Introduction to Data Science

(Lecture 4)

Dr. Mohammad Pourhomayoun

Assistant Professor

Computer Science Department

California State University, Los Angeles

Python Programming Language

- In this course, we work with **Python** and its libraries.

Python Programming

- In this class, we will briefly review the basics of python programming for beginners, and then we will more focus on python libraries and packages (such as scikit-learn machine learning library, …).
- For Python beginners, we highly recommend to boost your programming skills using at least one of the following references (or hundreds of other free tutorials available online):
  - Python Doc: https://docs.python.org/2.7/tutorial/index.html
  - Codecademy course: www.codecademy.com/en/tracks/python
  - Google's Python class: https://developers.google.com/edu/python
  - Python for Informatics: www.pythonlearn.com
  - …

Programming Expectations

- As a computer science student, you should feel comfortable with:
  - looking up python syntax in StackOverflow or other similar websites as needed.
  - checking Python documentation to learn more about syntax, commands, libraries, methods, … as needed.

Jupyter Notebook

Jupyter Notebook

- **Jupyter Notebook** (previously known as ***ipython Notebook***) is a *web-based interactive development environment*, in which you can combine code execution, text and notations (markdown), mathematics, and plots.

- The Jupyter Notebook will be run under your *browser*. It supports a number of programming languages including *Python, Julia, Octave, and R*.

- See http://jupyter.org/ for more information …

- In this class, we will be using Jupyter Notebook to do your homework and projects.

How to install Jupyter Notebook?

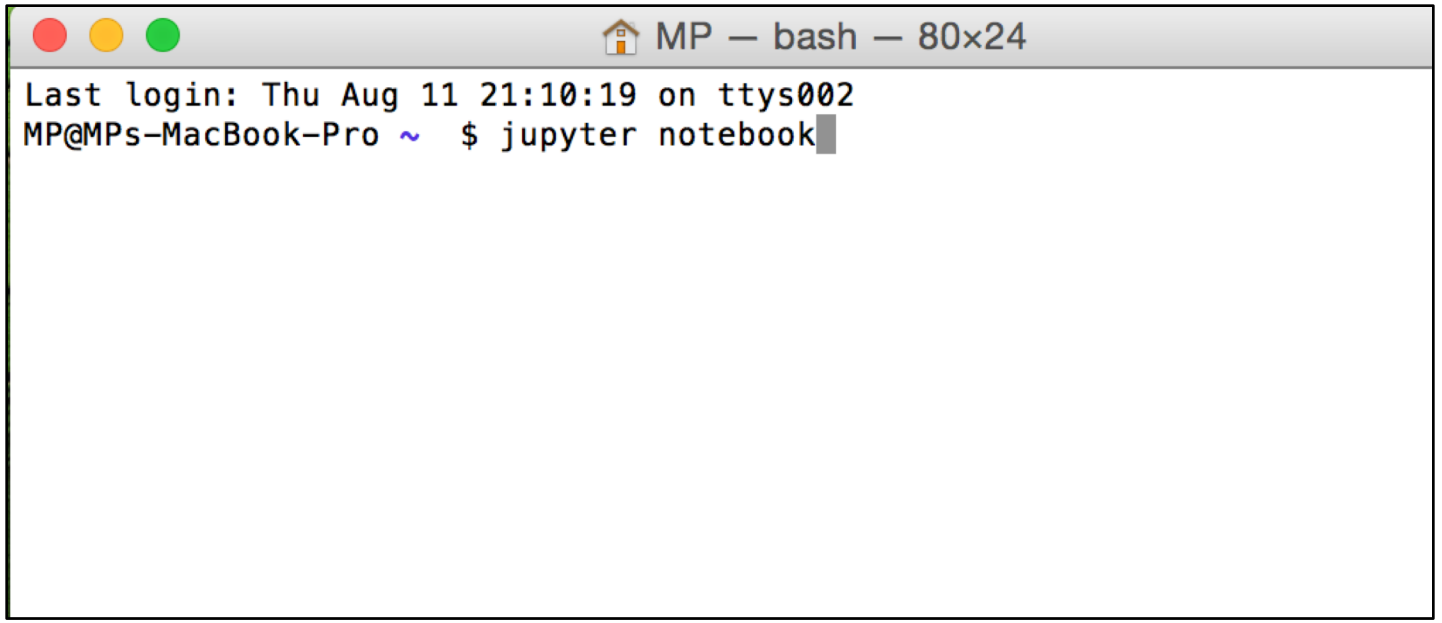
- There are various ways to install Jupyter. However, for new users, we recommend installing **Anaconda** package that includes Python and Jupyter.

