Project Milestone 2: Group 6

Introduction

Stock market data has long since attracted the attention of numerous people, while the mere stock trend charts themselves provide little, if none at all, assistance in the analysis and understanding of stock market data. In Milestone 1, our group have decided to construct a **design studio** on multiple possible influencing factors.

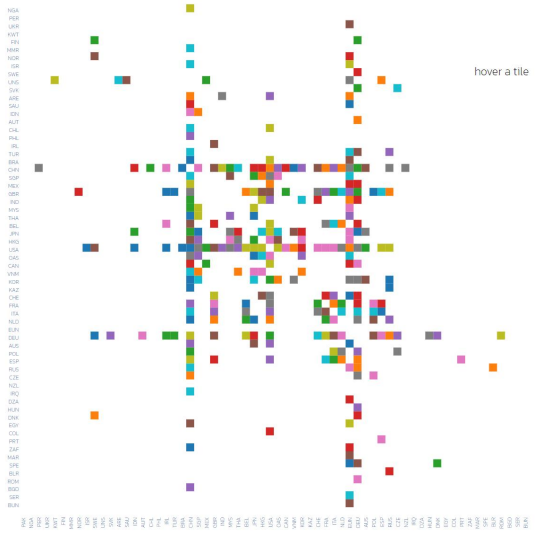
The factors selected are **Political and International Influence**, **Company Factors**, **News and Events**, **Macro-economic variables**. Our approach involves developing distinct interactive visualizations for each of these factors, using D3 and shiny app. In later plans these separate parts will be integrated with stock data visualization to provide a full view of stock trend analysis

International Influence

The first area of focus is the International Influence, predominantly through the lens of trade value. In an effort to provide a clear and interactive understanding of international trade dynamics, we employed D3 technology to create a comprehensive interactive heatmap.

This heatmap is based on data sourced from the International Trade Database on Kaggle(International Trade Database (kaggle.com)), provided by UN Comtrade for the World Integrated Trade Solution (WITS) platform. The dataset includes seven columns, each representing the export value of goods from one country to another in a given year. Our visualization specifically highlights the top 300 international trades of 2020, ranked by trade value.

This interactive graphic is designed to provide a detailed view of the international trade amounts between countries. Each tile on the heatmap represents a distinct trade, with detailed information about the trade value and the countries involved accessible by hovering over the tile. The x-axis of the heatmap indicates the exporting country, while the y-axis shows the importing country. An innovative feature of this visualization is the emphasis placed on neighboring tiles and axis labels when a user hovers over a specific trade, thus enhancing the interactive experience and providing a clearer understanding of trade relationships.

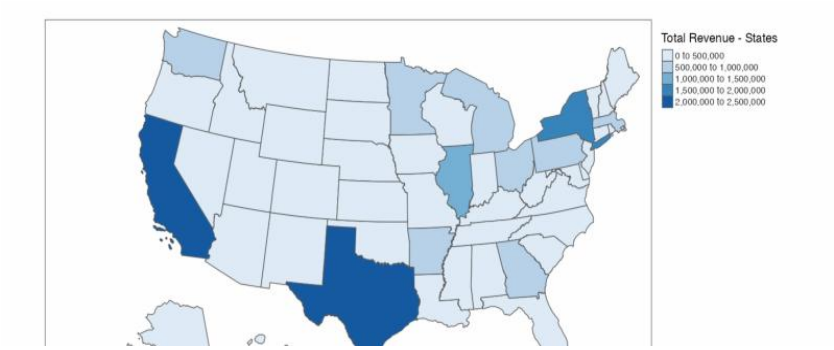


Sample graph shown above, video and code at project/milestone2/international.

Company Revenue

Another critical factor is the revenue performance of individual companies. To effectively visualize this aspect, we have planned to implement an interactive platform using the Shiny App, a tool known for its capacity to integrate multiple interactive plots on a single page. The data for this visualization is derived from the Fortune 1000 dataset, which includes comprehensive data on U.S. companies for the year 2022.

Our Shiny App visualization allows users to interactively explore various aspects of company performance. Users can manipulate two slider inputs to select specific ranges of company ranks and employee numbers, narrowing down the dataset to a more focused group of companies. The resultant visualization comprises multiple components: a pie chart and a box plot illustrating total revenue across different sectors, another pie chart showing the distribution of companies across these sectors, and a heat map on a U.S. map indicating total revenue across states. This multifaceted approach enables users to gain a holistic view of company revenues and their distribution across different sectors and geographical locations.

