Quick Python for Data Analysis

2022

Author: Ramzi Saouma
Email: Ramzi@Saouma.org

Course Pages: Course website to be updated here.

Main References: This is a list of recommended texts which we will use continuously during the course. Some texts are more advanced compared to the rest, but all convey the essential concepts and ideas eloquently.

- Wes Mckinney, *Python for Data Analysis, 2e*, O'Reilly, 2017.

Recommended Pre-Course Reading: The following article presents a broad overview of the developments and milestones for the past 50 years of data science.

David Donoho, *50 years of Data Science*. Based on a presentation at the Tukey Centennial workshop, Princeton NJ, September 18 2015. Available online at:
https://courses.csail.mit.edu/18.337/2015/docs/50YearsDataScience.pdf

Objectives: This course is primarily designed for anyone who wants to quickly learn the basics of python in order to implement data analysis projects. Broadly speaking, this course is designed to enable students to appreciate the power of Python and the versatility of the packages and functions that are build for different type of applications. The first five sessions are designed to introduce the student to the basic concepts and packages of python in a relatively superficial matter. This should give the user enough knowledge and confidence to begin with data analytic projects. During the next five weeks we will study in greater depth some of the concepts we introduced earlier. We will also apply most of what we learn on study cases which should prepare the student to real-world projects.

Prerequisites: An undergraduate-level understanding of probability, statistics, graph theory, algorithms, and linear algebra is assumed.

Tentative Course Outline:

Week 1:
Part I: Python Installation

- Installing Conda and Packages
- Getting familiar with Jupyter

Part II: Introduction to Python

- Semantics and Objects
- Types and Operators
- Control Flow

Grading Policy: Homework (50%), Final Project (50%)