



# Finding Correlations between Geographical Twitter Sentiment and Stock Prices



Juweek Adolphe\*, Ressi Miranda†, Zhaoyu Li, Dr. Yi Shang

\*University of Georgia, †Mount Holyoke College, University of Missouri

## Introduction

A company's future is almost impossible to predict, and the fluctuations of a stock's price are only testament to that. Ideally, one would like to be able to have some sort of indicator of how a company's stock is going to act. The goal of this project is to determine the validity of the claim that public sentiment found on microblogging sites, such as Twitter, can be used as indicators of a company's stock behavior. This project focuses on whether the Twitter sentiment of a geographic location has any correlation to the stock price of a particular domain.

## Previous Work

- Social networking sites have made it possible for millions of people to share their opinions
- Twitter has been used to obtain the public's mood for elections and other particular events
- Stock market prices are somewhat determined by public mood[1] by using Twitter
- It could be possible to predict stocks prices to a certain degree using sentiment analysis

## Objectives

- Our objective is to find out:
- If a certain geographical location's tweets correlates to a particular company's stock price?
  - If a certain geographical location's tweets has more weight on predicting a particular company's stock price?

## Methods

In order to compare Twitter Sentiment and stock price, we selected a domain, a geographical location, and a classification method.

## Domain Selection

- Selected a domain and picked a set of companies as representation
- Our company selection is from Standard & Poor's 500
- Our focus is on top restaurants & food companies
- We use a keyword search to find company related tweets

	Domain Type	Stock Ticket
Company 1	Restaurants & Food	\$MCD
Company 2	Restaurants & Food	\$SBUX

## Geographical Location

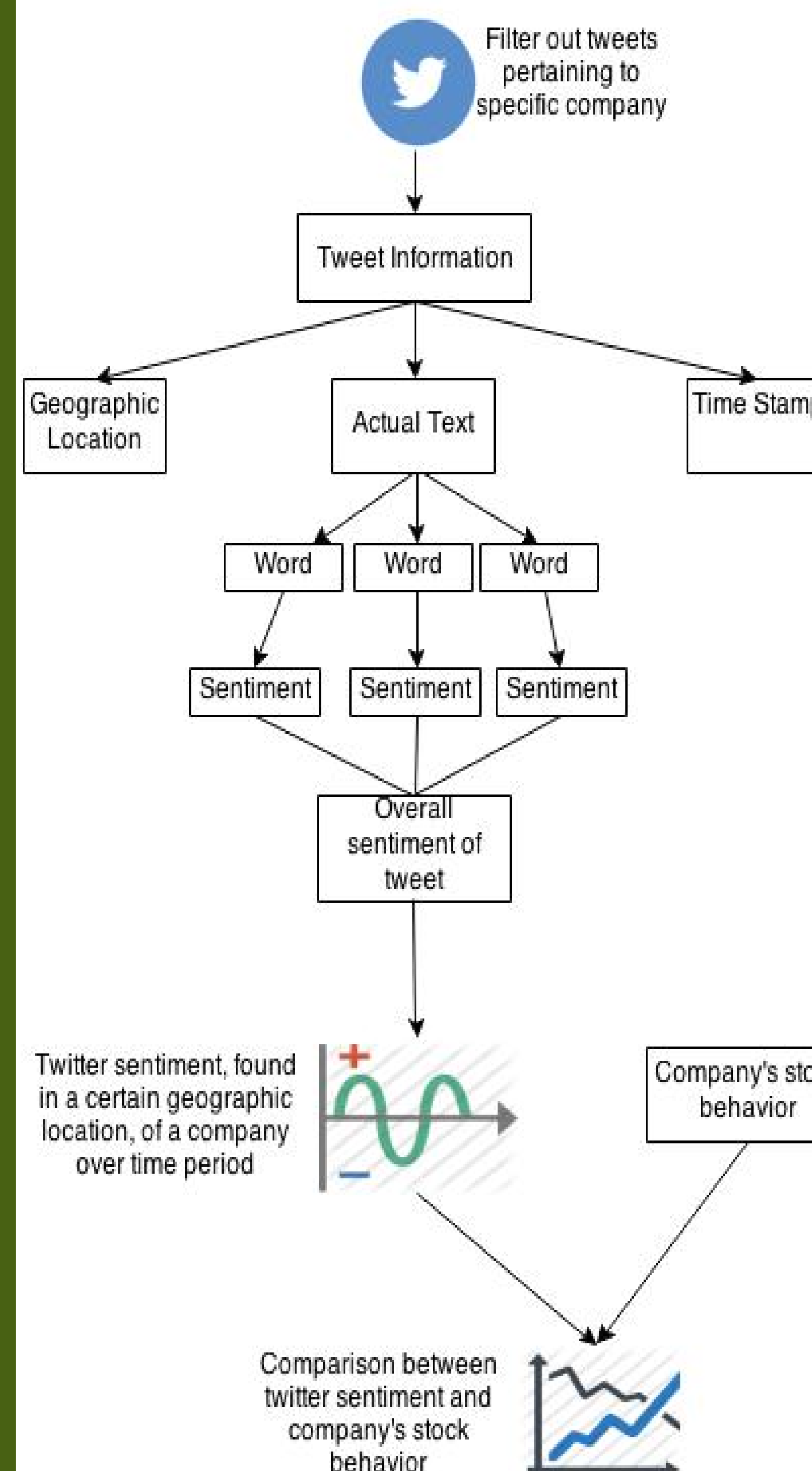
- We focus on the top cities that contribute to the GDP since previous work had shown that the distribution of Twitter users was concentrated in city areas.
- The Twitter population by state did not accurately portray a state's population[2].
- Cities chosen: NYC, Chicago, Huston, Washington DC, LA

