Python For Data Analysis
UCSC Extension, Silicon Valley
May 13 - 16, 2019

Marilyn Davis, Ph.D.
marilyn@pythontrainer.com
(650) 814-4435

This Course

This course is designed for programming for novices, that is, people who have absolutely no previous programming experience. However, you are expected to have high school level knowledge of simple statistics.

Your responsibility is to ask questions, even the same question over and over.

Programming is a lot like dancing, you have to practice, practice, practice. And, it takes time for new concepts to sink in and become natural. This short class is not nearly enough. But hopefully, it will be enough to discover your affinity, and awaken your talent, for programming to analyze data; and the art of data analysis via Python, itself, will motivate you to learn and do more.

Some Resources

- To brush-up on statistics:
  `https://sites.google.com/a/ahapps.anoka.k12.mn.us/blaine-hs-probability-and-statistics/textbook`
- To help you get started with Python:
  `http://wiki.python.org/moin/BeginnersGuide/NonProgrammers`
  has some
- *The Quick Python Book Second Edition*, by Vernon L. Ceder, ISBN # 9781935182207, published by Manning is my favorite for learning Python. Also, it is Guido Van Rossum's favorite. Guido is the author of Python.

- *Python For Data Analysis* by Wes McKinney, ISBN # 9781491957660, published by O'Reilly. Wes is the mail developer of the `pandas` library.

Getting Ready

For this class you need the Jupyter (iPython) Notebook Environment, the class material, and to link them together.

The Jupyter Notebook Environment

The best way to get the environment, without loading in libraries you don't need, is to install the Python 3 version here:

`https://conda.io/miniconda.html`

After the install:

Windows, you should see an "Anaconda prompt" on your **Start** menu. Use that:

or

On **Mac**, after the install, use the OS prompt:

```
conda install pandas

conda install numpy

conda install matplotlib

conda install jupyter

conda install xlrd
```

And then:

jupyter notebook

Class Material

1. Login into the class site. The url starts with 'https://ucsc-extension.instructure.com/courses/
2. On the left you see a menu. Go to *Files* and download 'PDA.zip'. *PDA* stands for Python For Data Analysis. `PDA.zip`. *PDA* stands for Python For Data Analysis.
3. Put the `PDA` directory where you want to do your class work. It will be your home directory for the class.

Linking The Class Material and Jupyter Notebook

1. See:
   `http://jupyter-notebook-beginner-guide.readthedocs.io/en/latest/execute.html`
   to change your Jupyter Notebook startup folder to be your `PDA` directory.

   On **Windows**, the directions are sketchy. You also need to right-click on your new Anaconda prompt, on your Desktop; then pick **Properties** to find the **"Target Field"** so you can make the required change. Be sure to click on **Apply**.

2. At your OS prompt:

   ```
   jupyter notebook
   ```

3. Click on:

   `Lab01_DataTypes.ipynb`

and let's get started.