Python for Research Computation: An Intermediate Workshop

Adam Fine

September 3-11, 2024

1 Introduction

This workshop is a five-part series on various aspects of Python I wish I had known at the beginning of graduate school. Some familiarity with programming is assumed; familiarity with Python (as well as modules such as NumPy, matplotlib) will be helpful but not strictly necessary.

2 Rationale

Python has somewhat of a bad reputation. It's slow. It's badly designed - who thought duck typing was a good idea? why does the Global Interpreter Lock exist? Many people internalize these criticisms and resign themselves to writing bad, slow Python code. This workshop is intended as a defense of Python for research computing - in my opinion, Python is just enough of a programming language to be extremely powerful, while avoiding a lot of baggage that scientists don't care about. Also, it can be fast. Seriously.

3 Learning Objectives

By taking this workshop, students will be able to:

- Code in a manner that reduces and simplifies the debugging of errors.
- Implement computational pipelines in a rigorous and reproducible manner.
- Recognize potentially slow code and implement faster alternatives.
- Produce highly customized figures in matplotlib and Seaborn.
- Manipulate non-numerical data with pandas.

4 Schedule

1. T 9/3 - Why Python? + Introduction to Python
2. Th 9/5 - Defense: defensive programming, reproducibility, creating workflows with `snakemake`.
3. M 9/9 - Introduction to `NumPy` + how to write fast code.
4. T 9/10 - Figure... Tuesday? - plotting numerical data using matplotlib
5. W 9/11 - Handling categorical data with `pandas` and plotting it with `Seaborn`

All lectures will take place from 11 AM - Noon in GCIS W105.

5 Assignments

One Jupyter notebook's worth of homework per lecture. Optional. Each homework is intended to take 1-2 hours. I intend to provide feedback on all submitted homework.

6 Github

All lectures, homework, and miscellaneous files can be found on my github here: https://github.com/enifmada/research-python-workshop/tree/main