ISyE 2028 – Basic Statistical Methods - Fall 2015
Bonus Project: "Big" Data Analytics
Proposal (or Final Report)

Kristiaan_Sheedy

For the final project, I would like to test the effect of news articles on the price movements of a stock to determine if news (aside from earnings reports) has a statistically significant effect. More specifically, I would like to test both the effect of the number of news articles about a particular company (both positive and negative) on the company's stock price, and the effect of the proportion of positive articles on a companies stock price. The hypothesis for these cases are as follows: As the number of articles about a company increases, it's stock price will not demonstrate a statistically significant diversion from its current momentum and the movement of the S&P; As the proportion of positive articles about a company increases, it's stock price will demonstrate statistically significant positive movement relative to the S&P and its previous momentum.

In order to do this, I will use a sample set of 14 companies from the S&P 500. Elements of this set will consist of 7 companies from the Nasdaq and 7 from the NYSE in order to balance the companies by exchange. Companies with larger market caps will be given preference in this selection as they will by nature be reported on more frequently. Some considerations that will be included in normalizing the results:

Each stock will have a moving average to compare to when analyzing our results. Each stock has a beta compared to the overall market, has an implied growth rate included in the price of the stock, and a current linear momentum which should be used as the normal basis to compare the results to. This is to say that while news articles may be overwhelmingly positive towards a particular company, if their past stock performance is not improved by a statistically significant margin following the release of good press (relative to their previous momentum, implied growth rate, and beta – accounting for their volatility relative to S&P 500), we should not consider the effect of the press to be statistically significant.

Another consideration is that the potential correlation between positive and or negative press does not imply causation with regard to the companies stock movement. Our results will have to be functionally interpret ted as such.

The objective of this project will be to determine if the news at time t influences the movement of the price following time t. Given the time span of this project, the time period for the initial news accumulation interval to use as our initial data will be set at 1 week. Reputable publications such as Barron's, Bloomberg, the NY Times, Fortune, CNBC, Reuter's, and CNN will be considered for the data set. With these considerations in mind, 14 companies which fit the aforementioned parameters will be selected. From the initial articles published on the selected companies, two variables, the total number of news articles about a particular company, and the proportion of positive vs. negative articles, will be established. Each day, the price movements of these stocks will be recorded (using starting and ending price for each trading day + gains or losses in the gap) and the proportion and number of news articles will be updated. This data (price movements, news articles, and proportions) will be collected for 2 weeks following the initial collection period. The data will then be analyzed with the aforementioned considerations being accounted for.