53)
[GCC 9.3.0] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>>
>>>
[nfitzkee@blargh: ~] ▮
```

- Python 3 is strongly preferred over Python 2
  - Most (not all) Linux systems will default to version 3
  - Most Macs use version 2 (legacy reasons)

5

5

Installing Python: **python.org**
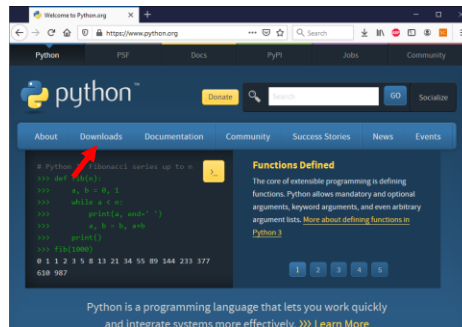
- **DON'T** use the Windows App Store:

- Windows App Store would be great, but:
  - Broken IDLE context menu
  - Asks you to log into Windows store (not really needed)
  - Maybe in a year or so?

6

6

3

Installing Python: **python.org**

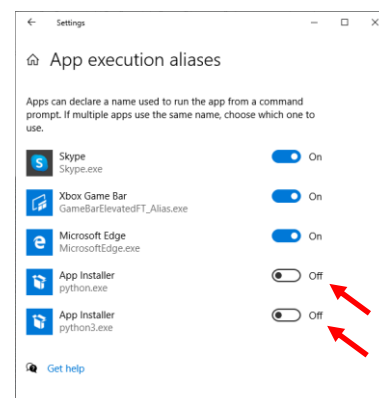- For Mac and Windows, download the latest version www.python.org:

- Linux will require administrator privileges, but anything $\geq$ 3.7 should be fine

7

7

Disable Windows "Python" Shortcuts

- Typing "python" at the command line on Windows 10 (even if Python is installed) brings up the app store window

- Search "app execution" in the start menu, disable app installers for python and python3

8

8

Make Sure Python Is In Your Path

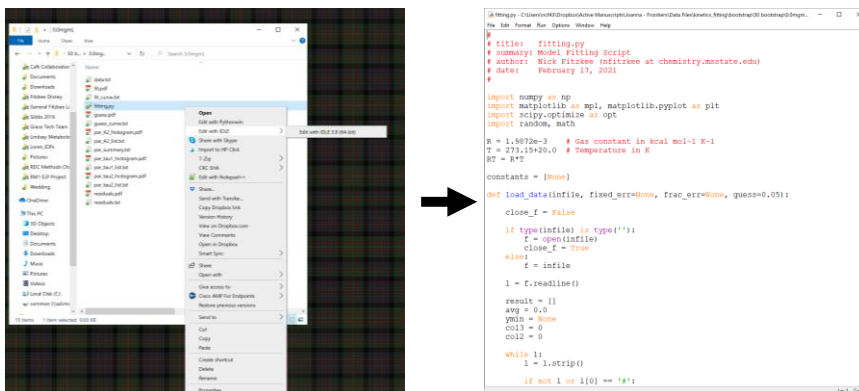- On Windows 10, most defaults are fine, but be sure Python is in your path:

- If you change other defaults, be sure to install pip, IDLE
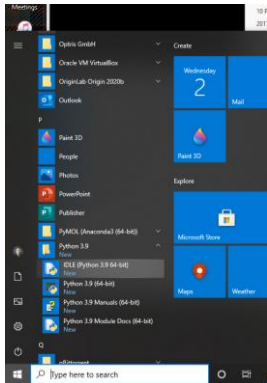
9

9

Starting Python on Windows

- **Method 3:** Right click on a .py file, then select "edit with IDLE"
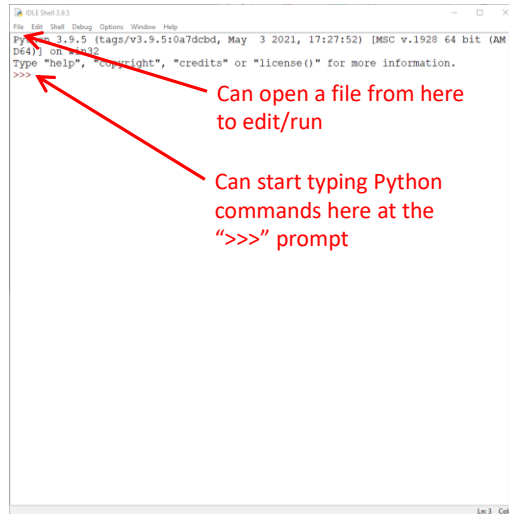 - Can run the file by pressing F5 or shift-F5

10

10

Starting Python on Windows

Can open a file from here to edit/run

Can start typing Python commands here at the ">>>" prompt

- **Method 2:** Open IDLE from the Start Menu

11

Python Modules

- The program *pip* is used to add modules & features to Python
  - Extend functionality, re-use code from others

- Key modules:
  - **Numeric Python (numpy,** https://numpy.org/**):** support for fast matrix/vector calculation
  - **Matplotlib (matplotlib,** https://matplotlib.org/**):** Create interactive graphs and PDFs from within Python
  - **Scientific Python (scipy,** https://scipy.org/**):** support for function optimization, numerical integration of differential equations
  - **BioPython (Bio,** https://biopython.org/ **):** Parse PDB files, manipulate DNA/protein sequences

12

Installing Key Modules

- First, update pip:
  ```
  pip install --upgrade --user pip
  ```
  - May need to type "`pip3`" instead

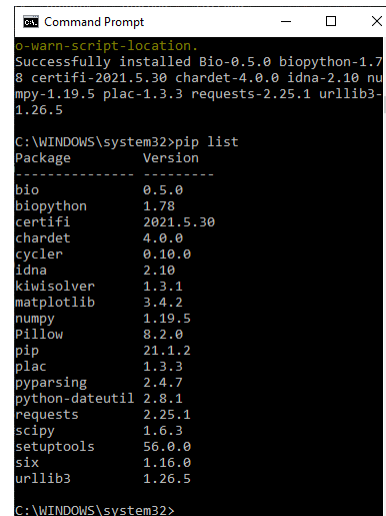- Then, install numpy, scipy, matplotlib, and Bio:
  ```
  pip install --user numpy
  pip install --user matplotlib
  pip install --user scipy
  pip install --user biopython
  ```

- Some additional packages may be installed; these are updated from web (Internet required!)
  - Can `--upgrade` at any time

13

13

Installing Key Modules

- If everything is updated, type "`pip list`" and you should see something like what is shown to the right

- Versions will vary, but notice: biopython, scipy, numpy, and matplotlib

