

A low-angle, upward-looking photograph of several modern skyscrapers with glass facades, reaching towards a cloudy sky. The perspective creates a sense of height and architectural grandeur. The text is overlaid on the central part of the image.

Analysis of Most Liquid Securities Using Sentiment Analysis on News Headlines

Vivian Liu

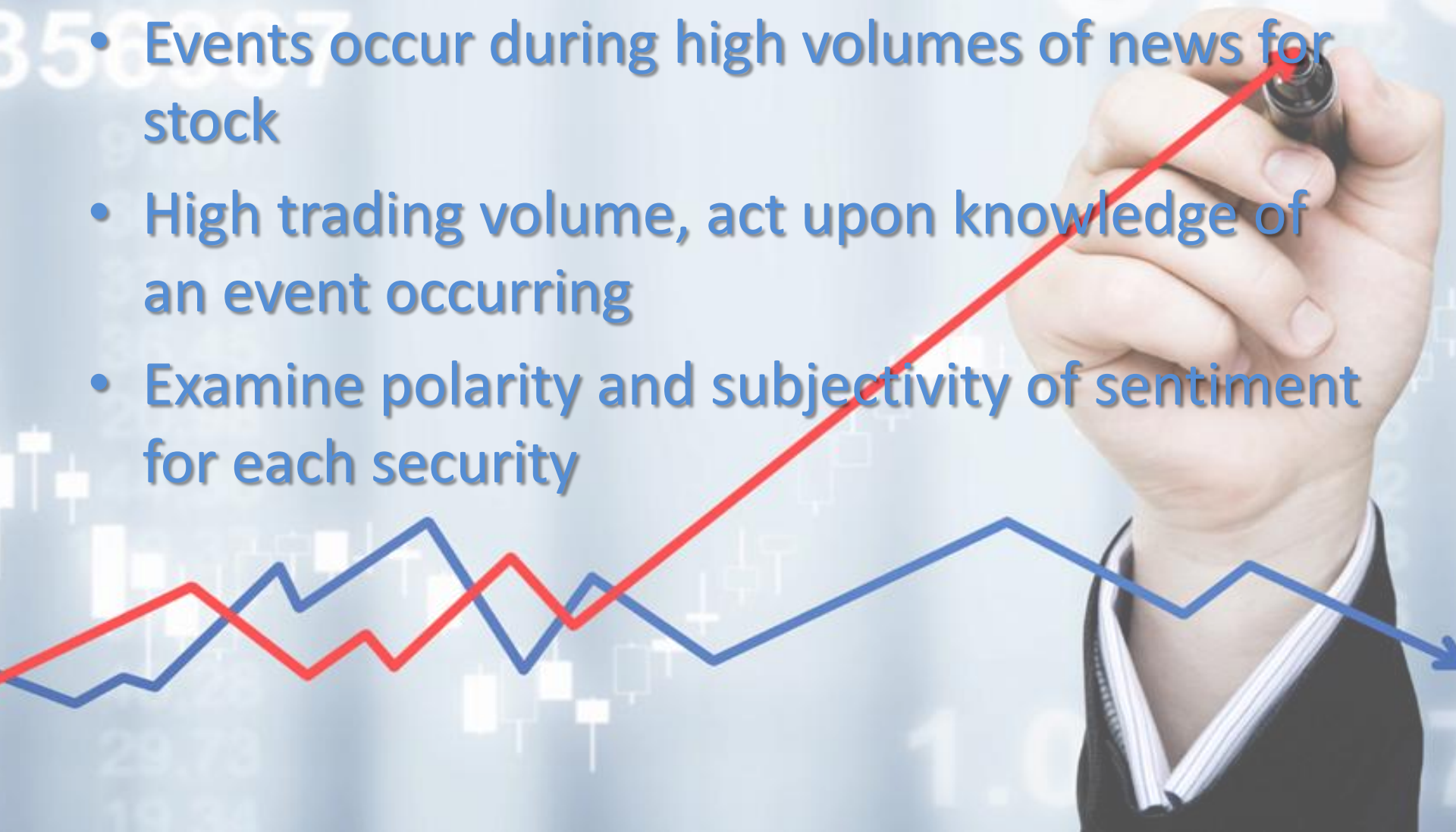
Why News?



