

15-388/688 – Practical Data Science: Jupyter Notebook Lab

Eric Wong
Carnegie Mellon University
Fall 2016

Outline

Setting up the Jupyter notebook environment

Using Jupyter notebook (tips & tricks)

Common questions

Announcements

Project proposal is due 10/19 (maximum 1 page and not more than 500 words). Use Piazza's 'Search for Teammates!' post to find a group of 2-3.

Lunch signup: will be released shortly on Piazza.

Reminders

Homework 1 is due next Wednesday, September 14

Office hours are [updates or changes will be announced on Piazza]:

Monday 3:30-4:30pm [Dhivya, GHC 6008]

Tuesday 3:00-4:00pm [Eric, GHC 8213]

Thursday 3:00-4:00pm [Eric, GHC 8213]

Friday 9:00-10:00am [Dhivya, GHC 6008]

Anaconda

For this course we recommend that you use Anaconda, a data science python platform that will contain most of the tools you will need in this course (including the Jupyter notebook environment).

Download the Python 2.7 version here:

<https://www.continuum.io/downloads>

Verify that you're using the Anaconda platform by firing up a Python interpreter.

```
$ python
Python 2.7.12 |Anaconda 2.5.0 (64-bit)| (default, Jun 29 2016, 11:07:13) [MSC v.1500 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license" for more information.
Anaconda is brought to you by Continuum Analytics.
Please check out: http://continuum.io/thanks and https://anaconda.org
```

```
>>>
```

Additional Packages

Sometimes we'll need more than the default packages (<https://docs.continuum.io/anaconda/pkg-docs>) that come with Anaconda.

If there is a conda package for it, then it is best to use that. For example, **pymc** is not included with your default installation, however you can install it easily with

```
conda install -c pymc pymc
```

For the first assignment, you'll need yelp-python which is not in the Anaconda package. You can install it via PyPI using

```
pip install yelp
```

Jupyter Notebook

Assignments for this class are distributed in Jupyter notebooks. These let you create documents that contain a mixture of code and text organized in cells.

Read more about Jupyter here: <http://jupyter.org/>

For now, we will soon jump into a live Jupyter notebook instance and go through the first problem of the assignment. To view notebooks, you'll need to run the following command and go to <http://localhost:8888/tree>

`jupyter notebook`

Jupyter Notebook (continued on notebook)

Writing code (modifying cells, adding and deleting cells)

Running code (running all cells, restarting or interrupting the kernel)

Testing code (writing tests, timing code, viewing output)

Plotting graphs (matplotlib, ipython magic functions)

Useful shortcuts

Escape / Enter: exit or enter edit mode

Shift-Enter: runs current cell

A: insert cell above

B: insert cell below

L: toggle line numbers

Ctrl-S / S: save

0, 0: restart kernel

Find more at Help -> Keyboard Shortcuts

Homework Tips

Carefully follow the homework problem specifications to match the output required by Autolab. (the scraper grader accepts both unicode and ASCII text output)

Test your code locally on more than just the examples we give you
Read the Autolab output to debug submissions and missing scores