S&P 100 Stock Insight using StockTwits API

FINAL PROJECT DESCRIPTION

Yuxiang Zhang, Elizabeth A. Syso, Bonan Yuan, Tianqi Jiang
{yz3079, eas695, by424, tj763 }@nyu.edu

Project page (on Github):
https://github.com/NYU-CS6313-SPRING2016/Group-18-INET-3-StockTweets
Video:
Working demo:
https://stockdashboard.herokuapp.com/

What is the problem you want to solve and who has this problem?

The stock market is unpredictable and there are many differing opinions about which stocks will do well in the future. Meanwhile, visualizations of market *opinions* are very rare. But, traders need a quick way to access the opinions of their peers without having to search through websites like Stocktwits for it. Also, market newcomers who want to buy stock but are not sure which companies are popular or trending need the wisdom of current traders as well. Our solution is to build a dashboard to get an insight into stock market through Stocktwits in a simple way.

What are the driving analytical questions you want to be able to answer with your visualization?

- Which stocks are being talked about?
  Find popular stocks and the topics related to them.
- How does the market feel about a specific company?
  For a specific stock, explore the change in public opinion in real time based on the number of buy/sell opinions gathered from Stocktwits.
- What equity does the market suggest we buy/sell?
  Make the buy/sell consensus of each stock comparable with each other.

What does your data look like? Where does it come from? What real-world phenomena does it capture?

The data about stocktwits is from Stocktwits API. Realtime server takes 200 hit per hour and 30 records returned by each hit. Price and other basic information of each stock is from Yahoo Finance API.

Attribute Name	Attribute Type	Description	Range	Derived
text	Categorical	Test content of the stocktwit	String	N
hashtag	Categorical	Hashtag of the stocktwit	String	N
timestamp	Quantitative	Timestamp of the stocktwit	MM/DD/YY HH:MM:SS AM	N
username	Categorical	Unique name of the user	String	N
sector	Categorical	Sector of stock	String	N
price	Quantitative	Current price of the stock	Array	N
prev close	Quantitative	Yesterday's closing price of the stock	Array	N
open	Quantitative	Today's open price of the stock	Array	N
volume	Quantitative	Market volume of the stock	Array	N
P/E	Quantitative	P/E value of the stock	Array	N
EPS	Quantitative	EPS value of the stock	Array	N
Hotness Index	Quantitative	Hotness of the stock	Array	Y
B/S Index	Quantitative	Market Buy/Sell sentiment of the stock	Array	Y

What have others done to solve this or related problems?

Here already exist many projects that use Twitter sentiment analysis for stock information, but none so far that utilize Stocktwits. Their aims are similar to ours, but by focusing on Stocktwits as a source of information. We are using data that is published by other traders, rather than the general masses of Twitter users.