- 70-80% of data in organizations is unstructured
- Accounts for kinds of risk not attributable to the market
- Large amounts of unstructured data untapped

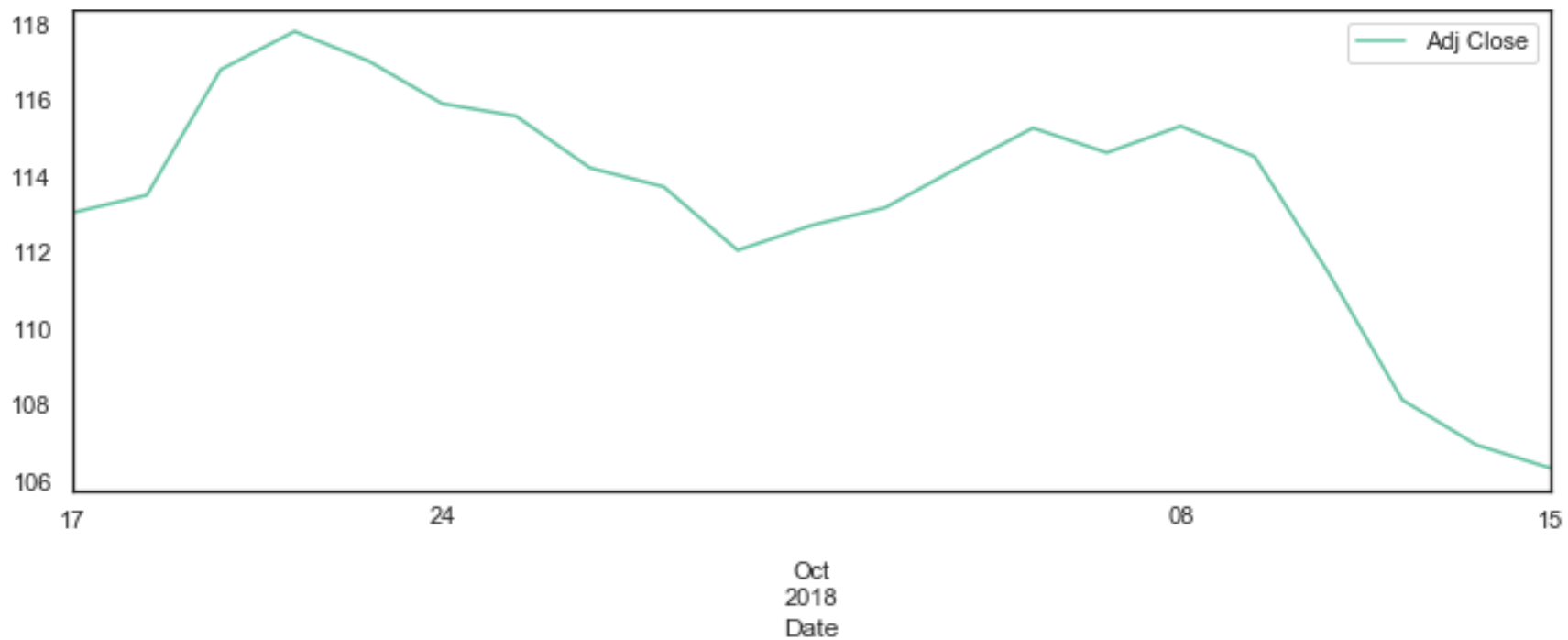
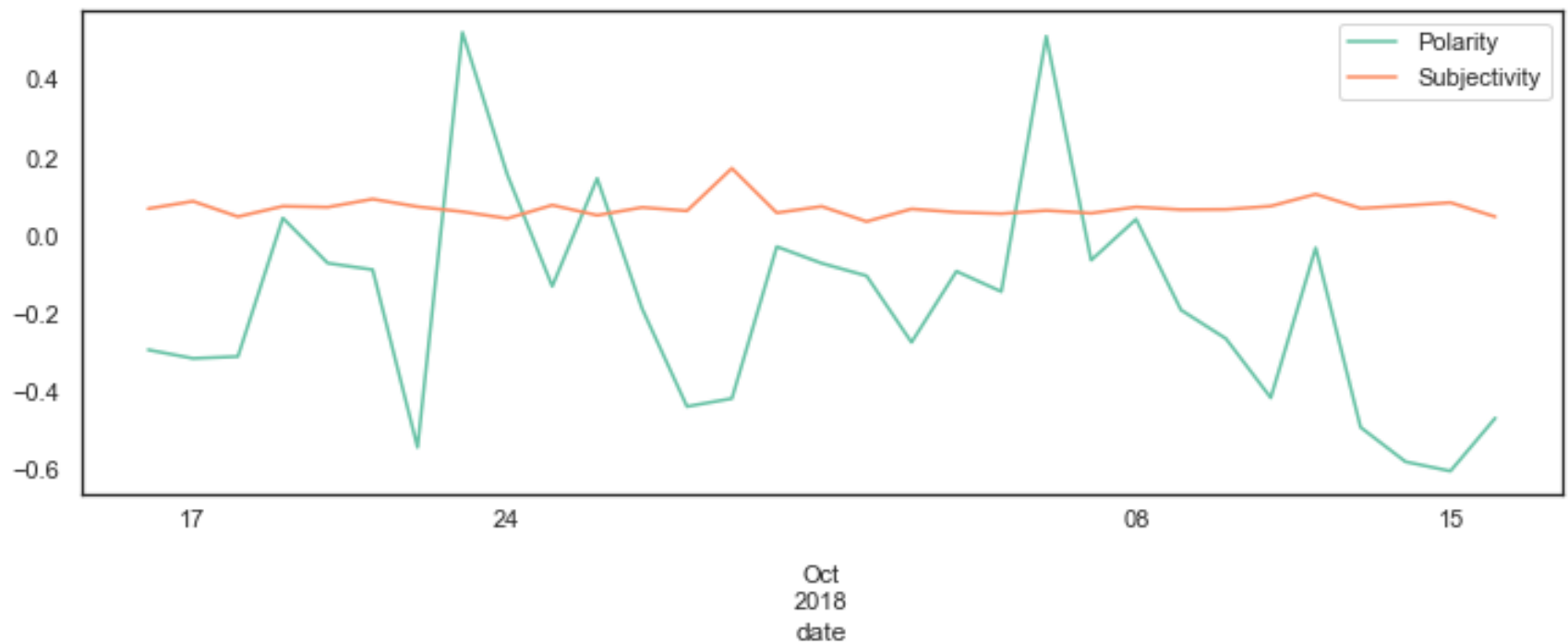
Volume as Indicator of an Event

