`numpy` and `matplotlib`

Ben Rose

`brose@nd.edu`

June 12, 2018

Goals

1. Learn about `np.array`s and what makes them more useful mathematical vectors/matrices
2. Be able to read data in from disk via `numpy`
3. Be able to make a line or scatter plot
4. Be able to make a histogram

Review Homework

- Import `nothwind.txt` then separate by word
  - Try to count how many times each word appears
- Calculate the average sunspot form `sunspots.txt`
  - Can you count the days who's number of sunspots fell with an arbitrary range?

Day 04 Handout

5. Update dot product to use `zip` or `enumerate`

Convention

```
import numpy as np
import matplotlib.pyplot as plt
```

Numpy

Create an array

All of these create the same array:

```
a = np.array([0, 1, 2])
b = np.arange(3)
c = np.linspace(0, 2, 3)
```

Random points

```
from random import random
points = [[random(), random()]
          for _ in range(100)]
```

```
points = np.random.rand(100, 2)
```

How would you separate these out into *x* and *y* arrays?

Aside

What is `_` in `for _ in range(100)` mean?

https://dbader.org/blog/meaning-of-underscores-in-python

Math with arrays

Basic math: `a + b`, `np.sqrt(a)`

Basic stats: `a.mean()`, `a.std()`

- for multidimensional arrays, specify an axis if necessary: `a.mean(axis=0)`

Linear Algebra: `a.dot(b)`, `np.cross(a, b)`

Manipulating arrays

- Transpositions with `a.T`
- Slices similar to lists, `a[:,1]`

Read Data

- We can use our regular python version, or use `np.loadtxt()`

```
a, b = np.loadtxt(filename, unpack=True)
```

Let's plot!

Basics plotting

```
plt.figure('MyFirstFigure')
plt.plot(x, y)
plt.show()
```

Need help? http://matplotlib.org/api/pyplot_api.html

Labels

```
plt.xlabel('Smarts')
plt.ylabel('Probability')
plt.title(r'$\mu=100,\ \sigma=15$')
plt.axis([40, 160, 0, 0.03])
plt.grid(True)
```

You can use LaTeX!

Histograms

```
plt.hist(a)
```

Look up the options and this example.

Sunspots

Total observed number of sunspots for each month starting in January, 1749

- Read in `sunspots.tsv`, and plot it
- Change the plot to *not* be a connected line graph
- On the same plot, include a moving average (window = 5)

Git review

Practice

- Handout
- Read in `sunspots.tsv` and make a histogram of the counts/month
- Read in `sunspots.tsv`, and make a scatter plot of the data
- Add a moving average to you
- Read Newman Ch 5