**PUI 2015 HW 5. Due W 10/14 at 5:00PM. Automatic scripts will fork your repositories at deadline time, no excuses or delays accepted.**

**Reading:**

Introduction to Statistics, David Lane
Chapter 15: ANOVA
http://onlinestatbook.com/2/index.html
Generally i try to give you a lighter reading than just a chunk of a text book, but i think you should be exposed to ANOVA, and we are running short on lecture time to cover it. In addition several of you have asked for a more basic reference then Statistics in a Nutshell: This is an e-book, it is free, and i think it is quite clear! if you have not had a lot of exposure to statistics before this is a better reference than Statistics in a Nutshell

**I have been trying to be more clear in the notebooks, but i also have included fewer pieces of code. In order to not to confuse you with a double explanation I am not including instructions here, just follow the notebook. Let me know if this is a better way to give you instructions.**
Use version control properly: do not wait until the last minute to upload but do frequent uploads at various stages of your work!
Remember your axes labels, with units when possible, add "captions" to describe your figures (in the cell below the figure), try to be PEP8 compliant.

**Assignment 1 : Data cleaning and merging, linear regression (1st and 2nd degree polynomial), log-scales, likelihood tests.**

Following the ipython notebook

https://github.com/fedhere/PUI2015_fbianco/blob/master/HW5/building_nrg_instructions.ipynb

**Assignment 1 : Regression, predictions**

Follow the ipython notebook

https://github.com/fedhere/PUI2015_fbianco/blob/master/HW5/income_by_gender-instructions.ipynb

**the data can be accessed here:**
**http://cosmo.nyu.edu/~fb55/UI_CUSP_2015/data/**

**please do not upload copies of the data on github. see my last announcement for data uploading policies. which in summary says: do not upload your data unless it is a unique dataset, and valid options are uploading it online in the ipython notebook (you would not be penalized if the link were to die after your submission), placing it in a directory referred to by the environmental variable $PUI2015, or placing it in your local directory (but do not track it on git!)**

**Delivery:** check the ipython notebooks in a new folder HW5 in your PUI2015_<yourname> repo. Once you create the folder on your station (inside of the folder tracked by git) and have and have some files in it add it to the github repo by
$git add HW5
$git commit HW5/* -m 'HW5 first submission
$git push

**(here is a link to APOD: the astronomy picture of the day, to cheer up your evening. http://apod.nasa.gov/)**