Initial Mockup

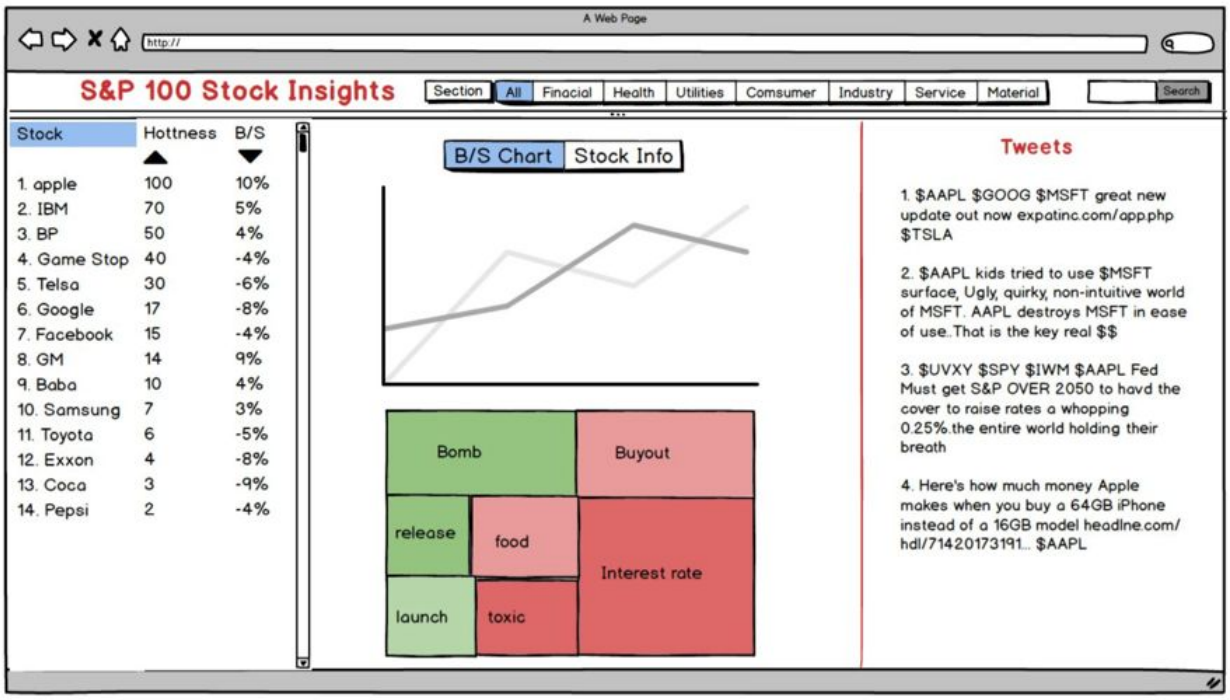


Figure 1: Overview

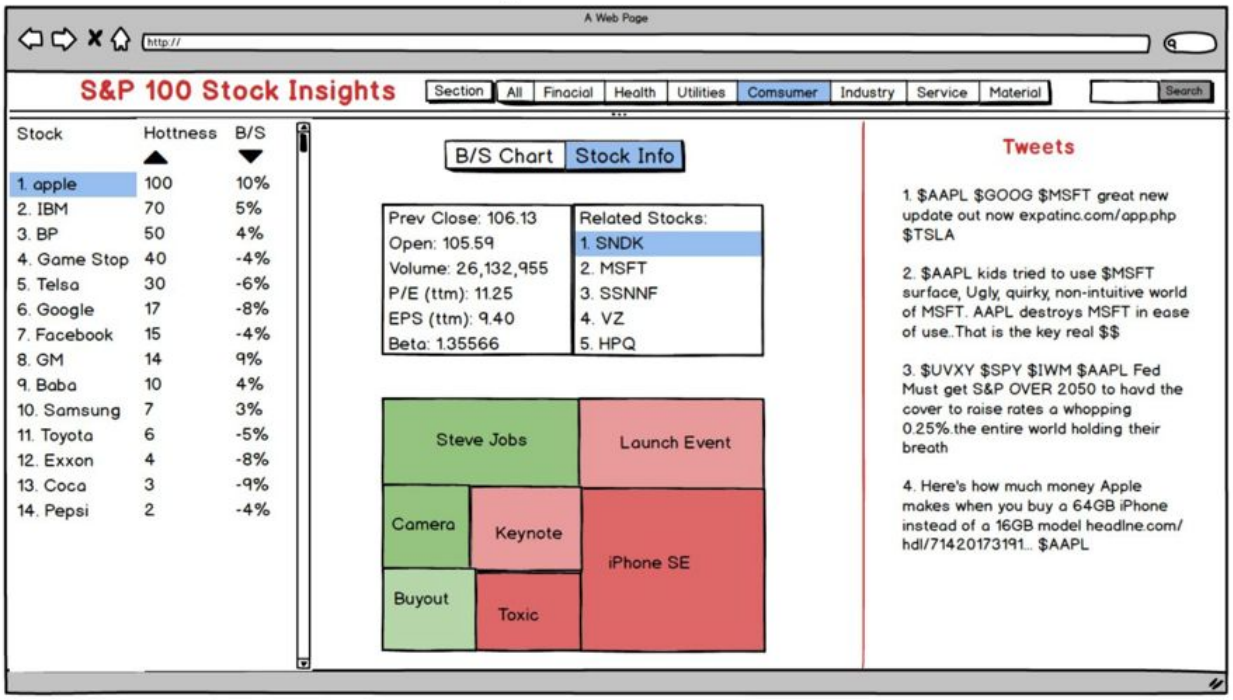


Figure 2: Insight of Apple

The title of the mockup is named “S&P 100 Stock Insights” which means this tool helps to visualize the 100 stocks in the S&P 100 Index. There are also stock search field and section selection field on the top of the page.

