Day 3: Setting Up Cont'd: Environments and Contexts

Raymond Yee

January 28, 2014 (`http://is.gd/wwod1403`)

Goals Today

- Administrative
  - Meet your tutor: AJ Renold
  - Schedule office hours – and correct location.
  - Let's keep using `http://bit.ly/wwod14classnotes`
- Basic orientation to Python environments
- Checklist: can everyone run a sample IPython notebook and upload to bCourses
- My goal: to make WwOD FUN, ENGAGING, and NOT OVERWHELMING

I want people to ask questions and flag stuff that's confusing

Meet AJ Renold, the Course Tutor

AJ is wonderful: kind, smart, and responsive. Get to know him.

His role:

- maintain activity on canvas and to create useful examples for students who are struggling with elementary python, shell, and git.
- holding office hours
- helping RY with eventually grading assignments (or working on automation tools)

Assumed Background Knowledge

- Certain skills that would be very useful for us to work on if they don't already have. I will actually use them in course.

The things I have in mind include:

- elementary Python
- working with the Unix shell
- source control, specifically git and github

Material you'll find in http://software-carpentry.org/v4/index.html for example.

- Please flag things that are unclear by posting questions to bCourses.
- Work with AJ to get up to speed.
- Should many people struggle with certain skills/knowledge areas, we'll consider providing more training and taking up some issues during class.

Office Hours

Thanks to everyone who took the Doodle: Office Hour choices. I'm going to leave my office hours at Tues, Thurs 3:30-4:30pm in **302 South Hall**. We start today and run until the end of the semester – unless I make an announcement canceling an office hour.

I'm willing to schedule other meetings and virtual office hours as the need arises. (For example, I'm a fan of Google Hangout, and the campus has a new video-conferencing system we can try out or the Web Conferences in Working with Open Data.

Faculty Office Hours | School of Information

For info about AJ's office hours, see bCourses: AJ's Office Hours for Working With Open Data.

Assignments from Last Week

Working with Open Data: Assignments – about 38 people submitted assignments.

When can I expect the result of you to submit assignments?

For this round, I will accept late submissions.

Results of survey of student background: Survey about Technical Background: Statistics

Fundamental Concepts or Things we Use/Components/Concepts

Lots of things happen when we run a Python program or run a cell in an IPython notebook. Most we can take for granted, but sometimes we need to understand some of the underlying complexity.

Much work goes into allowing us to work with **abstractions** but beware The Law of Leaky Abstractions.

The core of this course is using **Python** to work with **data** (particularly **open data**.)

But consider, for a moment, all the things associated with the "simple" act of running a Python program or an IPython notebook:

- the actual version of Python we are using (v 2.7.6 vs. v 3.3 vs. v 2.6 – what difference does it make?)
- operating system used
- using a Python shell, a command line, inside of IPython Notebook, etc?

RY's philosophy for learning how to work with data

I have a philosophy to guide ourselves in this course

- philosophy of immersion: often we jump in and then figure stuff out as we go.
- developing toeholds of the mind
- working code is gold – hang on to it. . . and leverage it and work hard to keep things working code working
- confusion is par for everybody, novice and expert alike
- experts learn to deal with confusion.
- Hang in there – your persistence will pay off: Things I Wish Someone Had Told Me When I Was Learning How to Code ? Learning to Code ? Medium
- coding is social practice: reflecting on Introduction to Python for Librarians

Learning to Program is a bit Like learning a natural language

- You got to practice to develop fluency
- We learn together
- We learn by making mistakes

Key Concept for Today: Execution Environment of Python

What is an execution environment? Python Essential Reference, Fourth Edition> Execution Environment : Safari Books Online

Python: (Core Language + Standard Library) + other installed packages

- versions matter: version of Python Python versions + versions of packages
- Python core language + any Python packages installed –> *often* run across different contexts (operating systems)

Typically, the quotidian question we have is: **how to install a given Python package**

Major Alternatives to Package Installers

I would say that pip is a central tool for installing packages – especially because of the great PyPI "repository of software for the Python programming language."

But there are times, it's actually really hard to get certain packages installed. . . and there's where I really appreciate Python distributions along with their package installer.

For this course, we recommend Anaconda: see IPython Installation Options · rdhyee/working-open-data-2014 Wiki.

(BTW, anyone can edit the course github wiki).

Wakari vs Anaconda

I recommended Wakari to start because it's easy to set up: create an account on wakari.io and you can go...

Think about the trade offs betwen running on your laptop (Anaconda) vs running on a virtualized Python environment running on the continuum.io servers (Wakari)

Wakari is good to have around but I think everyone will be happier having IPython notebook running on their own laptops.

I encourage you to install the anaconda python distribution even if already have another one running.

Conda

Conda — Continuum documentation is a Anaconda-specific alternative to virtualenv/virtualenvwrapper.

