

Graval: A Company Valuation Visualizer

By: Uma Desai, Anne Ku, Lauren Pudvan
Olin College of Engineering: Software Design Spring 2016

Background

We chose to combine components of data analytics and visualization to create a tool that visualizes the relationship between a company's stock price and public opinion (based on top google search results) during a given period of time. This program is designed primarily as a tool that allows the user to easily visualize the connection between interesting events that may have skewed public opinion of a company, such as recent outbreaks of foodborne illness at Chipotle or the BP's Deepwater Horizon oil spill, and the effect it had on the company's stock price. This could also be used by investors looking for insight into predictions of a company's future stock price based on current events or by companies looking to better understand the relationship our tool analyzes.

Features

Our correlation tool produces a graph that overlays a company's public opinion and stock values on an interactive scatter plot. The x axis represents the time period that the user inputs into our program, the left y axis represents the range of stock values, and the right y axis represents the range of public opinion values. When the user moves the cursor over a data point, they are able to see the numerical value of the data point and, for public opinion values, the top headline of google search results for the company that day. The graph also has zooming, panning, and selecting capabilities that allow the user to zoom in on the information for specific time periods on the graph. All of these features work to help the user draw a connection between events that skewed public opinion of a company and the resulting effect it had on the company's stock price.

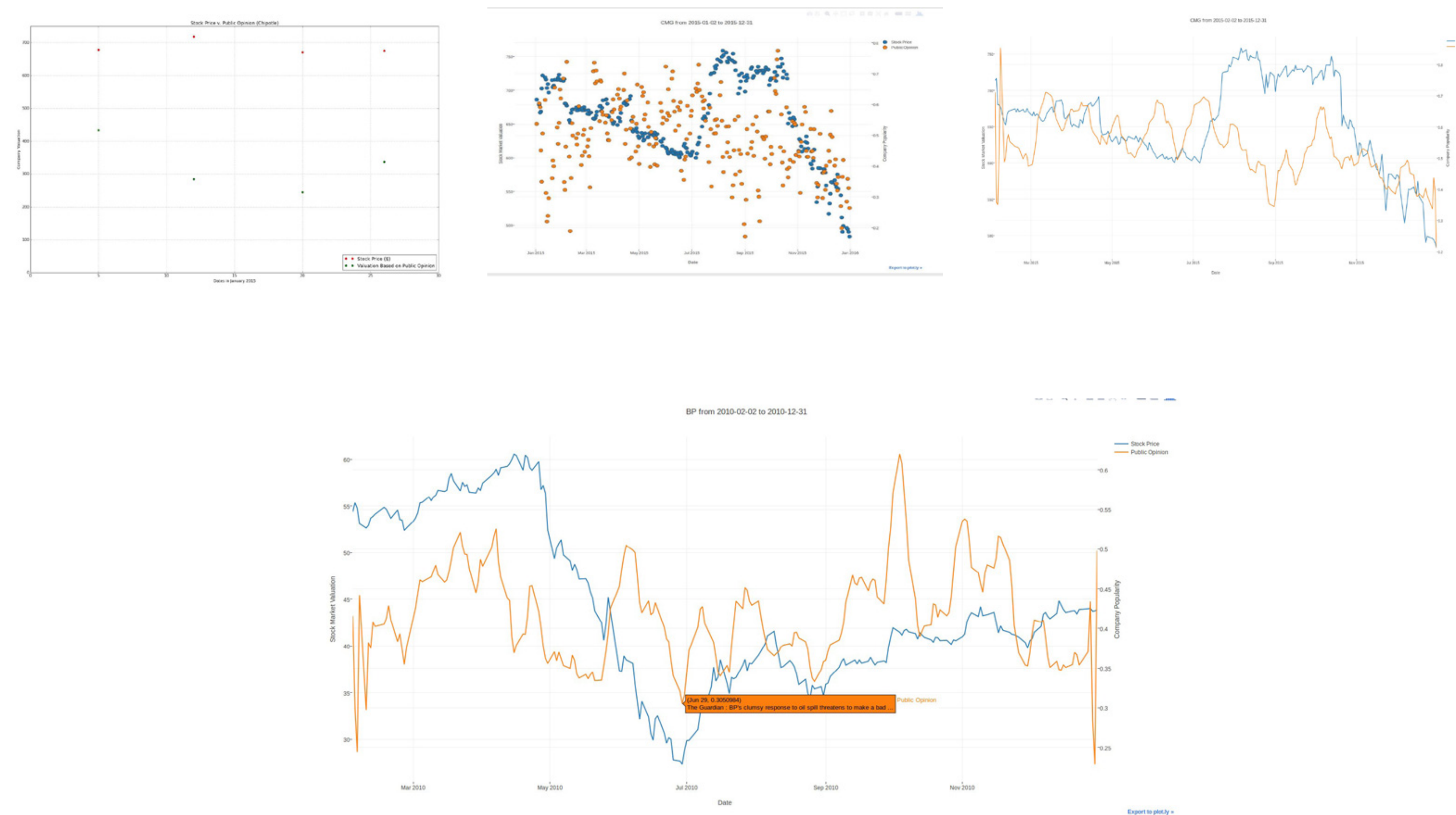
Assumptions

That a company's stock price is affected by its customers' positive or negative opinions of them.

That the polarity of Google search results for a company can be used as an indicator of that company's public opinion.

