CS591 Network and Markets **Final Project Report**short line

**Social Media Impact on Stock Market & Price**

Scarleth Estevez

Nathan Gallaway

Sang-Joon Lee

Dec 15 2016

# **Social Media’s Impact on Stock Market and Price**

A study of stock market behavior on news articles sentiments

Sang-Joon Lee, Nathan Gallaway, Scarleth Estevez

Department of Computer Science

Boston University

Boston, MA

email: {sangjlee, nwg, scarlet} @bu.edu

## **Abstract**

In a world dominated by the internet and social media, we wanted to determine what effect social media played in the stock market. We are researching on effect of news article in social media of companies listed in stock market and whether or not the sentiment of articles and spread throughout the network is correlated with the performance of stock market prices. In our study, our hypothesis is that there will be a positive correlation between positive news articles and the stock market price and a negative correlation between the stock market prices and the number of negative news articles.

## 

## **Keywords**

Social Media Analytics, Stock Market Prediction, Sentiment Analysis,

## 

## **INTRODUCTION**

#Todo

## **RELATED WORK**

#Todo

## **DATASETS**

# list in table list of stocks that we analyzed here

## **WEBHOSE.IO**

We are getting our data from webhose.io which is an api we can use that data scraps news articles and returns it to us in json format. We are going to use this for looking at the articles as well as looking at their social links and seeing their spread. Twitter and Facebook data for shares of articles.

#put example JSON format here

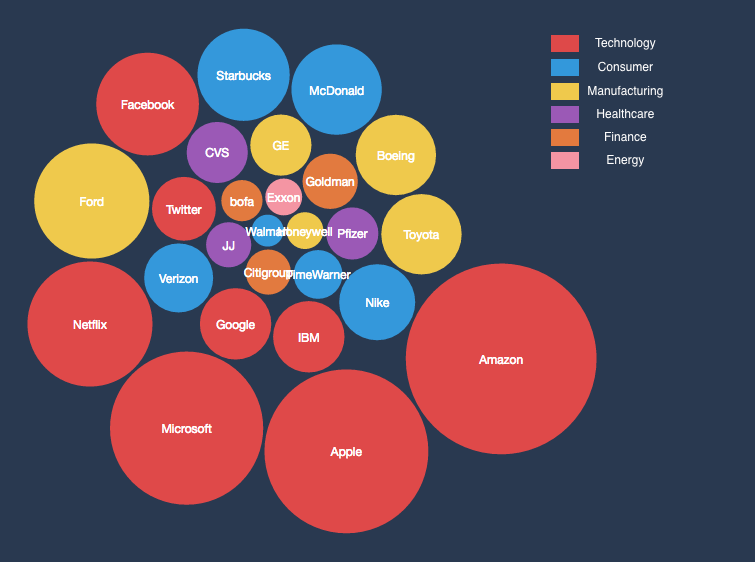
#mention over 60,000 news articles scrapped from 26 companies listed over 30 days

## **STOCK MARKET PRICES**

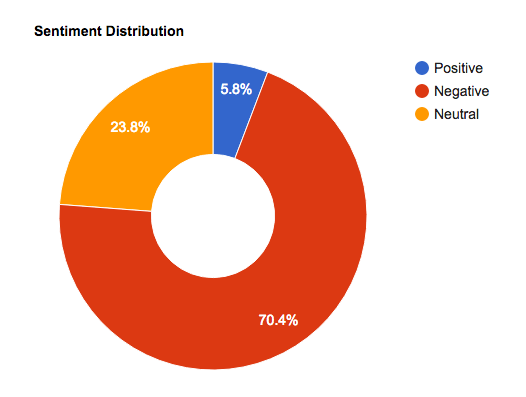
# say something about the source of the stock prices (yahoo finance) and duration of the data

