

Nathaniel Kong (915178616)
Ryan Lazzareschi (916355661)
Peter Lo (915185612)

Visualizing S&P 500 Companies Under the Impacts of Coronavirus

Data

Our goal for this project is to visualize the recent volatility of the United States stock market in light of the Coronavirus epidemic. The data for this project is the S&P 500 stock market of the United States (the S&P are the 500 largest stocks in the United States).

Most visualizations of the stock market are simple candlestick charts, but don't provide a larger picture of the stock market as a whole. We wanted to provide both a more holistic and an exploratory visualization method.

Initial Proposal

Illustrating how the stock market responds to the Coronavirus requires recent data, so our initial proposal used the Finnhub API to gather real-time data. The proposed visualization would fit the minimum requirements for data filtering, using an API to gather data, and data selection interaction.

The original visualization included a bubble map, parallel coordinates graph, and a difference chart. The bubble map would provide an overview of the stock market. Each bubble would represent one stock market sector, and bubble size would correspond to market capitalization size. On a click to a bubble, companies in the sector would be displayed on the parallel coordinates graph. The parallel coordinates graph would display variables such as market capitalization, CEO salary, and total price. From the parallel coordinates graph, you could select two companies to compare company percentage gains in the difference chart.

Visualization Evolution

After the presentation proposal and advice from Professor Ma and Maksim, we decided a tree map would be a stronger visualization than a bubble map. A tree map can display individual companies of a sector (in contrast to a bubble map that only displays the market capitalization of a sector) and removes unnecessary factors that would have otherwise been present in the bubble map, such as proximity and location of bubbles relative to each other.

We also added a candlestick graph so that a user could visualize a single company. The addition of a fourth visualization provides another level of granularity and a familiar way to see stock prices.

After we discovered the Finnhub API was limited in both functionality and scope, we utilized the IEX Cloud API and Alpha Vantage API instead. Due to the change in API, we changed the axes of the parallel

coordinates graph. We removed CEO salary, and added average total volume, P/E ratio, week 52 high, week 52 low, and percent change. The companies of the S&P 500 are mostly stagnant, so we opted to use an CSV to optimize loading time and calls to the API.

Design Justification

We implemented a tree map because S&P 500 sectors are categorical. The categorical data is also why we used color to label each sector. Furthermore, treemaps clearly illustrate sector size (market capitalization) and can display individual companies in a single sector.

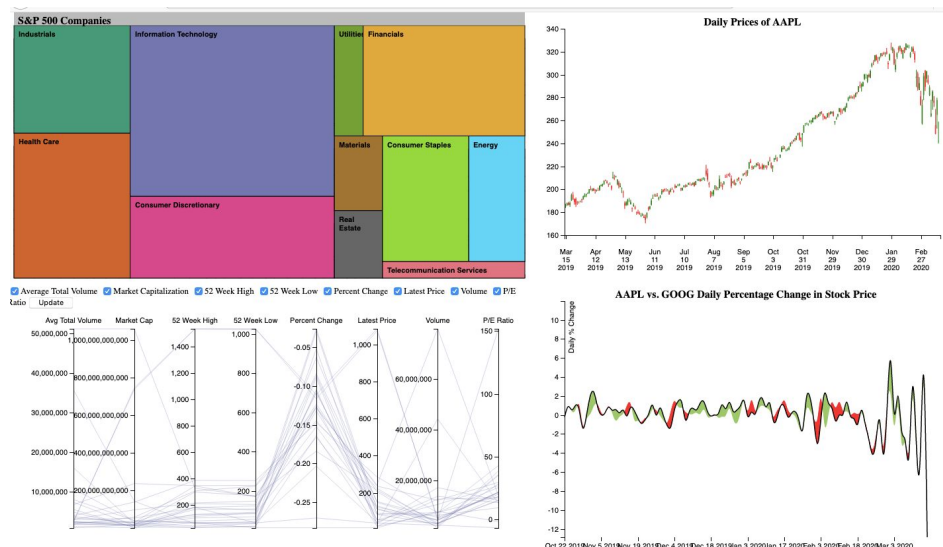
We implemented a parallel coordinates graph because the graph can display all companies without occlusion, and display multiple different aspects about a company. Additionally, the parallel coordinates graph is good at filtering data, so the parallel coordinates graph provides a good way to link the treemap and our other visualizations.