- Anaconda is a free platform that easily installs *Python, Jupyter Notebook*, and many other popular packages for data science such as *Scikit-Learn, Numpy, Pandas, Scipy, …*

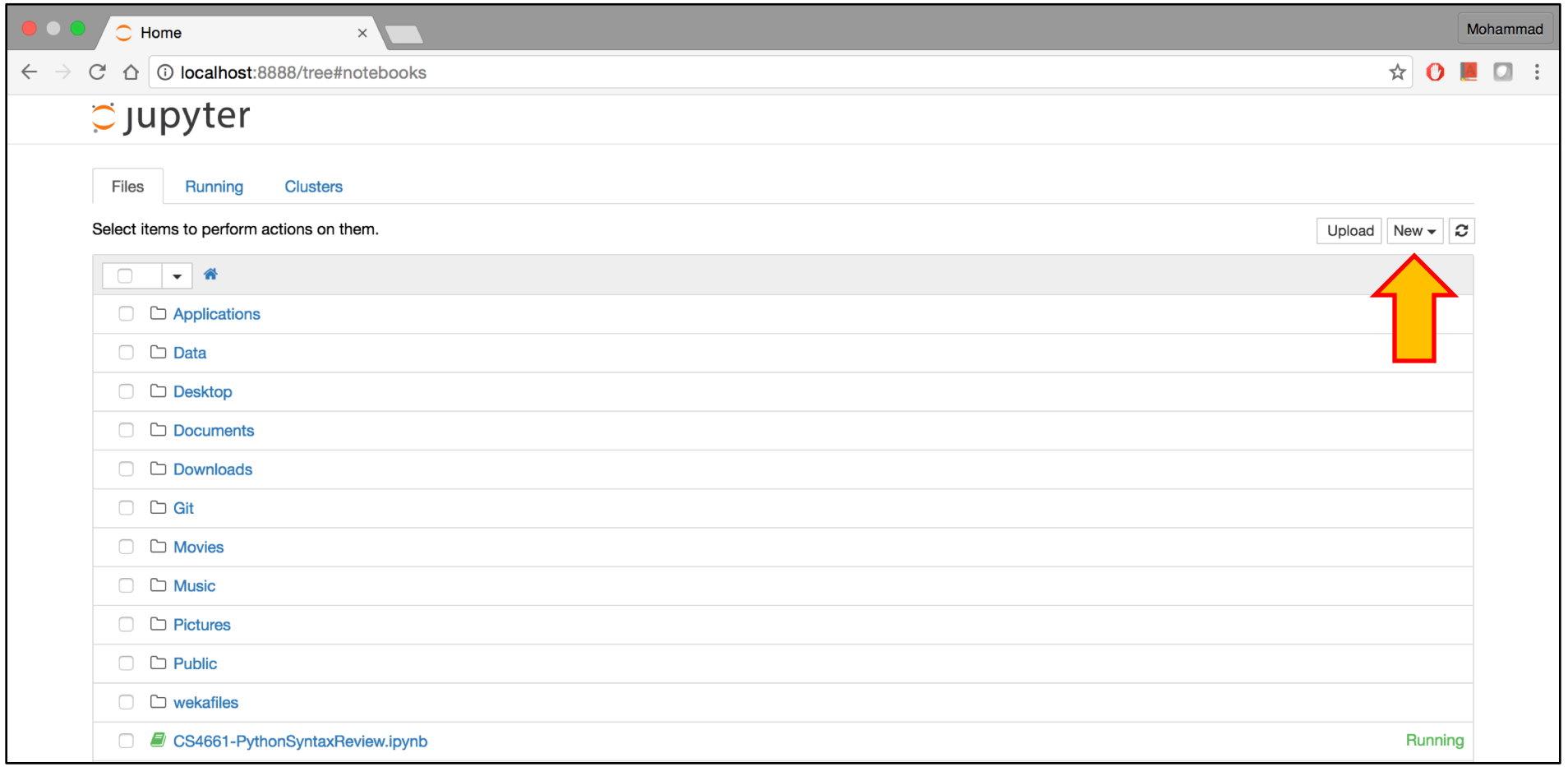
- Installation steps:
  - Go to https://www.continuum.io/downloads
  - Download and Install ***Anaconda*** with ***Python 3.x version***.
  - Now, you have Python, Jupyter Notebook, and many other useful data processing packages!