# say something about the interpolation of data for stock price during the weekend.

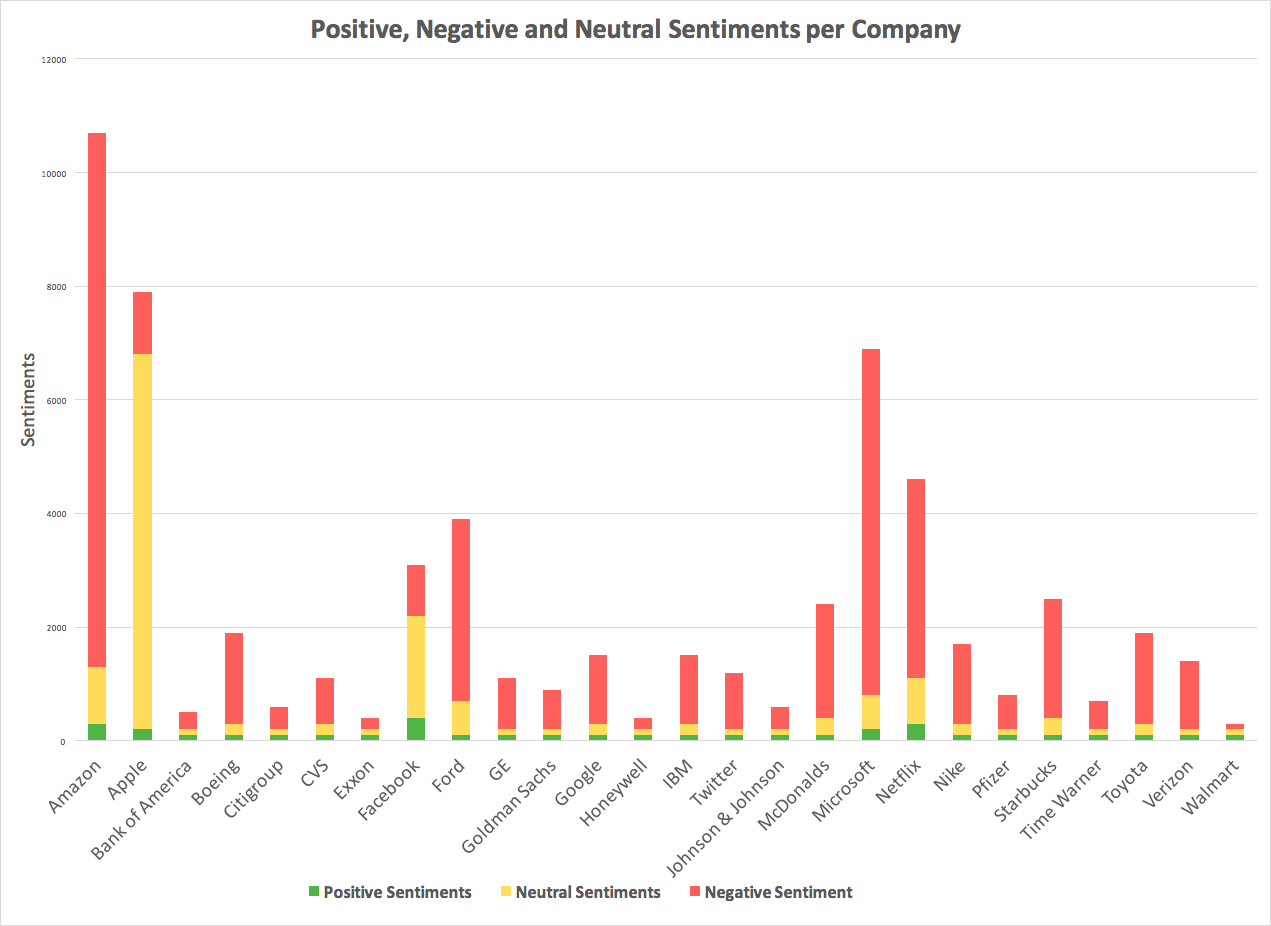
## **DATASET CHARACTERISTICS**



#todo: explain briefly what above graph shows – extrapolate the visual data into words such as volume and sector



