**Project Title**: Stock Analysis Utilizing Tweet Sentiment and Historical Correlation

**Team Members**: Andrew Kling, Joel Boyette, Charles Gaskins, TJ Zientek

**Project Description**: Use data from the SEC, twitter and the NYSE to track tweet sentiment and historical correlation regarding the top 50 companies in the NYSE by trade volume.

**Questions to Answer**:

1. Does news/financial commentators sentiment on twitter about certain companies correlate to overall company performance as measured through their stock price?
2. Does general user base sentiment on twitter about certain companies correlate to overall company performance as measured through their stock price? Is this better or worse than the “experts”?
3. How do the 50 largest companies stock prices move with the overall stock market (S&P 500/DOW/etc)?
4. Can we identify good predictors for stock price volume trade? Day of week? Quarterly earnings calls?
5. Does acquisitions (ie Microsoft purchase of Github) affect overall stock price? Is sentiment impacted by this acquisition? Days after IPO of new stock?

**Data Sources**:

* Historical stock data
* Twitter API using tweepy library
* IEX Stock API
* SEC Dataset -Edgar

**Tasks**:

* Joel and Charles will focus on questions 1 & 2
* Andrew will focus on question 3&4
* TJ will focus on question 5

BACKUP QUESTIONS TO ANSWER:

* Using historical Kaggle dataset:
  + What is the average amount of time for maximum profit?
    - Aka is time in the market more profitable than timing the market?
  + How many “false peaks” occurred during this maximum profit period?
    - False peaks would be where the stock price was going up and then started going down. Could count the number of new maxes before reaching the final one
    - Story: if large amount of peaks before final peak then this is more attributed to luck than anything
    - Correlate to SEC data?
  + Maximum change day over day and does this correlate with SEC data for that stock?