- Events occur during high volumes of news for stock
- High trading volume, act upon knowledge of an event occurring
- Examine polarity and subjectivity of sentiment for each security

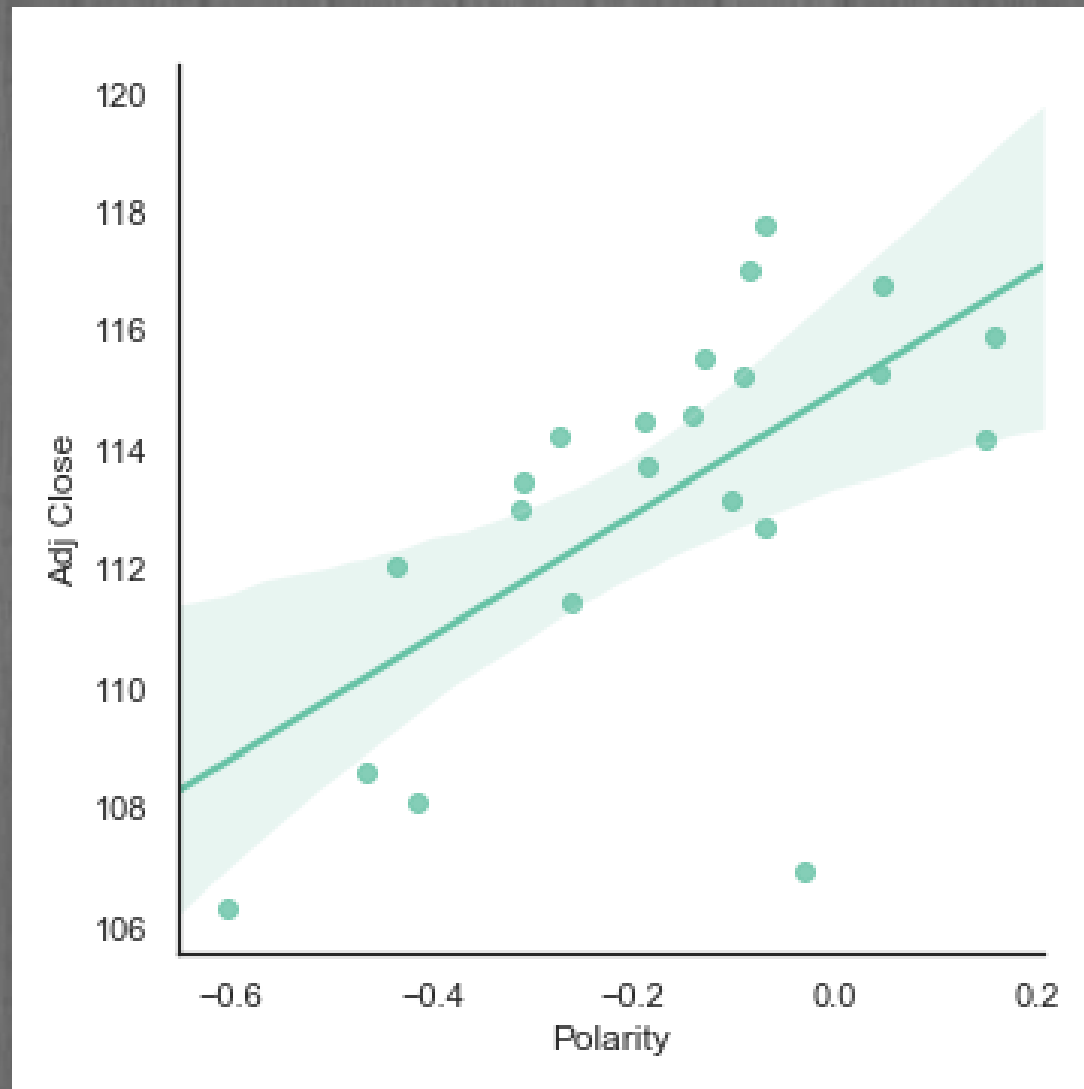