## Sentiment Analysis

- Used sentiment analysis, a method to classify text into a polarity, as a tool to represent our tweets behavior
- Using lexicon based approach to classify tweets
- We compare tweet sentiment with stock price

**Example:** *The coffee is repulsive in taste and in flavor.*  
#starbucks → *negative*

## Procedure



1. Streaming tweets (Tweepy)
  - 1) Select Tweets by chosen company related keywords (Example: \$MCD, McDonald's, etc.)
  - 2) Sort tweets by location, date, and then time
2. Filter out the top 5 cities
3. Preprocessing text (NLTK)
  - 1) Remove stopwords (e.g. and, but, etc.)
  - 2) Remove non-alphanumeric characters
  - 3) Lowercase all characters
4. Sentiment classification - Lexicon based approach - each word is assigned polarity scores [3]
5. Sum of positive and negative scores
6. Correlate sentiment with stock price

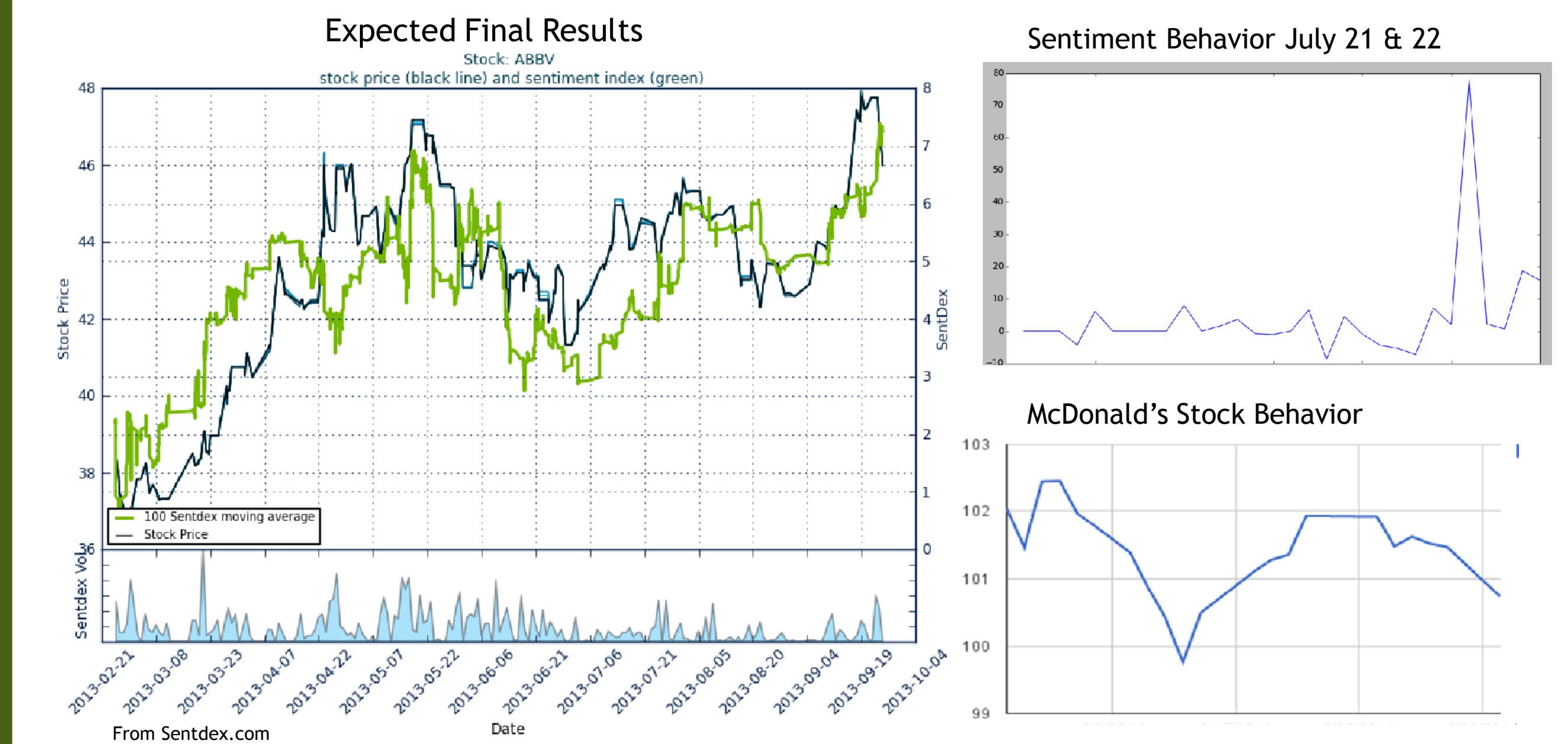
### Example of Processed Text:

**Original:** The chicken nugget is invisible to Jeff's eyes. It is sad what mcnuggets have become.

**Final:** chicken nugget *invisible* jeffs eyes *sad* what mcnuggets have become = positive = 0.125 + 0.75 = 0.875

*negative* = 0.5 + 0.75 = 1.25

## Current Results



- The right conveys our current results
- The left shows the expected results
- Currently, our work is still in progress and our results are inconclusive

## Future Work

Our results are inconclusive since we are collecting data. We hope to find stronger correlations with Twitter sentiment and a company's stock price, given a geographical location. We also hope to explore the following questions:

- How does twitter sentiment affect a domain's stock price?
- Does geographical twitter sentiment have any predictive power on a particular stock price?
- Is Twitter sentiment predicative of one domain's stock price over another domain's stock price?
- Is Twitter sentiment predicative of a particular company's stock price?

## Special Thanks

Special thanks to Professor Shang and Zhaoyu Li for their mentorship and for their help. We also would like to thank the National Science Foundation for funding our research experience and the University of Missouri for this great opportunity!

## References

- [1] Bollen, Johan, Huina Mao, and Xiaojun Zeng. "Twitter mood predicts the stock market." *Journal of Computational Science* 2.1 (2011): 1-8.
- [2] Mislove, Alan, et al. "Understanding the Demographics of Twitter Users." *ICWSM 11* (2011): 5th.
- [3] <http://sentiwordnet.isti.cnr.it/>