Intro to Data Analytics and Visualizations

Lecture 11 – Intro to Python for Data Analysis
Fall 2014, September 22

Outline

1. Why Python
2. Install Python
3. Install editors
4. Install pandas
5. Intro to Python Basics

Why Python

- Can use for multiple tasks: data analysis, data handling, general purpose programming, web apps
- Good data wrangling/munging capabilities
- Many programmers are familiar with it as they traverse into the Data Science field coming from pure Software Development

3

Python Installation

- Install Enthought Canopy, by following the book instructions.

https://store.enthought.com/#canopy-academic

- Choose the "full" installation that comes for free if you are in the academic field; make sure you sign up with your vt email address; that will give you access to the "full" version of Enthought Canopy.
- Follow the instructions at pages 8-9 of your textbook, for Windows and Mac respectively.
- We are using the Python 2.7 version.

4

Python Installation for MAC

1) Install Xcode (from APPLE site);
2) Launch Terminal (Applications > Utilities);
3) Write gcc (and press Enter); Should see: "i686-apple-darwin10-gcc-4.2.1: no input files".
4) Install Enthought Canopy full installation provided free for Academics.

https://store.enthought.com/#canopy-academic

- Choose the "full" installation that comes for free if you are in the academic field; make sure you sign up with your vt email address; that will give you access to the "full" version of Enthought Canopy.
- Follow the instructions at pages 8-9 of your textbook.
- We are using the Python 2.7 version.

5

Test your Enthought Canopy Python Installation

Windows:

*GO to Start;
*Run CMD;
*In the CMD window, write: python
* Press Enter.

MAC:

*Open Terminal
* Write: Python
* Press Enter.

You should see something similar to: "*Python 2.7.3 |EPD_free 7.3-1 (32-bit)| (default, Apr 12 2012, 14:30:37) on win32*
Type "credits", "demo" or "enthought" for more information."

6

Editor Installation

- For Windows, download and install Notepad++ from

http://notepad-plus-plus.org/

- For Mac, download and install TextWrangler from

http://www.barebones.com/products/textwrangler/download.html

Pandas Package Installation

- Go to:

https://pypi.python.org/pypi/pandas
and read the package description.

Download:

- pandas-0.14.1.win-amd32-py2.7.exe (md5) if you have Windows 32 bits or pandas-0.14.1.win-amd64-py2.7.exe (md5) if you have Windows 64bits. Double click on it and run it.
- Follow the instructions on textbook's page 9 to install pandas for MAC.

Test Your Pandas Installation

Windows:

1. Go to Start
2. Write “cmd”; window opens.
3. Write: ipython –pylab
4. At the numbered line prompt write:

```
import pandas
plot(arange(10))
```

MAC:

1. Go to Terminal
2. Write: ipython –pylab
3. At the numbered line prompt write:

```
import pandas
plot(arange(10))
```

Test your Editor: First Python Program

Windows:

1. Open Notepad++ application.
2. Create new file “PythonBasics”. Save it with .py extension.
3. Write

MAC:

1. Open TextWrangler application.
2. Create new file “PythonBasics”. Save it with .py extension.

Note: Save your file in the folder with iPython installation; For me that is: C:\Users\Denisa

Comments and strings print

- *Use your PythonBasics.py file. It should be open in your editor of choice (Notepad++ or TextWrangler).*
- *Copy the following:*

```
#first Python program
#Run with iPython
print "Welcome to Python!"

"""
There is something going on here.
With the triple quotation marks.
We'll be able to type comments.
How many lines we want. """
```

Comments and strings print (cont.)

1. Copy and paste the program exactly as presented.
2. Notice that in Python extra space/indentation has meaning!!
3. Notice two ways to create comments throughout your code
4. Notice that strings/characters are introduced by “” just like in R.
5. Command “print” just prints the string to the output window, similarly to R.

How to run your program

Windows:

*GO to Start;

*Run CMD;

*In the CMD window, write: iPython

*Write: %run PythonBasics.py

MAC:

*Open Terminal

*Write: iPython

*Write: %run PythonBasics.py

How to Exit

* To exit iPython environment: exit()

* To exit CMD (Windows) or Terminal (MAC): exit

*To exit editors (Notepad++/TextWrangler):

-Save program in the correct directory;

-Close window.