Obtaining the Data

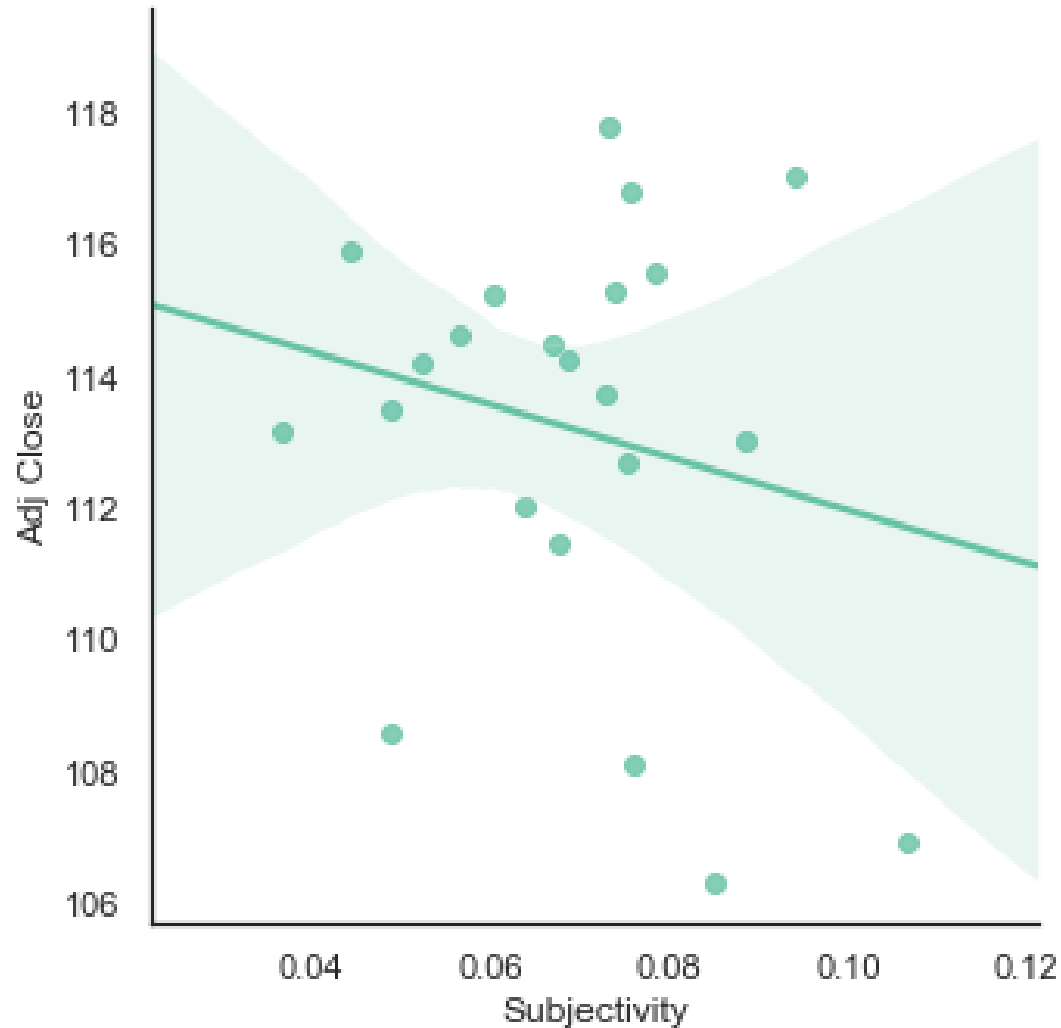
- Scraped most liquid stocks
- Use a news API for 30 days of news headlines and descriptions, newsapi.org
- Use Pysentiment package for its financial sentiment dictionaries



