

Working with Pandas

Instructions

Aloha! MDST is trying out a new format for team meetings. There's gonna be an interactive tutorial at the beginning of the meeting and as part of it we ask that you find 20 minutes or so to prepare a little in advance so we can make the most of our time in the meetings.

First off, you're gonna want to do the following ...

1. Copy `test.csv` and `train.csv` into the folder called `data`.
 - You can just copy the data from the top level of the `tmp` directory into the `./data` folder here.
2. Install the IPython environment, as well as the `matplotlib`, `scipy`, `sk-learn`, and `pandas`.
 - New to programming in python? Fear not, just install [Anaconda \(https://www.continuum.io/downloads\)](https://www.continuum.io/downloads). It's a scientific computing package for python that contains basically everything you'd need for MDST.
3. Once you've set up the environment you can test to see if it's working by navigating to the project directory and running `ipython notebook`. Your browser should pop open and show the directory in the IPython Notebook web interface.

If you have any issues with the set up you can obviously drop either me (thealex@umich.edu) or Jared (jared.a.webb@gmail.com) an email and we'll get back to you as soon as we can. Additionally, anyone who still needs help can show up a few minutes early to the 5pm meeting and I'll be ready to help ya'll out.