Python for Predictive Modeling

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 Predictive Modeling, is the first step to obtaining the blue belt certification.  This course will teach you how to create advanced models with Python, including:

* Monte Carlo Simulation
* Linear and Logistic Regression
* Time Series Analysis
* K nearest neighbors
* And Principal Component Analysis
* Random Forest and Boosted Tree

## Expectations and Goals

By the end of this class, you will able to build upon your data analysis knowledge and build powerful model that can predict outcomes for your customers.

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 “Introduction to statistical learning” by Gareth James, freely available on the USC website.

# 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.

* Introduction to Data Modeling
* Monte Carlo Simulation
* Linear Regressions
* Cross Validation and Bias Variance Trade Off

**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.

* Logistic Regression
* K nearest neighbors
* Time series analysis
	+ Seasonal Decomposition
	+ Holt-Winters model
	+ Exponential smoothing

**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.

* Decision Trees and Random Forests
* Recommender Systems
* Natural Language Processing