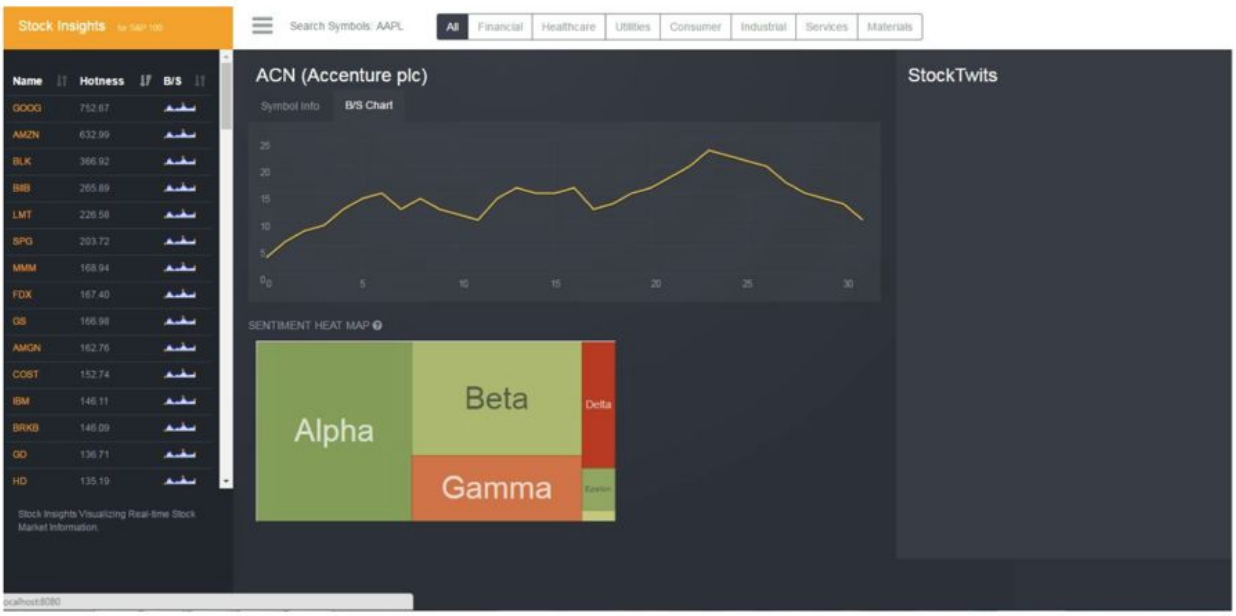
The left panel of the mock up is by default the list of the Top 10 stocks that have the highest rank in either “Hotness Index” index or “Buy/Sell” index. The “Hotness Index” is the count of times each stock is mentioned in StockTwits’ hashtag in real time. The “Buy/Sell Index” is a calculation of StockTwits’ community’s buy/sell opinion on that stock, which is derived through sentiment analysis on Twits message. User also can filter the list according to their interested sections, such as 'Financial', 'Healthcare', 'Utilities', 'Consumer Goods', 'Industrial Goods', 'Services', 'Basic Materials', or 'Technology '.

In the middle part, there are up and down two charts. The top one has two tabs, “Buy/Sell Chart “ and “Stock Information”. “Buy/Sell chart” is a line chart represent the trend of the whole market. And “Stock Information” introduce the Stock Market PE ratio, volume, etc. The bottom one is a heat map of popular keywords among the stock market.

And the right part of the web is a waterfall of StockTwits message. Onloading the main page will give user a waterfall of the most recent messages.

Once you click on certain stock in the left part(or search in the input field above), the upper chart in the middle part will become line charts representing the price trend and sentimental sell/buy trend of that stock. The “Stock Market Information” tab will become “Stock Information” tab, which contains additional information of related stocks and their price and Buy/Sell Index. And the bottom chart in the middle part will change to a heat map representing a sentimental analysis of what StockTwits’ users talked about this stock. If you click a word/topic in heat map of sentimental analysis in middle part, the right part of Twits message will show you the massages relative to that word.