The candlestick chart and the difference chart both are ways to illustrate finer granularities of our dataset. We implemented the candlestick chart so that the user could analyze the price movement for a standalone stock. The difference chart is implemented so the user can compare individual performance to another stock performance. We limited the scope of the candlestick chart to one year and the difference chart to three months because the developments in the coronavirus epidemic are recent.

Final Visualization

Startup

We chose to view all four visualizations side-by-side so the user can see a comprehensive picture. The candlestick chart, difference chart, and parallel coordinates chart are all displayed on startup but change through interaction.



Implementation Details

The tree map is a static visualization based upon the S&P 500 companies with financial information pulled off datahub.io in the form of a csv file. According to datahub.io, the source of their dataset was obtained from the official S&P Dow Jones Indices website. We implemented zooming in upon a click based upon company information at stock market-level and sector-level. We utilized a hovering feature that illustrated companies in the hovered sector, and linked the hovered company to highlight the hovered company in the parallel coordinates graph. On a click to a single company, a candlestick chart is drawn and the company is selected in the parallel coordinates graph.

The parallel coordinates graph utilized calls to the IEX Cloud API, and formatted JSON data like csv data. When a treemap sector was clicked, we passed the color and the corresponding companies to draw each path. We used another hovering feature for the parallel coordinates graph, that also selected and highlighted the same treemap company.

To implement parallel coordinates interactive remapping, we used html buttons to indicate whether a user wanted to draw the axis. On a call to draw the parallel coordinates graph, we used if statements to check if a box was clicked. When the event listener for the 'Update' button was called, the graph was redrawn.

After two stocks have been selected from the parallel coordinates graph, the difference chart makes calls to the Alpha Vantage API and receives time-series stock data. We convert this data to daily percent change so that comparison in the difference chart is easier.

The candlestick also utilizes the Alpha Vantage API, retrieving time-series stock data for the stock that is chosen from the treemap. Using candlesticks informs the user of open, high, low, and close price for that stock on any given day within the past year. The user can also zoom in and out of the candlestick chart by scrolling to focus on a specific time period. By dragging horizontally, the candlestick chart will also shift the time scale so the user can see stock data for surrounding dates.

Visual Encodings

The marks in the treemaps are areas, the marks in the candlestick chart and difference chart are areas and lines, and the marks in the parallel coordinates chart are lines. The channels encoded are size, color, and position.

The color encodings in the visualizations are for categorical data, so color is used to label. We chose the colors to distinguish themselves from the white background. Similarly, the colors in the parallel coordinates chart corresponds to the selected sector color.

In the candlestick chart, red represents the price dropping for the given day and green represents an increase in price. For the difference chart, color encodes the performance of one stock relative to another.

Green area would represent the first stock outperforming the second stock while red area means the first stock underperformed the second stock.

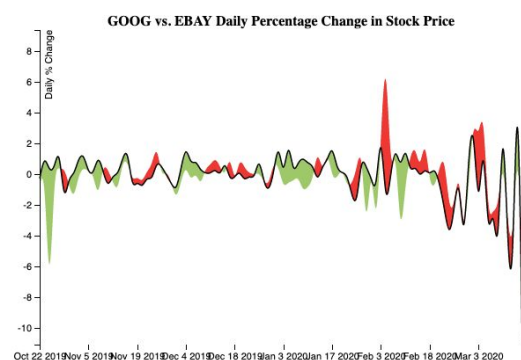
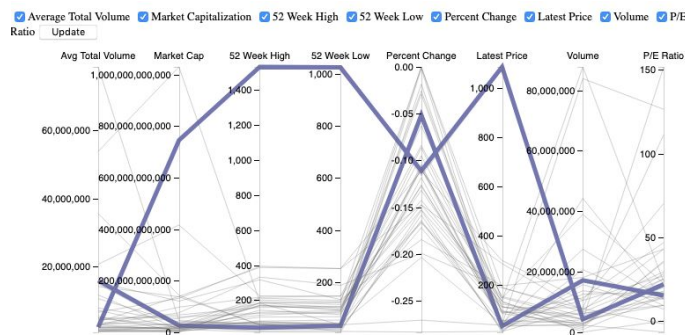
When hovering over the treemap and the parallel coordinates, the stroke-width and opacity channels of the currently hovered-over element changes, indicating to the user the position of their cursor and the element's corresponding point in the other graph. For example, hovering over a treemap company rectangle highlights both that rectangle and its line on the parallel coordinates graph, and vice versa.

