**Web Data Analytics Final Project: Idea 2**

**Team 202: Shan Lin, Nick Molter, Surya Gundavarapu, Ta-Wei Yen**

**Executive Summary**

**Our Goal**

Our goal for this project was to determine what effect, if any, earning calls have on whether a mutual fund chooses to add or remove a company from its holdings. With this information, we can better predict what is likely to happen to a company’s stock price once it releases its earning call. We can also utilize this information to try and determine what strategies are being utilized by different mutual fund, and thus which strategies are most effective.

**Our Process**

To accomplish this goal, we have to first start with the mutual fund. Once we have selected the mutual fund, our Python code takes over and extracts the 13F filings from the SEC website. Next it compares the filings from the previous two quarters and determines which companies the fund chose to add, and which it chose to drop. With these list, we do two things: we collect the earning calls for the past four quarters of each company in the list; and we collect some additional financial data (revenue, net income, total assets, and quarterly stock prices). The next step is to conduct sentiment analysis on each of the earning calls we collected. For this, we chose to utilize Google’s natural language sentiment analysis tool. Next we combined the sentiment and magnitude scores with our data set containing the financial information. Finally, we performed a logistic regression to determine how well of a predictor each of our variables were.

**Our Results**

After performing our analysis, we were able to conclude that the sentiment and magnitude scores where not statistically significant in the analysis of our selected mutual fund. At this time we were unable to run sentiment analysis all of the mutual fund we intended to due to the limitations of Google’s tool. It is likely that these earning calls play a more substantial role in the decision making process with other mutual funds which deploy different strategies. In our analysis, the only variable that was determined to be statistically significant was the stock price, this is to be expected for obvious reasons. To extend our analysis further, we would like to try many things including: expanding our reach to international companies; performing sentiment analysis on many more mutual funds’ holdings; collecting more financial data; performing sentiment analysis on Twitter feed related to the company; experimenting with other sentiment analysis tools and customizing their dictionary weights among other ideas.