How to run Jupyter Notebook?

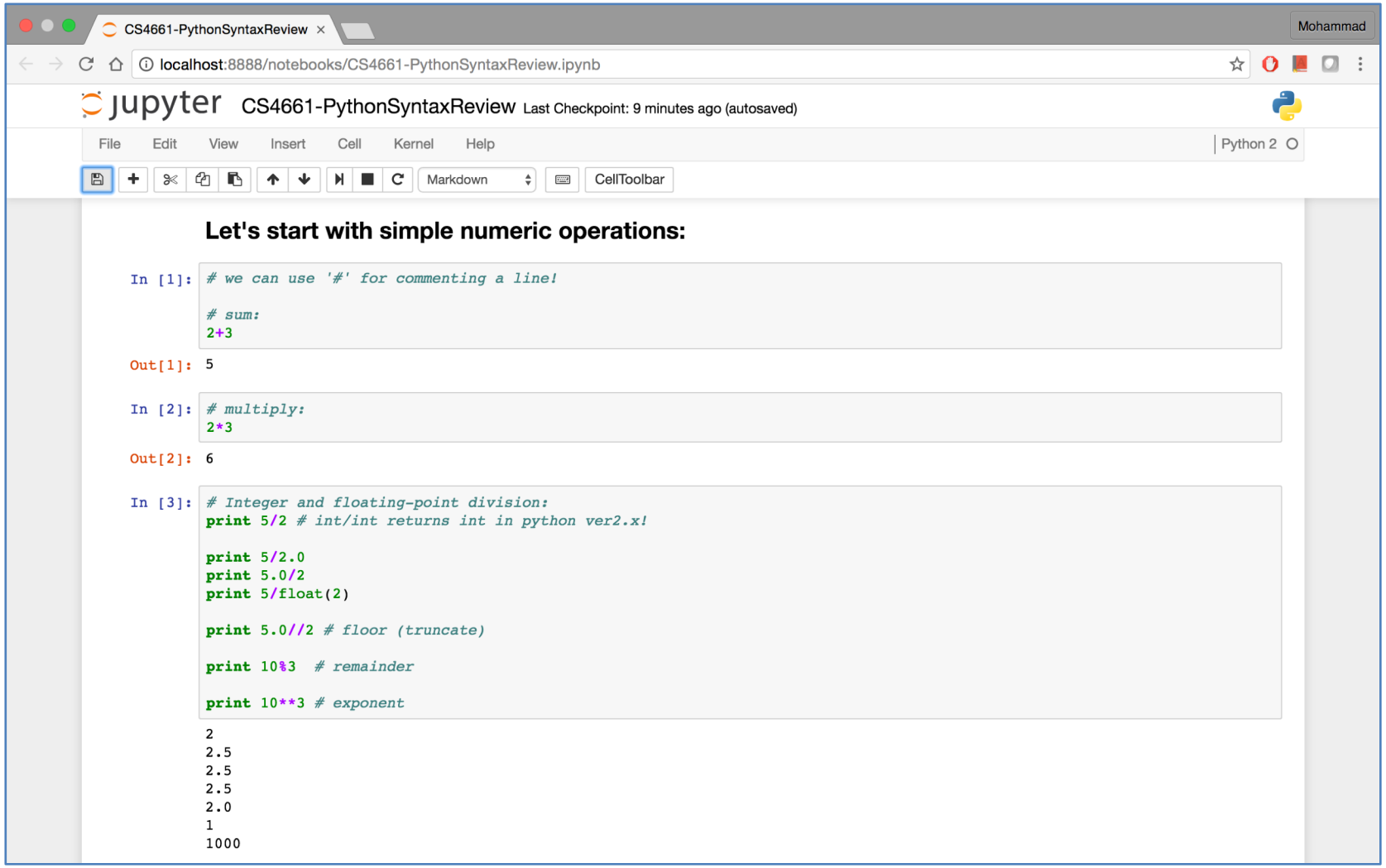
- To run the Jupyter notebook:
  - Type *"**jupyter notebook**"* at the terminal (command line) to open the dashboard.
  - Don't close the terminal window while the Notebook is running.



Jupyter Notebook



Jupyter Notebook



Jupyter Notebook

- The Jupyter Notebook include cells. Each cell can be executed independently. You can add/remove cells, or arrange the order of them.

- A piece of code, or several lines of the code that we want to run together can be written in a single cell.

- Now, Let's open file ***CS4661-PythonSyntaxReview.ipynb***

  in Jupyter notebook to continue python tutorial.