Day 4: Numpy & Pandas Intro

Raymond Yee

January 30, 2014 (`http://is.gd/wwod1404`)

Goals Today

- how to use the *Python for Data Analysis*, textbook for the course
- getting oriented to the basics of `numpy`, `pandas`, and `matplotlib`

majot topics in PfDA

- pandas with numpy implicitly
- IPython notebook
- matplotlib
- nice example data sets which we will use (e.g., baby names)

Work through code in PfDA

A notebook I will talk about: Day_04_A_PfDA.ipynb

- I will highlight important concepts
- frankly, it took me a lot of trial and error in working through pandas, which has a large and complicated structure

one of my favorite tip sheets on Pandas: Pandas and Python: Top 10 - Curiosity

10-minute tour of pandas on Vimeo – impressive demo by Wes McKinney

Some Pandas videos to check out

PyData 2013 | New York, NY | Nov 8 - 10 video: A Practical Introduction to IPython Notebook and Pandas - Julia Evans on Vimeo

notebook as Wakari bundle: `https://www.wakari.io/sharing/bundle/jvns/PyData%20NYC%202013%20tutorial`

in context of some incredible resources:

- PyData's Videos on Vimeo
- SciPy 2013 :: Home

numpy: foundation of the scientific Python stack

Intro to NumPy on Vimeo

and slides: Introduction to NumPy (PyData SV 2013)

Thinking Through Numpy basics and Pandas Series

Day_04_B_numpy_and_pandas_series.ipynb

2010 census brief

Overview of Race and Hispanic Origin: 2010

nice tutorials on census

Mapping Census Data

Tutorials - US Census API

Our Census Notebook for Day 4

Day_04_C_Census.ipynb

Assignments / Homework

No new items to submit. But I do expect you to be wrestle with the materials in the `PfDA`, these notebooks, and talk to others in the class.

Please, please, please ask questions! :-)

Don't forget Day_02_A_US_Census_API.ipynb – due tomorrow 11:59pm.