Background: In the age of social media, companies are increasingly using social media platforms, such as Twitter, in order to alert followers of what they have accomplished as well as what they aim to do in the future. The affect of the language they publish on social media provides a valuable source that can potentially indicate the overall financial health of the company; for example, a Tweet that reads “Excited to launch our new product!” contains positive sentiment, indicating financial well-being. Many sentiment analysis techniques exist that can be used to measure the sentiment of tweets. This sentiment information can then be correlated with macroeconomic indicators.

At the same time, a company’s financial health can be tracked using another valuable source of information in the form of company-level news text. Often this news text is longer and more thorough than the information published on social media, opening up the opportunity to not only measure sentiment but also forecast information like stock returns.
**Purpose:** I aim to develop a stronger understanding of modern natural language processing techniques and to track company economic indicators using information from company news and Tweets.

**Project:** I will analyze Tweets provided by the School of Management, extract sentiment from them, and correlate the sentiment with several macroeconomic indicators. Using this information as a guideline, I will analyze company-level news text in order to forecast stock returns. The sentiment of the Tweets and the content of the news text will provide an overall picture of the company’s financial health, and allow for specific predictions to be made about the company’s stock returns.

**Procedures:**

1. Using sentiment analysis, extract sentiment from the collection of Tweets provided by the School of Management.

2. Discuss and finalize macroeconomic indicators with the School of Management. Correlate the sentiment with those indicators.
3. Scrape company-level news text. Extract sentiment and analyze key content in order to make predictions for stock returns.

**Deliverables:**

- Project documentation
- Data sources (i.e. Tweets, news text)
- Code for performing analyses, extracting sentiment, and predicting stock returns
- Conclusions about financial health for some example companies