Here we conduct Granger-causality tests between the financial variables and sentiment data from Twitter. We normally standardized the time series first and then include a lag of 2 time intervals for the financial variable and 1 time interval lag for the sentiment variable in our regression equation.

Considering we can only get past 7 days' twitter data, the amount of our data is even smaller than the number of independent variables, which will return Null t - value when we conduct regression. So here we choose to use 4 hours as time interval to integrate data and then we can have enough data to regress. Also, because CSCO stock has very frequent tweets, we also conduct statistical analysis based on one hour interval.

After that, we consider the following regression models for the excess of return, which equals the log return of stock minus the log return of SPY etf. We don't use index here because lack of data in WRDS.

$$ER(t) = \alpha + \beta_1 ER(t-1) + \beta_2 ER(t-2) + \gamma_1 G(t-1) + \gamma_2 B(t-1) + \gamma_3 SR(t-1) + \varepsilon_t$$

Similarly, we have regression models for volume and volatility:

$$VOL(t) = \alpha + \beta_1 VOL(t-1) + \beta_2 VOL(t-2) + \gamma_1 G(t-1) + \gamma_2 B(t-1) + \varepsilon_t$$

Finally, according to the regression results, we investigate the statistical significance of the sentiment variables in regards to movement in excess returns, volume, and volatility and we can find the results differ a lot depending on the stock we choose.

The detailed results are as below:

The stock list where positive tweets have a significant positive influence (p<0.1) on stock excess of log return: WM

The stock list where relative sentiment has a significant positive influence (p<0.1) on stock excess of log return: ABT, FLT
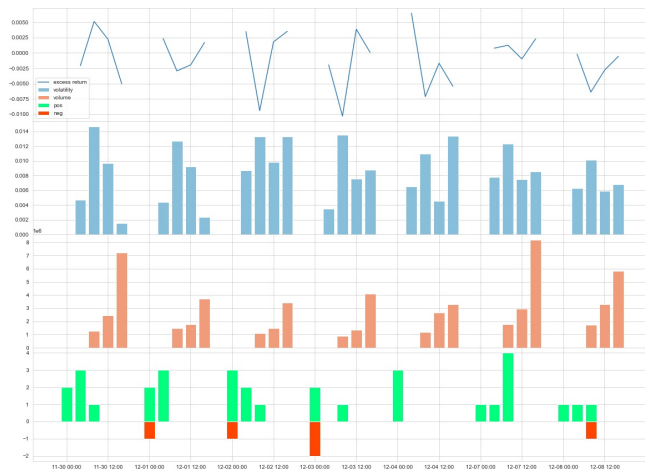
The stock list where positive tweets have a significant positive influence on stock trading volume: None

The stock list where negative tweets have a significant negative influence on stock trading volume: None
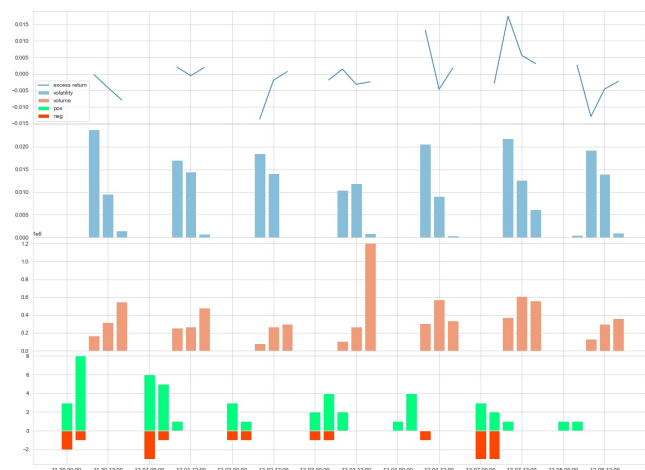
The stock list where positive tweets have a significant positive influence (p<0.1) on stock price volatility: HCA

Therefore, we can see from the statistical results that only 3 companies have a significant positive relationship between excess of log return and twitter sentiment. The trading volume does not seem to be influenced by the sentiment and only HCA company shows a positive influence of sentiment on the stock price volatility. The graphs shown below may help us better understand the relationship between financial variables and sentiment. Also, considering the limit of amount of data, the results maybe biased. So we can find other way to get more twitter data in the future to see if our conclusion changes in the future.
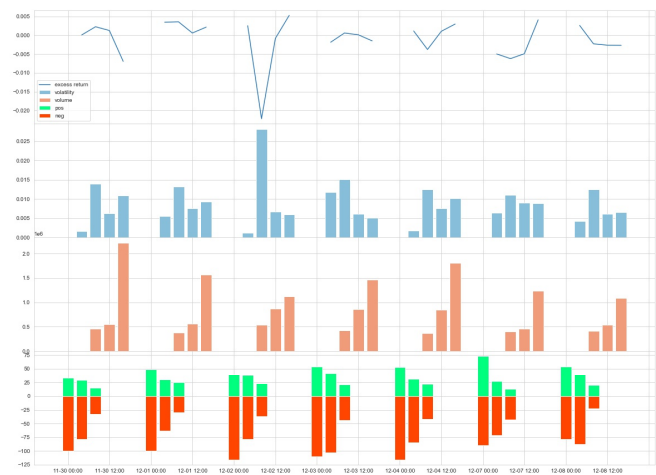
ABT:

FLT:

WM:

HCA: