Prep Notes - Meeting 03

Dylan McDowell, Andrew Benson

April 26, 2018

Data camp has a number of courses. After doing some research it was determined that these courses would be the most needed for students to compete some of the homework assignments.

Below is an outline of the topics covered in the Data Camp Python Courses organized by course. This list will be updated as needed as additional courses are explored.

1 Intro to Python for Data Science (3.5 - 4 hours)

The Intro to Python for Data Science course covers many of the basics of procedural programming.

1.1 Python Basics

- Using Python as a basic calculator
- Variables and variable (data) types
- Comments in code
- Variable assignment
- Calculations with variables
- Python's type() function

1.2 Python Lists

- What is a list?
- List syntax
- Lists with different variable types
- List of lists
- Subsetting lists
- Slicing lists
- List manipulation
- Pitfalls of copying lists (idea of pass by copy vs. pass by reference, although I wouldn't expect them to understand this, explanation was too shallow to remember or apply this idea)

1.3 Functions and Packages

- How to use Python's help feature on a function.
- Multiple arguments
- String methods
- List methods
- Packages
- Package importing
- selective import

1.4 NumPy

2 Intermediate Python for Data Science (4 hours)

2.1 Matplotlib

Data Visualization is a key skill for aspiring data scientists. Matplotlib makes it easy to create meaningful and insightful plots. In this chapter, you will learn to build various types of plots and to customize them to make them more visually appealing and interpretable.

2.2 Dictionaries & Pandas

Learn about the dictionary, an alternative to the Python list, and the Pandas DataFrame, the de facto standard to work with tabular data in Python. You will get hands-on practice with creating, manipulating and accessing the information you need from these data structures.

2.3 Logic, Control Flow and Filtering

Boolean logic is the foundation of decision-making in your Python programs. Learn about different comparison operators, how you can combine them with boolean operators and how to use the boolean outcomes in control structures. You'll also learn to filter data from Pandas DataFrames using logic.

2.4 Loops

There are several techniques to repeatedly execute Python code. While loops are like repeated if statements; the for loop is there to iterate over all kinds of data structures. Learn all about them in this chapter.

3 Potential Areas of Supplement

This section will be a list of topics that are not covered in the Data Camp courses that students would need to know to either fulfill requirements of CS 101 or for this course.

- Data import
- Solving systems of equations