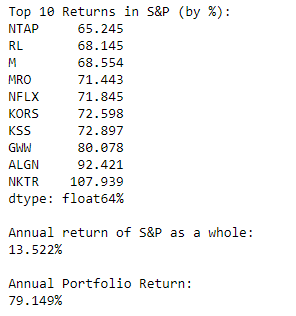
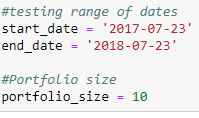
Ari Bailey

Momentum Regression Analysis:

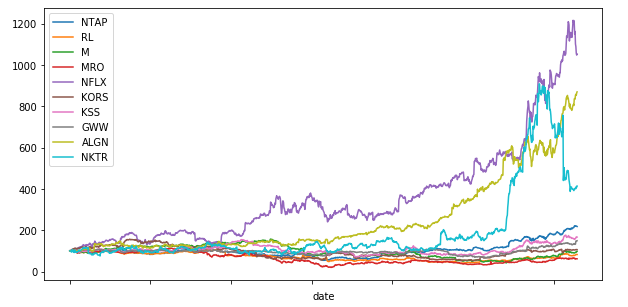
A python project that takes advantage of **IEX**, **Alpha** **Vantage**, and **Bloomberg** for data sources on stocks in the S&P 500 and **scipy** and **statsmodels** to preform regressions on this data.

First, variables for the range of dates to look at, and the “portfolio size” (top x stocks for momentum – measured by annual return – in the S&P) are instantiated. Historical closing data is collected from IEX for each of the stocks in the S&P, and data for the entire index is collected from Alpha Vantage. The annual returns of the S&P over the specified period and for each of the stocks, and the top x stocks with the highest returns are displayed. Stocks are weighted by last closing price, and the variance and volatility of the “portfolio” is calculated and displayed:

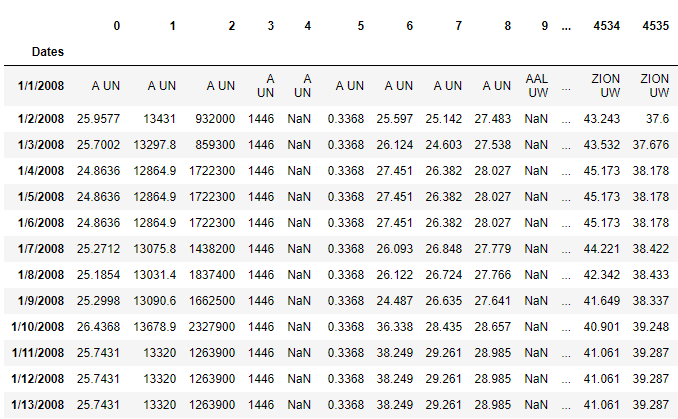








Now, daily data from the Bloomberg excel plugin is imported for each of the stocks in the S&P 500, going back to January of 2008, including: **Closing Prices, Current Market Capitalization, Price Volume, Revenue, Estimated Price Per Share GAAP, Net Debt to EBITDA**, and **10 30 and 90-day Volatility.**



Then, regressions are run. Over the indicated time period, regressions are preformed, per stock in the S&P, with each variable compared against the closing price, and a multivariable regression is preformed (excluding current market cap, which is directly related to stock price and would cause a bias, and 10 and 90-day volatility, because the volatilities may cause a false positive in correlation with each other). The r-squared and p values of these regressions are averaged, leaving 10 r-squared and p values. Also, these regressions are again preformed, but this time only on the top momentum stocks found earlier. The results are printed:

|  |  |
| --- | --- |
| Total S&P 500 Data  Market Cap R Squared Value 0.94133  Market Cap P Value 0.0  Volume R Squared Value 0.21894  Volume P Value 0.09426  Revenue R Squared Value 0.14649  Revenue P Value 0.07873  Estimated Earnings Per Sale GAAP R Squared Value 0.14291  Estimated Earnings Per Sale GAAP P Value 0.06742  Net Debt to EBIDTA R Squared Value 0.08226  Net Debt to EBIDTA P Value 0.06473  Volatility 10 Days R Squared Value 0.05229  Volatility 10 Days P Value 0.075  Volatility 30 Days R Squared Value 0.03488  Volatility 30 Days P Value 0.07312  Volatility 90 Days R Squared Value 0.02499  Volatility 90 Days P Value 0.06854  Profit Margins R Squared Value 0.01835  Profit Margins P Value 0.07021  Multivariable R Squared Value 0.53552 | Top 10 Momentum Stocks in S&P Data  Market Cap R Squared Value 0.99917  Market Cap P Value 0.0  Volume R Squared Value 0.26235  Volume P Value 0.05346  Revenue R Squared Value 0.28098  Revenue P Value 0.04197  Estimated Earnings Per Sale GAAP R Squared Value 0.33945  Estimated Earnings Per Sale GAAP P Value 0.0382  Net Debt to EBIDTA R Squared Value 0.19686  Net Debt to EBIDTA P Value 0.05061  Volatility 10 Days R Squared Value 0.14852  Volatility 10 Days P Value 0.06556  Volatility 30 Days R Squared Value 0.12242  Volatility 30 Days P Value 0.09294  Volatility 90 Days R Squared Value 0.11664  Volatility 90 Days P Value 0.08132  Profit Margins R Squared Value 0.07741  Profit Margins P Value 0.08748  Multivariable R Squared Value 0.65255 |

Unsurprisingly, Market Capitalization has a very high correlation to stock price. Save the multivariable regression, the rest of these values aren’t very significant. The only values that are near a r-squared value of **.3** – and only for the momentum stocks - are **volume**, **revenue**, and **estimated earnings per sale GAAP**. While this is certainly interesting, keep in mind that these correlations aren’t very strong still, and are likely influenced by the fact that “momentum stocks” often must be doing well financially to gain momentum.

However, the multivariable regression r-squared values are **0**.**535** and **0.652**! This may seem great, except that for this result to be useful to anyone trying to predict the momentum of stocks, they would need to somehow predict future values for every variable going into the multivariate regression before they could get an estimate of future stock price.

Closing Thoughts

I went into this project knowing that the odds of finding any useful results from my code were very low, especially with the two weeks’ time I had to complete it, and my lack of prior knowledge in Python, finance, and statistics. I am very happy, and surprised by, the fact that I got results at all. This was a great experience for me: learning not just python, but python for finance, and how to gather financial data from free API’s and the Bloomberg (using excel), the fundamentals of momentum stocks, and some basic statistical analysis. Even if I had succeeded at finding some correlation, which seems unlikely given that people with master’s in mathematics spend months to create (albeit more complex and useful) models, the process in getting to my results would still be more valuable to me.

However, if I was to go back with my current hindsight and take another attempt at my project, I would do some things differently. To list a few: First, I would skip trying to get free data from API’s entirely and go straight to the Bloomberg with the excel plugin (which is a wealth of any financial data you could ever want). Second, I would be a lot more purposeful when picking independent variables for my regressions. I would try to figure out, by running some test regressions (like this one), or by doing some research, which variables would be most likely to yield a useful or novel correlation, and I would use those. Third, I would try to get a better understanding of what momentum is, by trying to find a better measure of momentum than rate of return, which may have led to better regression results or decisions on input data.