

CÓDE PROGRAMMING BOOTCAMP

May 2019

Instructor:	Tom Needham	Time:	MTuWTh 3:30 – 5:00
Email:	needham.71@osu.edu	Place:	Scott Lab E0103
Office Hours:	416 Math Tower	Office Hours:	MW 2:30 – 3:30

Course Description: This course serves as an introduction to the Python programming language, with a focus on techniques for data analysis. The intended audience is graduate students and postdocs who are studying in technical fields and who want a quick introduction to Python to aid in research and career development. No prior coding experience will be assumed.

Course Page: <https://github.com/trneedham/Python-Bootcamp-2019>

Course Requirements: You need to bring a laptop to class in order to follow along and to work on in-class assignments. **Important:** Before the first class, please download and install the Python 3.7 version of the Anaconda platform <https://www.anaconda.com/distribution/>. Please also clone the GitHub repository from the course page listed above.

Rough Schedule: The following is a rough list of topics and the order in which we will cover them. This schedule is subject to change depending on time and on student interest. I will do my best to accommodate student requests to cover topics not listed here.

Week 1: Python Basics

We will start with basic setup and usage instructions for for GitHub and Jupyter notebooks. This week will be focused on the basics of the Python language: expressions, operations, functions, conditionals, loops and dictionaries. Practice with basic syntax will be built up by exercises involving simulations and statistics.

Week 2: Statistics and Data Analysis

Now that we understand the basics of the language, we move on to understanding statistical algorithms. At first the focus will be on creating and analyzing our own algorithms. We will then integrate standard Python packages such as pandas, scikit-learn and matplotlib in order to analyze and visualize real-world data.

Week 3: Advanced Data Analysis and Command Line Programming

We will continue with more sophisticated techniques from statistics and machine learning. During this week we will move away from working in Jupyter notebooks to writing packages and working from the command line.

Week 4: Introduction Topological Data Analysis

In Week 4 of the bootcamp, instructors of all sections will present lectures on more specialized topics. For my special lectures, I will give a brief introduction to the field of *Topological Data Analysis (TDA)*. This rapidly-developing approach to data analysis uses techniques from metric geometry and algebraic topology to produce summaries of complex datasets. In two lectures, I will cover:

1. A light introduction to the mathematics behind TDA.
2. Barcodes and persistence diagrams: topological features of a data set.
3. Working with topological signatures: qualitative analysis, distance-based clustering and machine learning approaches.
4. Python packages for TDA such as Ripser and Persim.
5. Examples using TDA for shape classification and analysis of noisy images.