

**Course:**

Foundation of Data Science – CS-GY 6053  
Fall 2021

**Instructor:**

Aline Bessa, PhD

**Homework 1****1. Install Jupyter Notebook**

The Jupyter Notebook is an interactive Python programming environment. The main instruction page for Jupyter is here:

<http://jupyter.org/install.html>

The Anaconda distribution is highly recommended in order to install Python, the Jupyter Notebook, and other commonly used packages for data science.

If you have a problem, please speak to the TA so that you can have this installed before next class.

**2. Jupyter and Python overview documents:**

If brand new to Jupyter, this page contains details resources:

<http://jupyter-notebook.readthedocs.io/en/latest/>

There are several ways to get more familiar with Python. The O'Reilly book suggested for the course is a nice handbook. Other good resources include:

<https://docs.python.org/3/>

<https://stackoverflow.com/>

**3. Review the following notebooks to catch up on Python basics (some we also reviewed in class).** The following describes basic Unix commands and core Python libraries that will be helpful for use in this course. The notebooks are condensed tutorials compiled from several guidebooks.

<http://nbviewer.ipython.org/github/jattenberg/PDS-Spring-2014/blob/master/>

[ipython\\_notebooks/](#)

[Basic%20Unix%20Shell%20Commands%20for%20the%20Data%20Scientist.ipynb](#)

[Python Intro: ExampleNotebooks/Lecture\\_SimplePythonExample\\_3\\_6053Fall21.ipynb](#)

[Python Numpy: ExampleNotebooks/Lecture\\_NumPyBasics\\_3\\_6053Fall21.ipynb](#)