Project Update



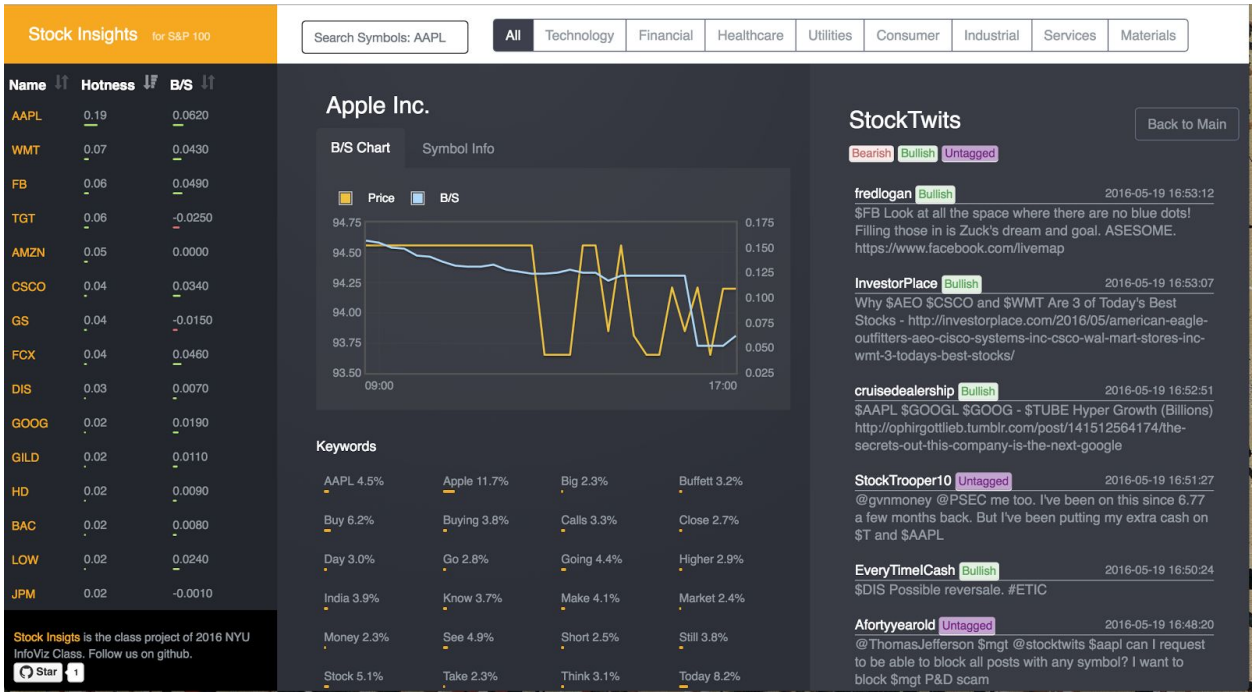
In updating our visualization plan, we added another dimension to our solution for question # 3 “ What equity does the market suggest we buy/sell?”. Previously, we had only planned to answer that question with a line chart on the page for that specific stock. In order to see the trendline for buy/sell sentiment for any given stock, the user would have to navigate to that stock’s page in our app. Currently, we have added small sparklines to the “hotness” list on the left hand side of our visualization, which allows the user to see the general trendline of a stock’s buy/sell sentiment at a glance, without navigating away from the overview page. The user can then navigate to the “stock details” for a larger, more detailed version of that buy/sell trend.