Size in the tree map represents market capitalization, and has no representation in the parallel coordinates graph. For the candlestick chart, stem length represents the volatility of a stock for a given day while the size of the rectangle represents the percent change for a stock for a given day. In the difference chart, size of the red and green areas represent the difference in percentage gains for a given day. For example, if there were more green area than red area in the chart, that would indicate that the first stock saw a higher average daily percent gain.

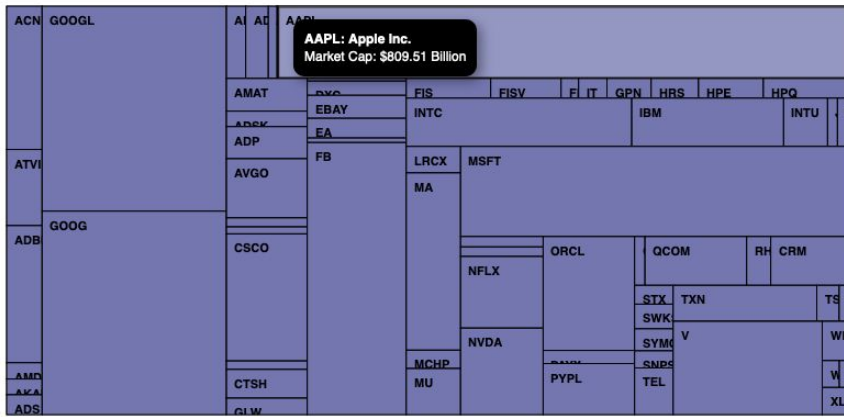
In the parallel coordinates graph, candlestick chart, and the difference chart, position is a channel. For these cases position encodes the variable on the scale, telling the user how large or small a variable is.

Interactions

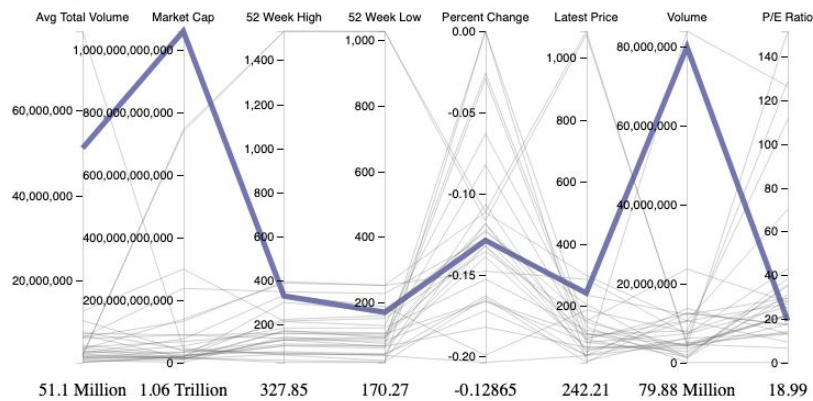
A user can select a sector to zoom in the treemap. Once in the sector-view of the treemap, a user can hover over a company in the parallel coordinates graph or treemap to see its position in the corresponding graph (see second figure below). Furthermore, a user can click on a company in the treemap to display on the candlestick chart. Once a first company is selected, this company will show up in the candlestick chart, and will be highlighted in the parallel coordinates graph, while other companies will be greyed out. A user can select a second company that will also be highlighted, to display in the difference chart (see first figure below).