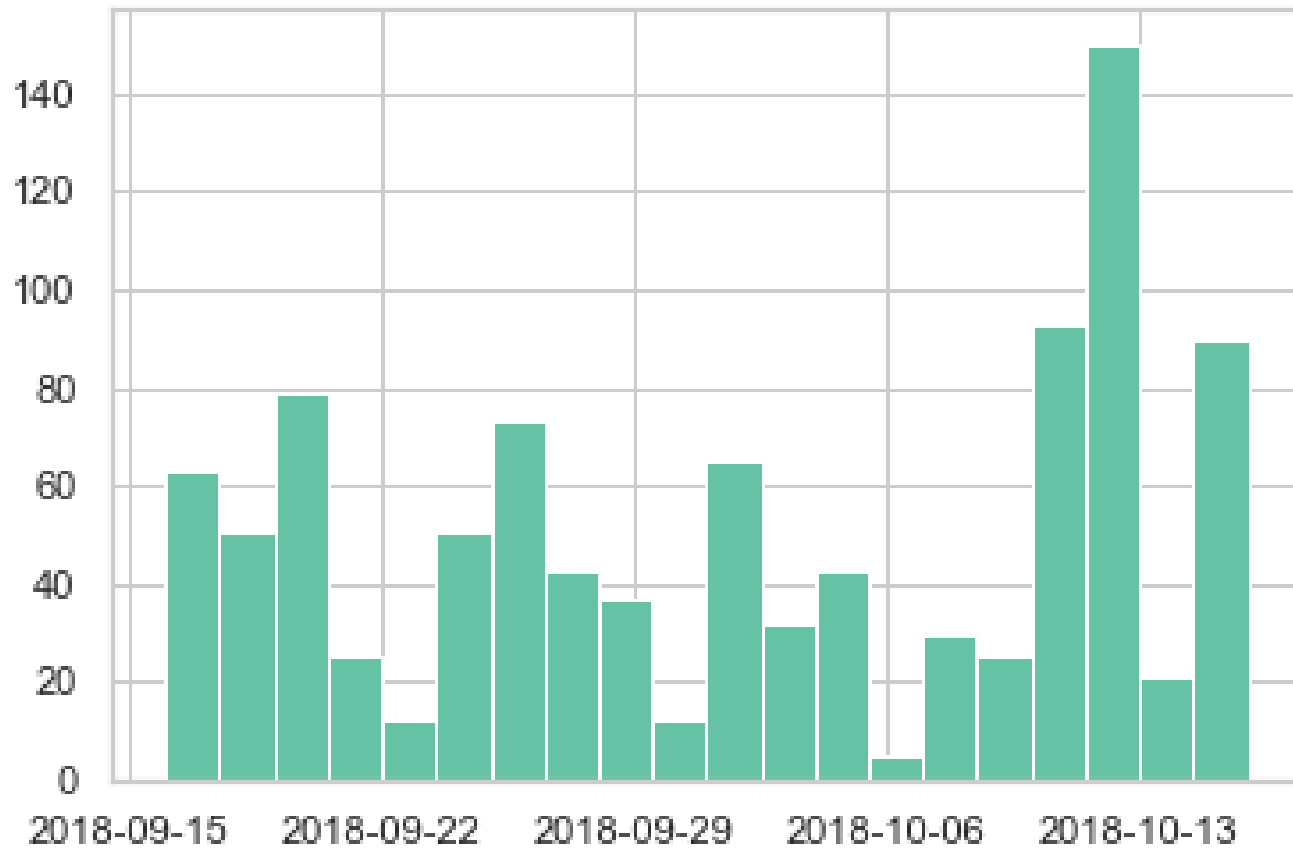
Closing Price v. Mean Daily Polarity



Closing Price v. Mean Daily Subjectivity



Histogram of JPM News



Conclusions

- Sentiment analysis on unstructured data can yield desirable results for day trading
- Larger volumes of news indicates an event which may lead to large rise or drop in price of asset

Improvements

- Barrier: requires much more data which will be expensive
- Generate alphas with SEC income statements, news, and other kinds of unstructured data
- Use deep learning