

Topic Outline: Pandas 1: Data Input

Revised: March 2, 2016

Materials

- Today's handouts: this outline, book chapter, code practice, stickers
- All posted on *Topic list & links* page of website (except the stickers).

Preliminaries

- Tools and buzzwords
 - Packages, import, dataframe, csv files, series, columns, index, API.
- What is this?
 - `x = 'abcd'`
 - `x[1]`
 - `'abcd'[1]`
 - `['x1', 'x2']`
 - `d = {'x1': [1, 5, 3], 'x2': [7, 2, 4]}`
 - `d['x1']`
 - `1 > 0`
 - `type(1>0)`
 - Tab completion
 - `object.method?`
 - Function returns
 - Objects and methods
- Prelaunch checklist
 - Start Spyder
 - Open blank file `scratch_pad.py` to use in class
 - Save Python code `bootcamp_pandas-input.py` in your `Data_Bootcamp` directory
GitHub ⇒ Code ⇒ Python (second half)

Pandas 1: Data input

We'll follow the book chapter.

- Packages
- Importing packages
- Reading internet files
- Properties of dataframes
- Working with variables
- Dataframe methods
- Reading files from your computer
- Examples: Penn World Table, World Economic Outlook, PISA education data, UN population data and projections, income by college major
- APIs: FRED, World Bank, Fama-French equity returns
- Review
 - Put red sticker on, replace with green when done

Examples

If time permits, we'll go through some examples of economic and financial data and its properties. It's an IPython notebook that you can view on the GitHub repo: the `Code/IPython` directory, file `bootcamp_examples.ipynb`.

After class

- Required
 - Code Practice #3 (and **note the code file we posted**)
- Recommended (always)
 - **Write:** After class, write down everything you remember without using your notes
 - **Review:** Reread the chapter and fill in anything you missed.
 - Ask for help if you get stuck.