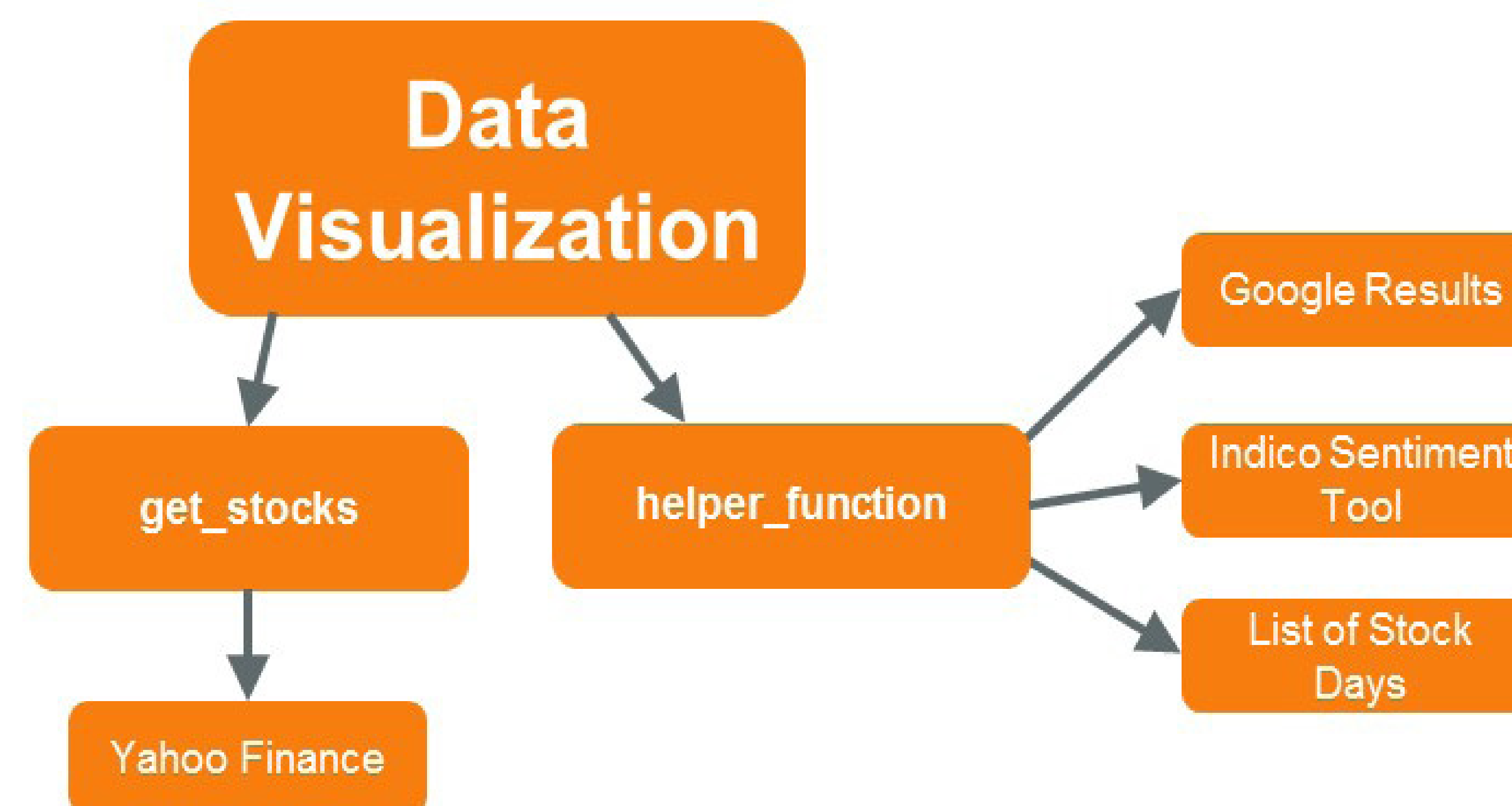
That the most relevant Google results are at the top. So we utilized a weighting system where we multiply the first sentiment value by 9, the second one by 8, etc., add them up, and divide by 45 (since there are 9 results per page).

Iterations



This is an example of one of the graphs that our tool can output when the inputs 'BP', '2010-02-02', and '2010-12-31' are entered into the interactive_scatter_plot.py function as the company name, start date, and end date.

Program Structure



Algorithm

This program obtains stock data for a company for every day the stock market is open through the get_stocks() function. Public opinion values are obtained by running Selenium Webdriver to crawl Google searches, using Indico's sentiment processor to analyze the first page of google news results (for the given day and company), and outputting an average public opinion value for that day. The data visualization component of the project uses Plotly's python graphing library to plot an interactive graph overlaying the stock data with the sentiment results. It is executed in one function that calls get_stocks() and get_sentiments() to obtain the data that is plotted. This function takes the company name, start date, and end date as inputs and is the main function that a user would interact with.

Future Work

Add features to the graph that make it more user-friendly and provide more interesting information (i.e. range slider, dropdown menu, etc.).

Create a webpage that makes our tool more accessible to users

Continuing to improve the accuracy of public opinion values by filtering out unrelated content and incorporating other sources

Attributions

Yahoo Finance API
Selenium Webdriver
Indico Sentiment Processor
Plotly's Python Graphing Library: Used to produce the graph that overlays the stock price and public opinion values