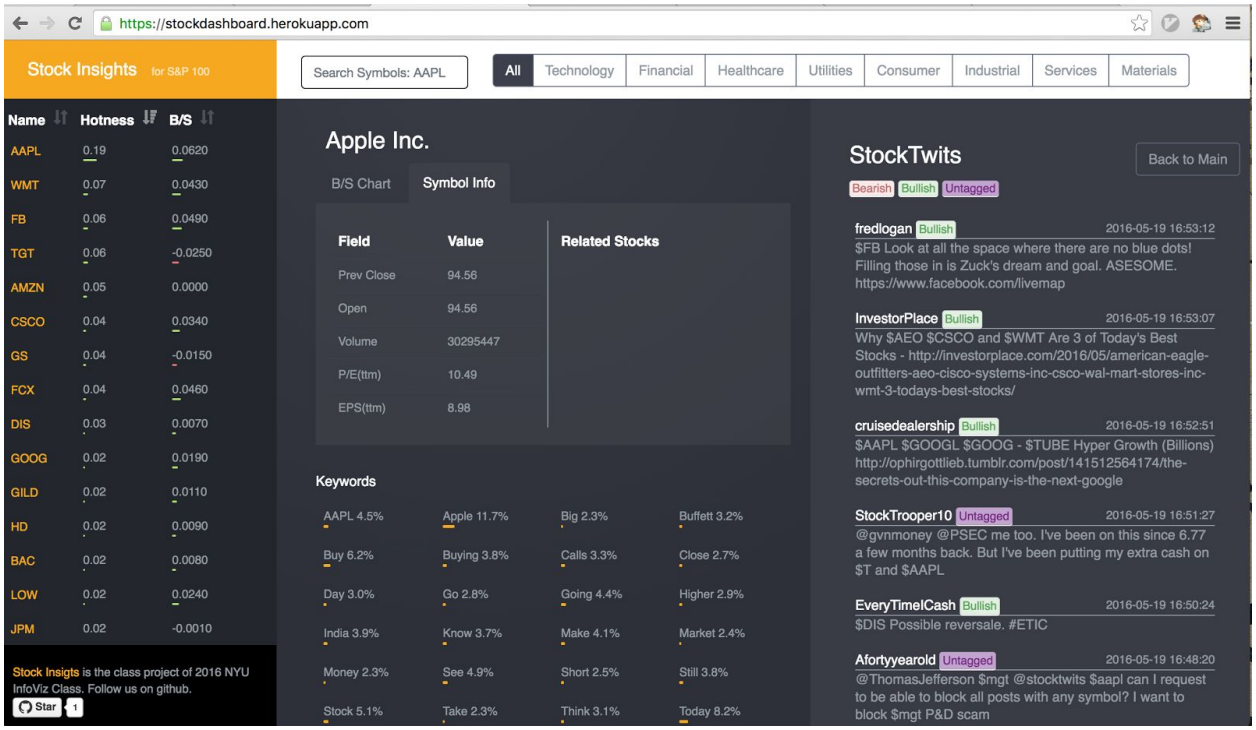
The left panel of the mock up is by default the list of the Top 10 stocks that have the highest rank in “Hotness Index”. The “Hotness Index” is the count of times each stock is mentioned in StockTwits’ hashtag in real time. The “Buy/Sell Index” shows a sparkline trend of the StockTwits’ community’ buy/sell opinion on that stock, which is derived from the buy/sell tags on the Stocktwits.

In the middle panel, there are two charts. The top one has two tabs, “Buy/Sell Chart” and “Symbol Info”. “Buy/Sell chart” is a line chart represent the trend of the specific stock/sector. And “Symbol Info” will include values for each stock like the current market price, stock volume, open price, closing price, etc. The bottom one is a tree map of popular keywords among the stock market.

The right panel will be a waterfall of Stocktwits messages. Loading the main page will give the user a scrolling waterfall of the most recent messages.

Final Visualization





In the final visualization, we almost achieve all the functions mentioned in the initial mockup.

On the head bar, next to the title “Stock Insights for S&P 100”, there is a search box which allows users to type in specific stock name. Beside the search box, there is a selection bar which allows users to select specific sector or whole market(All).

In the left panel, there is a table of stocks with their “Hotness Index” and “Buy/Sell index”. By default, the table shows stocks that have the highest rank in “Hotness Index”. Users can also rank the table by “Buy/Sell index” by clicking the sort button beside B/S title. In the table, below each number of “Hotness Index” and “Buy/Sell index”, there is a bar shows the percentile of the number, which makes the number could be compared easily.

