Python for Data Analysis

Summer & Fall 2020

# Instructor Information

|  |  |  |
| --- | --- | --- |
| Instructor | Email | Office Hours |
| Keenan M. | keenan@thepythonacademy.com | MWF – 8 to 8:30 pm, after class on Google Hangouts |

# General Information

## Description

This Course, Python for Data Analysis, is the next step to broadening your skills in Python. In fact, this is the first course that will allow you to be hired in the Python job market for data analysis. This course teaches you the foundations of data analysis with Python using industry-recognized libraries such as:

* Numpy and Pandas for data analysis and operations
* Matplotlib and seaborn for statistical visualization
* And Sklearn for regression analysis.

## Expectations and Goals

By the end of this class, you will able to analyze large datasets using python’s industry-standard libraries, and create meaningful, information-rich graphs that will further your analysis.

We expect you to attend class, pay attention, and do your homework. If you don’t do your homework or code along in class, you will fall behind and get frustrated. In return, we promise you to give you 100% effort on giving you the most up-to-date material and experience you will need to be successful in the data science field.

# Course Materials

## Required Materials

* You will need to have your own laptop or desktop (we only support troubleshooting for Windows, our staff has limited knowledge of MACs)
* You will receive a free copy of the “Python for Data Analysis” by Wes McKinney, the creator of the famous python library “pandas”.

# Course Schedule (next page)

**Week 1**

The following topics are covered in this class. All topics are covered with real life examples and applications of the topic. Homework is always given after the class and reviewed at the beginning of the next class.

* Welcome to NumPy!
	+ NumPy arrays
	+ NumPy Operations
* Introduction to Pandas
	+ Series
	+ Dataframes
	+ Missing data

**Week 2**

The following topics are covered in this class. All topics are covered with real life examples and applications of the topic. Homework is always given after the class and checked at the beginning of the next class.

* Introduction to pandas (continued)
	+ Group by
	+ Merging, joining, concatenating
	+ Operations
	+ IO
* Pandas exercises
* Python for Data Visualization
	+ Matplotlib
	+ Seaborn

**Week 3**

The following topics are covered in this class. All topics are covered with real life examples and applications of the topic. Homework is always given after the class and checked at the beginning of the next class.

* Linear regression with sklearn
* Geographical Plotting
* Data Capstone Project