**USDA Commodity Calculator**

**project overview**

I work for a team of commodity buyers. They routinely negotiate pricing contracts for produce; they would like to have historic data to look back on when going into the negations. The USDA collects and publishes data daily on several hundred produce commodities moving throughout the US but true to form, the government website where it is stored is difficult to navigate and confusing.

They would like to be able to produce quick analysis of a specific commodity over the last year by state and a prediction if the price on that commodity is going to rise or fall over the next year.

**data**

* Here broken down by State, Commodity, and Timeline as well as commodity attributes (organic, field grown, hydroponic etc)
  + <https://www.marketnews.usda.gov/mnp/fv-report-config-step1?type=shipPrice>
* Challenges: it is a huge amount of data and an incredibly large number of commodities, most of which we aren’t interested in
  + My initial thought was to make it interactive (buyers could select their own variables and produce a report) but that is running into some complexity issues
  + In order to address that I have scaled back, produced one year long report on tomatoes for California. I will do the data analysis and predict pricing for that one commodity and then move on and see how different other commodities would be (will one model fit all?)
  + Phase 2 will be to go back and see if I can modify it to pull data automatically on other commodities and see if the predictive model holds
* Positives: Data is available in excel downloadable reports making it very easy to feed into Pandas
  + Once I figure out the syntax to make it variable (allow user selection of commodities) will be able to build a pretty cool calc

**where i’m at today**

* Data (for tomatoes) is pulled and I have started exploring it (see usda calc v3 on github)
* Plans for the next week will be to start modeling and test out models
  + Initial gut check says Linear Regression but I’m interested to try a few models and see which one gets me closer