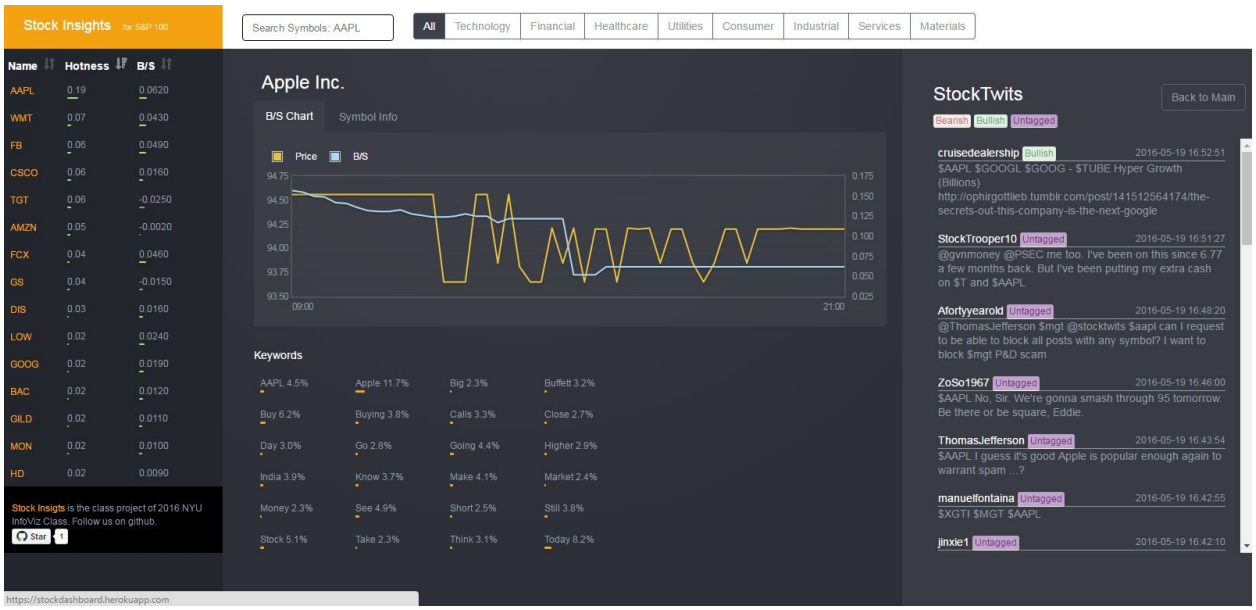
In the middle panel, there are two charts. The top one has two tabs, “Buy/Sell Chart” and “Symbol Info”. “Buy/Sell chart” is a line chart represent the consensus trend(based on B/S index) with price trend of the specific stock from 9:00 am to present time today. And “Symbol Info” will include values for each stock like the current market price, stock volume, open price, closing price, etc. The bottom one is a table which shows top 24 popular keywords among the tweets about specific stock. Numbers in this table are also with percentile bars making easily comparable.

In the right panel, there is a waterfall of Stocktwits messages. Loading the main page will give the user a scrolling waterfall of the most recent messages of whole market. Each message shows user name, timestamp, message content and a symbol showing author’s attitude of this message. The attitude could be “Bearish”, “Bullish” or “Untagged”.

For interaction, once you click on certain stock in the left part(or search in the input field above), the upper chart in the middle part will become line charts representing the price trend and sentimental sell/buy trend of the elected stock; the “Symbol Info” tab will also show the information about that stock; the bottom table in middle panel will shows top 24 popular keywords about the selected stock. (By default, the charts in the middle panel will show the information of APPL); the tweets messages in the right panel will also show the messages about the selected stock. Besides, once you move your mouse on the line chart, a tooltip showing time and value of that point will show up.

Data Analysis

Our visualization tool is meant to be accessible for anyone as a means to gain insight into market's current opinion. It is a stepping stone for someone who is interested in gaining an idea of where to start in their financial research. For a newcomer looking at our tool, the first screen they encounter is this:



This screen displays the full market hotness list, the 30 most recent twits, and the day's trendlines for Apple, one of the most recognizable companies in the world today, and also the first stock listed (alphabetically).

Say they wanted to know more about Apple's stocks,they would most likely be drawn to the center of the page, where they would see the trendlines and the keywords.

The trendlines show the interaction and changes in the stock price and our calculated B/S ratio throughout the business day (9 AM and on). *What is that huge drop in the B/S ratio?* They may think to themselves, before realizing that it probably occurred during the market close. Once the market has closed, traders share their trading tips at a slower rate.

Then, their eyes would move down to the keywords below the trendlines:

While scanning the co-occurring terms that appeared in twits relating to AAPL stocks, a few of them may jump out. *What does Warren Buffett have to do with Apple? Why is India a related term today?*

These are the connections we want to create for our audience. A quick scan of the related keywords generated for them has given them the perfect direction to further research Apple's current events.

Next, they can move to the right-hand side twits panel to see what traders are saying, verbatim, about these current events or do a simple Google search of the terms *"AAPL Warren Buffett".* If they did so, they would see that the news recently broke that Buffett's money management team has invested deeply in AAPL stock, injecting the market with confidence that Apple will perform well (because Warren Buffett thinks it will).