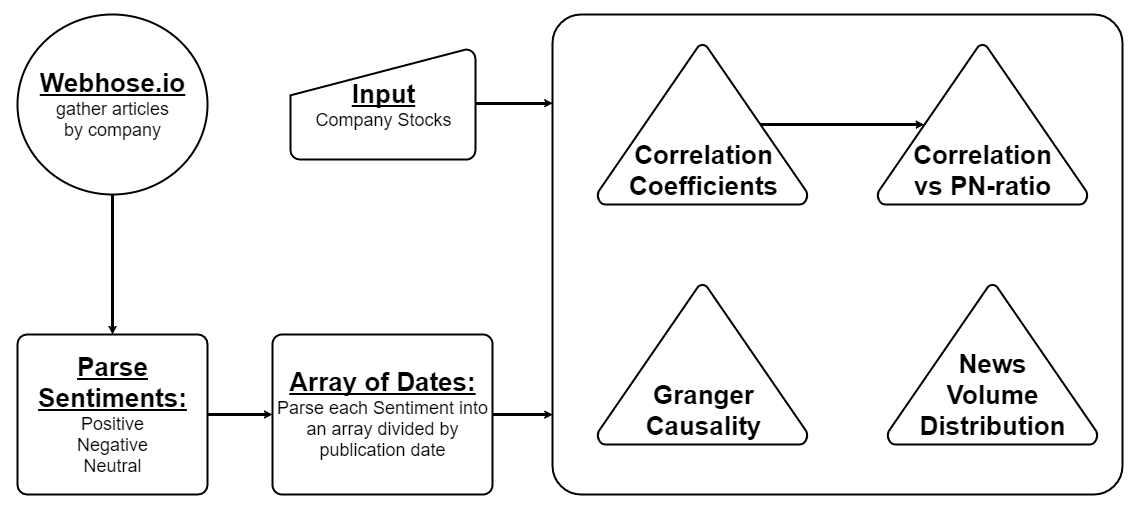
#todo: explain there are more negative news over last 30 days than positive news



# explain that this is breakdown of new volume over 26 different companies over different sectors

## **METHODOLOGY**

#brief description about what type of analysis was used and details in the section below



## **Correlation Analysis**

Identify how closely linked the following datasets are with each other

1. Positive articles & Negative articles
2. Positive articles and Stock Prices
3. Negative articles and Stock Prices

## **Granger Causality**

We analyzed two events over a period of time and determined if the first event can predict what the second event will do in the immediate future. For example, if there is a large increase in positive news articles with stock market price rise the following day.

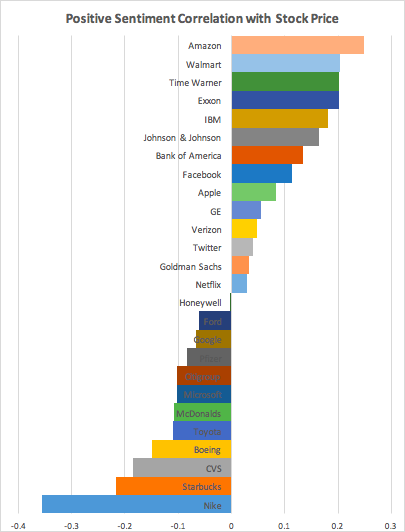
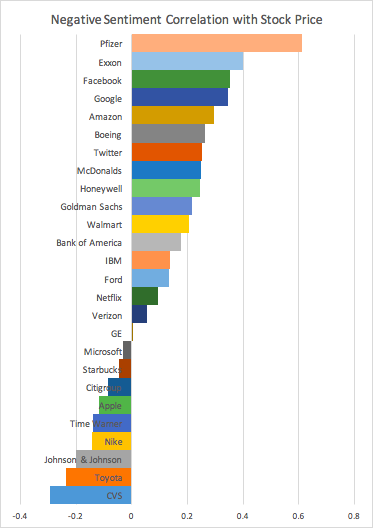
# more information about what the granger causality is and how to determine if two time series data have causal effect.