```
o-warn-script-location.
Successfully installed Bio-0.5.0 biopython-1.7
8 certifi-2021.5.30 chardet-4.0.0 idna-2.10 nu
mpy-1.19.5 plac-1.3.3 requests-2.25.1 urllib3-
1.26.5

C:\WINDOWS\system32>pip list
Package         Version
--------------- ---------
bio             0.5.0
biopython       1.78
certifi         2021.5.30
chardet         4.0.0
cycler          0.10.0
idna            2.10
kiwisolver      1.3.1
matplotlib      3.4.2
numpy           1.19.5
Pillow          8.2.0
pip             21.1.2
plac            1.3.3
pyparsing       2.4.7
python-dateutil 2.8.1
requests        2.25.1
scipy           1.6.3
setuptools      56.0.0
six             1.16.0
urllib3         1.26.5

C:\WINDOWS\system32>
```

14

How I Code (Try This First)

- Open a google window; lots of good examples are on the internet (e.g., "how to sort a list python")

- **Windows:** Open a text file in IDLE, edit the file using IDLE, run using F5
  - Test commands can be run in the Python shell

- **Linux or Mac:** Open a text file in an editor in the background (e.g. Xemacs), run it by repeatedlys aving and typing "`python <file.py>`"
  - Remember to use ampersand (&) when starting the editor!

- Let's try it both ways!

15

15

The Dirty Secret

- I can't teach you how to code in one 75-minute session
  - You must teach yourself or take a class to master coding
  - With experience, you will know syntax and be able to interpret error messages

- Establishing a workflow and having all the tools ("development environment") is the major hurdle for new users

- **Good News:** You now have a development environment and can start playing
  - **Yes, it's play.** Learning starts with play.

16

16

Where to Go From Here?

(all of these take *time*; *invest* and *reap dividends*)

- **How to Think Like a Computer Scientist**
  http://openbookproject.net/thinkcs/python/english3e/
  – Focuses on problem solving with Python, very complete list of topics

- **Automate the Boring Stuff Using Python**
  https://automatetheboringstuff.com/
  – Website that focuses on using Python for routine stuff (as scientists often do!)

- **Python Module Index (for Reference)**
  https://docs.python.org/3/py-modindex.html
  – All built-in module documentation, can be very useful for reference, e.g. the `time` and `math` modules contain functions for converting time and performing simple math operations
  – Primarily a last resort, but don't forget that *all* aspects of the core language are documented!

17

17

Where to Go From Here?

(all of these take *time*; *invest* and *reap dividends*)

- **NumPy, SciPy, Matplotlib, etc. Websites**
  – These modules may not be documented as well, but they are all documented to a certain extend

- **Internet Forums**
  https://stackoverflow.com/questions/tagged/python and
  https://www.reddit.com/r/Python/
  – Most people are very friendly and will help if you write a well-phrased question and have done a basic Google search first

- *Success goes to the bold*; if you don't ask questions, you are unlikely to be successful!

18

18

A Final Note: Sharing Code

- Many email systems filter/remove .py files from attachments for security

- Rename .py files to .txt (or .py.txt) before sending them as attachments!

19

19

The Rest of our Time

- Template for Python

- Basic Python Operations

- **Advanced Examples:**
  - Modify a PDB B-factor Column (BioPython PDB parsing)
  - Protein denaturation fitting script (Numpy, Scipy, Matplotlib)

20

20