Course Introduction

Course Agenda

- Python Crash Course
- Data Analysis:
  - **NumPy**
  - Pandas
- Data Visualization:
  - Matplotlib
  - Seaborn
  - Pandas
  - Plottly and Cufflinks
  - Geographical Plotting
- Machine Learning
  - Linear Regression
  - Logistic Regression
  - K Nearest Neighbors
  - Decision Trees and Random Forests
  - Support Vector Machines
  - K Means Clustering
  - Recommender Systems

NumPy

NumPy (or Numpy) is a Linear Algebra Library for Python, the reason it is so important for Data Science with Python is that almost all of the libraries in the PyData Ecosystem rely on NumPy as one of their main building blocks.

Numpy is also incredibly fast, as it has bindings to C libraries.

Numpy Installation

It is highly recommended you install Python using the Anaconda distribution to make sure all underlying dependencies (such as Linear Algebra libraries) all sync up with the use of a conda install.

If you have Anaconda, install NumPy by going to your terminal or command prompt and typing:

conda install numpy

pip install numpy

How we use NumPy

NumPy arrays are the main way we will use Numpy throughout the course.

Numpy arrays essentially come in two flavors: vectors and matrices.

Vectors are strictly 1-d arrays and matrices are 2-d (but you should note a matrix can still have only one row or one column).

Numpy Agenda

- Arrays
- Numpy Indexing and Selection
- Numpy Operations

Course Agenda

- Python Crash Course
- Data Analysis:
  - Numpy
  - **Pandas**
- Data Visualization:
  - Matplotlib
  - Seaborn
  - Pandas
  - Plottly and Cufflinks
  - Geographical Plotting
- Machine Learning
  - Linear Regression
  - Logistic Regression
  - K Nearest Neighbors
  - Decision Trees and Random Forests
  - Support Vector Machines
  - K Means Clustering
  - Recommender Systems