* Prediction of Residential Energy Consumption in US
  + Goal: Inform policy makers where and what type of housing units in the US would most impact energy efficiency based on energy usage trends. Output would include recommendations of energy efficiency measures and subsidies based on the region and income level of resident.
  + Data ( Residential Energy Consumption Survey (RECS) ) http://www.eia.gov/consumption/residential/data/2009/index.cfm?view=microdata
    - Note: The link is only for 2009 housing stock data, however the EIA has viable data for 2009, 2005, 2001, 1997, 1993, 1990, and 1987
  + Questions:
    - The RECS measures A LOT of parameters for each individual housing unit surveyed (the 2009 survey reported about 900 characteristics for each of the roughly 12,000 housing units). While I can guess which characteristics have the most impact on energy efficiency, is there a statistical test I can implement to analytically determine which characteristics I should take into account from the beginning? (How a characteristic impacts kBTU/sqft)
    - Quite a bit of cleaning is required for the microdata. I’ve been trying to figure out a better way to read an excel file into Jupyter and haven’t made much progress after a lot of digging around. I’ll show you the issue I’m having by just reading it into Jupyter as a .CSV file. (you can probably see what my problem is from looking at my Github. Link below)
    - Though there are RECS for other years, can this analysis be done with just the variation in the 2009 data?
* Another Kaggle thing for a clearer-cut option: Global Energy Forecasting Competition 2012 - Load Forecasting
  + https://www.kaggle.com/c/global-energy-forecasting-competition-2012-load-forecasting
  + Data sets are reasonably-sized
  + Appeals to personal interest
  + Largest file is easily read into Jupyter even on my little underpowered computer.
  + Questions

I loaded up some of the data for both ideas to take a look at:

https://github.com/mohi7621/Springboard/blob/master/capstone/Initial-Examination.ipynb