Conda [to quote the Conda docs]:

- "primary interface for managing Anaconda installations."
- "It can query and search the Anaconda package index and current Anaconda installation"
- "create new Anaconda environments"
- "install and update packages into existing Anaconda environments"

Examples — Continuum documentation

How I use conda

- In my specific case, I'm running OS X 10.6.8. When I downloaded and installed anaconda, all my anaconda files end up in `~/anaconda`.
- A really nice feature: "installs into a single directory and doesn't affect other Python installations on your system. Doesn't require root or local administrator privileges."
- On my Mac, the key to swapping between using Anaconda environments and my virtualenvs is through **manipulating the $PATH environment variable**. If I want to use anaconda, I need to get `~/anaconda/bin` to beginning of `$PATH` – see Anaconda FAQ — Continuum documentation conda environments
- I have an alias defined in my `~/.profile` to make this easier to do on the fly:

```
# have alias for using anaconda
alias use_conda='export ANACONDA_HOME=$HOME/anaconda; e
```

- My conda environments

installing census package

Wakari custom environments

Custom Python Environments in Wakari

Geographical hierarchies in Census

use of FIPS codes in census

Time to show project ideas? Yes

- campus map at very least

Next steps with world population

List of countries by population (United Nations) - Wikipedia, the free encyclopedia

get list of wikipedia URLs and flags

API vs bulk download – constant theme

"Simple" Maps We'd Like to Make

- World Map (with various useful projections), in which we can color countries and hover over to get info –> nice to get ability to pick center and zoom level
- state map of the USA –> choropleth with hover info
- county map of the USA
- leaflet.js / Google Map API with
- plotting a few dots with bubbles
- plot a few polygons on realistic background

Later: making our own tiles

Various Mapping Services

To learn about the world

I looked at Maps & Earth Apps: Top 5 Free Alternatives to Google Maps and23 Google Maps Alternatives [time for update :-)]

- especially because of generous non-commercial limits for Bing Maps API Profile
- Yahoo Maps, Driving Directions and Traffic –> but Yahoo Maps API: So Long, Old Friend
- Wikimapia / About – WikiMapia - Wikipedia, the free encyclopedia –> new to me. . . not related to Wikipedia or OSM as far as I can tell but very interesting project to look at Wikimapia API. (To my untrained eye, it looks like Wikimapia is using OSM data. . . .is that right? is it legal? Wikimapia mass-uploading OpenStreetMap data (Page 1) / Questions and Answers / OpenStreetMap Forum)
- Google Earth and Google Code Playground and Examples - Google Earth API ? Google Developers
- I remain curious about Features | ArcGIS for Developers. also

sidebar: How to learn about countries

- draw your own maps by hand
- jigsaw puzzle
- online games e.g., World Maps - geography online games

Google Maps

Classic Google Maps URL

alternatives to Basemap

I shouldn't give up on Basemap. . . .

So You'd Like To Make a Map Using Python looks promising but I'll have to work out how to do the installation via ansible on Linux (probably). . . don't think I can rely on conda alone. Why? Use of libraries to compile like Fiona 1.1 : Python Package Index, Shapely 1.3.0 : Python Package Index, descartes 1.0.1 : Python Package Index, and Welcome ? pysal v1.6.0 Reference Guide

Also try Enthought. . . .

Simple Start: Use Matplotlib and do simple projection

Equirectangular projection - Wikipedia, the free encyclopedia

from List of map projections - Wikipedia, the free encyclopedia

Come back to my early examples of displaying HTML/CSS/JS in IPython notebook

nbviewer.ipython.org/gist/rdhyee/6131622

World Map: Google Maps Charts

Visualization: Geochart - Google Charts ? Google Developers

Google Code Playground

Steps:

- make it work separately
- ideally turn into a Python class that has *repr*html_ to make it show up in IPython notebook using nbviewer.ipython.org/github/ipython/ipython/blob/1.x/examples/no 5 - Rich Display System.ipynb
- possibly package up into wwod package

d3.js geo

vega /vincent geo world map

Vega Live Editor:Choropleth example

easy to run locally – I cloned the repo (trifacta/vega) and ran to get http://localhost:8000/examples/editor/

Leaflet.js: Hello World

Quick Start Guide - Leaflet - a JavaScript library for mobile-friendly maps

Leaflet.js: Choropleth for states (and counties?)

Interactive Choropleth Map - Leaflet - a JavaScript library for mobile-friendly maps

Census Data – we have experts in our class!

cmgerber/CensusMapper

CensusMapper/python/acs_extract.py at master ?
cmgerber/CensusMapper

NHGIS

I have an IPython notebook on this. . . .and had downloaded some data. . . .find it

Data Availability | NHGIS

`/Users/raymondyee/D/Document/Working_with_Open_Data/working`

Where I had downloaded from: NHGIS data finder

`/Users/raymondyee/Downloads/nhgis0002/nhgis0002_shape/nhgis`

John MacKenzie, UD, has good census data

Mapping Census Data

http://www.udel.edu/johnmack/frec682/census/

census_form.png (1106?872)
http://www.udel.edu/johnmack/frec682/census/geog_hierarchies.png

The Intrepid Data Nerd has 2010 shp files

Freebase as a source of location data

Commons/location/location - Freebase

example of geocode query:
`https://www.googleapis.com/freebase/v1/search?indent=true&filter=%28all+type%3Arestaurant+%28within+radius%3A1000ft+lon%3A-122.39+lat%3A37.79056%29%29&output=%28geocode+practitioner_of%29`

via Search Output - Freebase API ? Google Developers

also. . . look up my county location example from wwod13. . . still works?

Come back to (99+) Freebase Discuss - Google Groups and Freebase - Google+ to ask whether Freebase is a good source for location data.

What would JakeVP do?

D3 Plugins: Truly Interactive Matplotlib In Your Browser # working with topojson

Command Line Reference ? mbostock/topojson Wiki thematic mapping blog: Converting shapefiles to TopoJSON + a GitHub secret

Looking Ahead

Would like to plot population density of countries: scrape and plot data from List of sovereign states and dependent territories by population density - Wikipedia, the free encyclopedia

Figure out other stuff

- Mapbox | Design and publish beautiful maps
- Create beautiful dynamic data driven maps | CartoDB
- TileMill | Mapbox (and v2: . . .)

How to build your own geo-stack

In Class Activity

Assignments / Homework