S&P 500 Companies / Information Technology

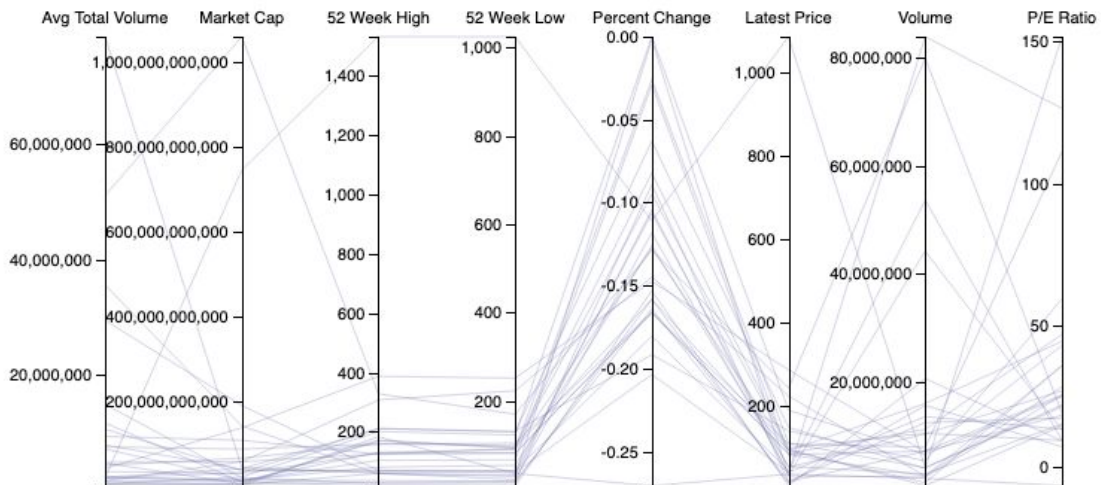


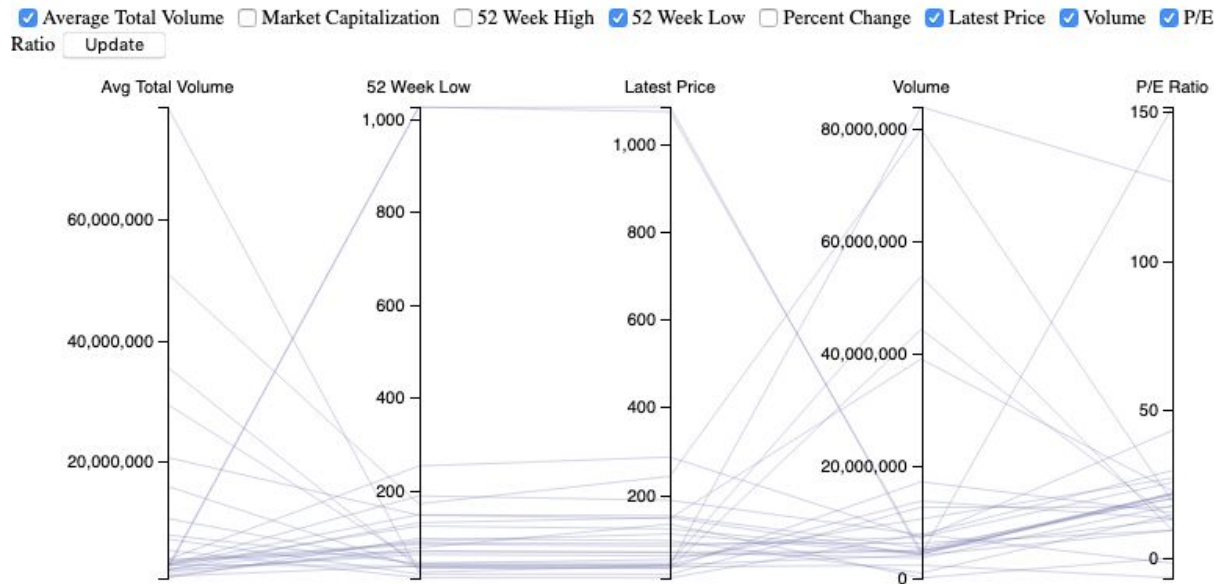
☒ Average Total Volume
 ☒ Market Capitalization
 ☒ 52 Week High
 ☒ 52 Week Low
 ☒ Percent Change
 ☒ Latest Price
 ☒ Volume
 ☒ P/E Ratio



Additionally, the parallel coordinates graph includes interactive re-mapping of attributes. A user can select which axis they want to see displayed on the parallel coordinates chart (see figures below)

☒ Average Total Volume
 ☒ Market Capitalization
 ☒ 52 Week High
 ☒ 52 Week Low
 ☒ Percent Change
 ☒ Latest Price
 ☒ Volume
 ☒ P/E Ratio





We also implemented tooltips for each visualization. The tooltip of the treemap displays company name and market capitalization of sectors and individual companies, which is the channel for size. The tooltip for the parallel coordinates graph displays the stock ticker, the company name, and the exchange the stock lives on. The tooltip for the candlestick chart displays the opening, closing, high, and low of a price for a particular company on a particular day. Finally, the difference chart displays the percent change for a given point for each company being displayed.

Task Division

Tasks were mainly divided by visualization, although interactions and linking were not necessarily bound by each individual in the group. Peter created the Treemap, Nathaniel created the Parallel Coordinates Chart and the report, and Ryan created the Candlestick and Difference charts.