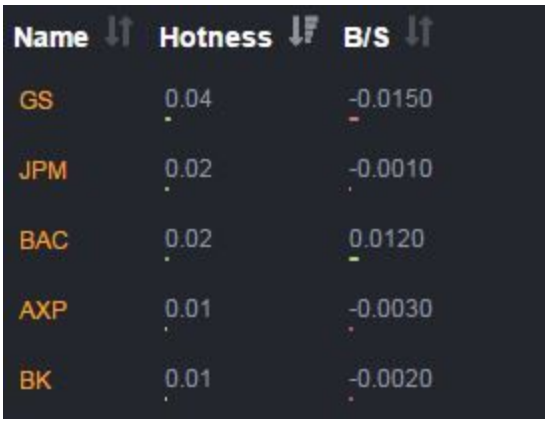
This is the exact phenomenon our visualization hopes to capture. So much of financial trading is centered around how well people *think* a company will do, rather than actual concrete success of the company itself. If a big financial player becomes invested in a firm, the market buzzes

with positive opinions about that firm, simply because they *believe* in someone else's predictions.

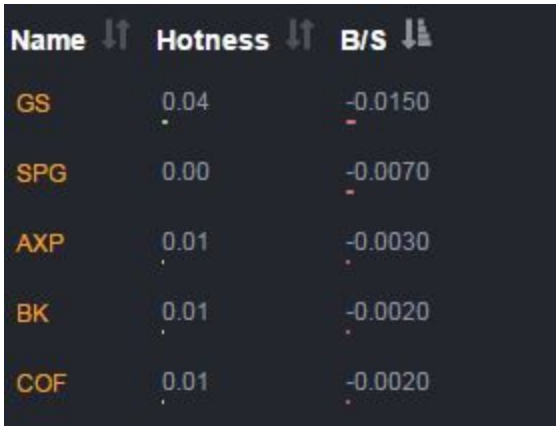
Our visualization tool also incorporates some capabilities for a more experienced trader. We quantify the market's confidence in a stock as its buy/sell ratio.

Supposed that a trader wanted to build or change their current portfolio, by selling one of their current equities and buying another. First, they might choose along the top panel what financial sector they were looking to trade in. Fund managers, in particular, will build equity portfolios around a specific sector to maximize their profits from the industry's growth, and they will advertise shares of their funds as such.

Then, they may use the overview list to look at the relative hotness (popularity) and B/S ratio (confidence) for a specific stock. For a hedge fund manager looking to switch around their financial sector equities, they would be able to see on this list that Goldman Sachs (GS) is being twitted about the most (highest hotness), and most traders are saying to *SELL*(most negative B/S ratio).



Name	Hotness	B/S
GS	0.04	-0.0150
JPM	0.02	-0.0010
BAC	0.02	0.0120
AXP	0.01	-0.0030
BK	0.01	-0.0020



Name	Hotness	B/S
GS	0.04	-0.0150
SPG	0.00	-0.0070
AXP	0.01	-0.0030
BK	0.01	-0.0020
COF	0.01	-0.0020

If the trader currently held GS equities in their portfolio, they may do further research to decide whether to offload a stock that is expected to underperform. The trader in question can use this very same hotness list as a tool to shop for a suitable replacement for their portfolio.

If they don't already hold GS stock, they may want to buy up if up while the price has depreciated.

Limitations and Future Works

What are the major limitations of your project at this stage (it's fine to talk openly about what does not work well yet and what you have not been able to realize)? If you had more time to develop your application further, what would be the next steps?

One of the major problems we didn't solve is that the stock price information that we plot in the chart doesn't show a good representation of the real market information. The stock price occurs to be unchanged for some time periods and we found that Yahoo Finance API for Python seems to experience maintenance problems as we read from the logging history from our server side that the api experience request error (YQL Query Error: Query failed with error: "'No definition found for Table yahoo.finance.quotes'") for some reason. We expect to find a good alternative for getting accurate and real-time stock price data.

B/S data is also not sufficient to support the analysis in terms of the correlation between stock price and B/S index. The B/S index we have is based on the twits' sentiment API from StockTwits, which has Bearish/Bullish tag for about 10 percent of the twits. The Bearish/Bullish tag is helpful for our project and we expect to proceed with building sentimental analysis algorithm for untagged twits which could improve the insights into the community consensus.

Our visualization tool provide an idea of how the dashboard may help users to work and deal with stock market. Given more time, we will conduct field test with both traders and new-comers to verify the underlying problems and improve our solution as the dashboard.