Python for Scientific Data Analysis

1. Syllabus

2. Course Topical Overview

We will cover the Basics of Python and Python Data Structures first, then we will discuss specific packages -- NumPy, SciPy, Matplotlib, Pandas, and AstroPy. The plan is to then finish the course discussing more advanced data analysis techniques -- e.g. PCA, K-Clustering, Hypothesis Testing, MCMC, and 'maybe' Machine Learning.

Below are some more details about these topics.

Basics of Python

- Printing/variables
- Prompting, Type Conversions, Argument Passing, and Reading/Writing
- Functions
- If-Then Statements, Looping

Data Structures

- Tuples, Lists, Arrays, and Dictionaries

`d=[5,6,7]` # a list

```
dtuple=tuple(d) #now a tuple
```

```
c=np.array([[1,2,3],[4,5,6]])
```

- Slicing

```
a[3,4,5,6]; b=a[1:2]
```

- Sequence Functions, Comprehensions, and Lambda Functions

```
zipped=zip(seq1,seq2)
```

NumPy

- NumPy Arrays

```
a=np.array([np.e,np.pi],dtype='float128')
```

- Array Arithmetic and Universal Functions

```
arr = np.array([[1., 2., 3.], [4., 5., 6.]]); arr2=1/arr
```

- Array Slicing and Reshaping

```
print(arr2d[:2, 1:]); arr.reshape(2,3)
```

- More Array Operations and Array Broadcasting

```
np.vstack([newarr,newarr2]); np.tile(arr,(2,1))
```

- Basic Linear Algebra with NumPy and SciPy

```
linalg.solve(a,b)
```

Basic SciPy

- Stats
- Optimization
- Root Finding
- Interpolation
- Signal Processing

Matplotlib

- Plotting Basics
- Subplots and Axes Configurations
- Shadings, Histograms, Contour Plots, and Images
- Other Plottable Things

Pandas

AstroPy

- FITS Files and Image Display (with Matplotlib)
- Units, Constants, and Coordinates
- AstroQuery (Working with Databases)
- Time Series Data
- Simple Photometry and Spectroscopy

Statistical Tests/Hypothesis Testing

PCA, K-Clustering

Markov Chain Monte Carlo Methods

??? (Machine Learning?)

3. Other Notes ...

- This is a new course!
- Computer Setup
- Markdown
- Jupyter Notebooks
- Github
- Feedback is strongly encouraged