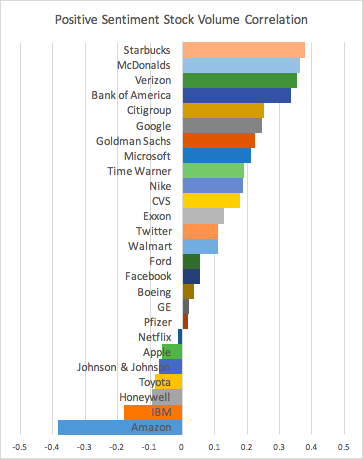
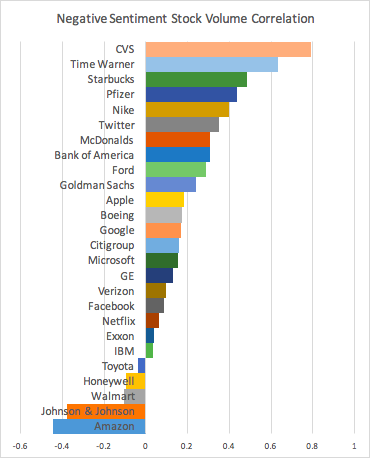
## **RESULTS**

#todo: need to explain the analysis result here

## **Correlation Results**

## **../../Downloads/volume%20scatter.png**

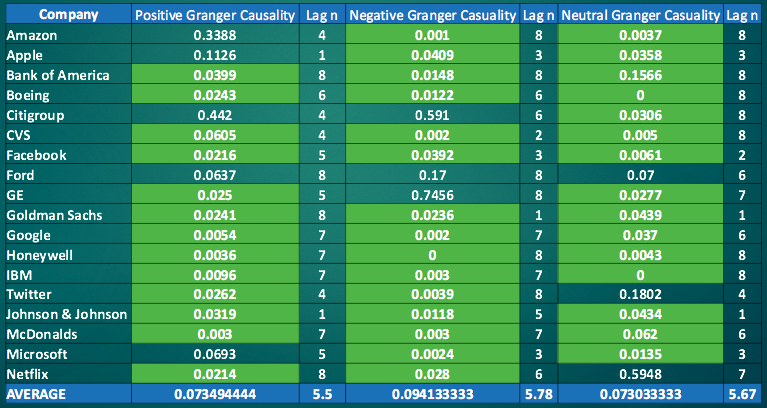
 

#todo: Add results for correlation over different sectors

#todo: Add results for correlation over different market cap

## **Granger Causality Results**

#todo: complete all granger causality results and explain them



## **CONCLUSION & FUTURE WORK**

#todo: some conclusion about how they are correlated or NOT

#todo: acknowledge that there are error in the findings due to nature of scrapped articles, number of articles, binary characteristics of sentiment output, duration of data – only 30 days

## 

# future work

Extending data for months rather than 30 days

Scrapping twitter data and comparing the lag of days in the granger causal effect

Analyzing twitter data for spread and diffusion of the sentiment of news against the behavior of stock market

## **ACKNOWLEDGEMENT**

# acknowledge Professor John Byers

## **APPENDIX**

# add link to github for source code and data

<https://github.com/cs591B1-Project/Social-Media-Impact-on-Stock-Market-and-Price>

## **REFERENCE**

[1] Bollen Johan, Huina Mao, and Xiaojun Zeng, “Twitter mood predicts the stock market”, Journal of Computational Science 2.1 (2011):1-8

[2] C Castillo, M Mendoza, Barbara Poblete (2011), Information credibility on Twitter, Proceedings of the 20th international conference on World Wide Web., page 675-684

[3] Asur Sitaram, and Bernardo A. Huberman. “Predicting the future with social media.” Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE WIC/ACM International Conference on. Vol. 1. IEEE, 2010.