**Intro to Python for Data Science**

Version: 3.x

IPython: Interactive python

type(var)

**List**

* *l* = [0, 1, 2, 3, 4, …]
* Can contains mix of variable types. Can also contain list in them.
* Length: len(*list*)
* **Slicing** : *l*[ start\_index : (end\_index+1) ]
* **Adding elements**: list + [ new\_values, … ]
* **Removing elements**: del(list[index])
* List1 = List0 #reference
* List1 = list(List0) #element-wise copy

**Functions**

* help(func\_name)
* Methods: functions that belong to objects. In Python, **everything is an object.**
  + List Methods
    - List.index(item)
    - List.count(item)
    - List.append(item)
    - List.reverse()

**Packages**

* Directory of Python scripts. Each script=module
* from numpy import array <= Might cause confusion

**Numpy**

NUMeric Python. NumPy Array: Alternative to Python List

Perform operations over entire arrays. Easy and Fast. Easy to perform element-wise operations.

* np\_array = np.array(python\_list)
* Conditional Subsetting: np\_array[ condition on np\_array ]
* 2D Numpy Arrays
  + np\_2d = np.array(list\_1, list\_2, …)
  + np\_2d.shape
  + np\_2d[row][col] OR np\_2d[row, col]
    - slicing works
* Mean: np.mean(np\_array[ :, 0 ])
* Median: np.median(np\_array[ :, 0 ])
* Correlation: np.corrcoef(np\_baseball[:,0], np\_baseball[:,1])
* Binding Columns: np.column\_stack((col1, col2))