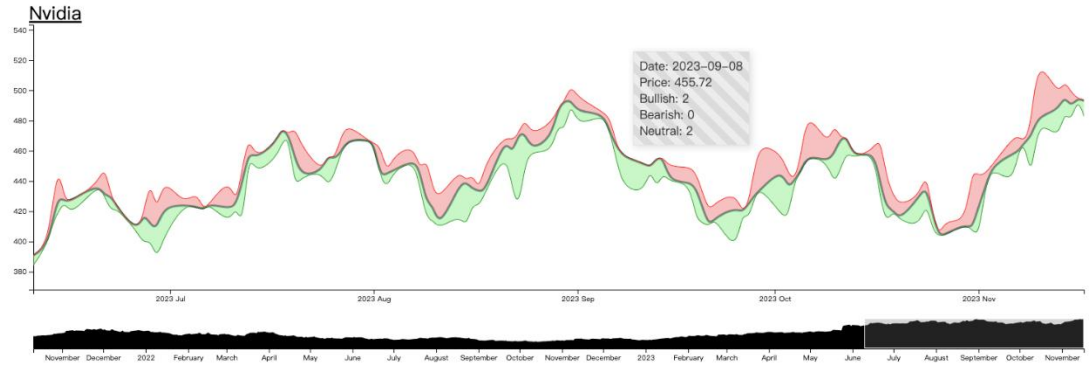


Sample graph shown above, video and code at project/milestone2/company.

News and Events

The stock market is also significantly influenced by specific events such as product launches, mergers and acquisitions, insider trading, and more. However, the visualization of these events' impact on stocks is often not sophisticated enough. Traditional methods employed by brokerage firms, like annotating timelines or presenting data in separate modules, are somewhat limited in their effectiveness. Our approach aims to revolutionize this aspect by leveraging the capabilities of large language models for a more nuanced analysis.

We source our data from seekingalpha.com and the Sina Finance API and utilize a Python script to scrape news data for individual stocks. To determine the sentiment of news and its potential impact on stock prices, we employ GPT-4 Turbo. The resulting data is then visualized using an interactive line chart created with D3. This chart not only displays the overall stock price movement but also incorporates a unique feature: the representation of the quantity of bearish (red) and bullish (green) news. Users can interact with the chart by creating and moving a brush on the stock price timeline, which then updates the line chart to reflect the selected time frame. This innovative visualization tool allows even those without a financial background to quickly grasp the complex interplay between news events and stock price movements.



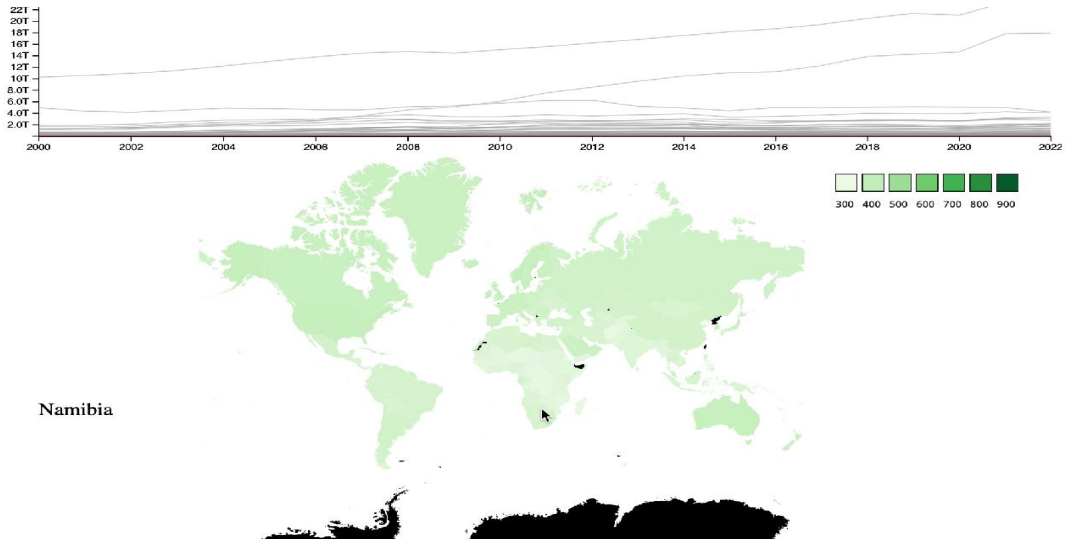
Sample graph shown above, video and code at project/milestone2/news.

Macro-economic Variables

Macro-economic variables, such as GDP growth, unemployment rates, and inflation, are widely acknowledged as major influencers of stock market trends. To address this, we have proposed creating interactive charts using D3, which would overlay key economic indicators with a company's stock price for a more comprehensive analysis.

In demonstrating our approach, we chose to visualize the macro-economic indicator of GDP. The data for this comes from the World Bank, as available on Kaggle. Our visualization combines a line plot and a map, allowing users to explore the data from multiple perspectives. The line plot provides a time series view of economic indicators, revealing trends and patterns over the years.

In parallel, the map offers a spatial analysis, enabling users to compare GDP and other metrics across different countries. A unique feature of our map visualization is the color filling based on GDP per capita, which uses a symlog scale to accommodate a wide range of values.



Sample graph shown above, video and code at project/milestone2/macro_economy.

Trade-offs

For International Influence, we chose a heatmap over alternatives like a line chart. The heatmap excels in showing dense trade relationships at a glance but may not capture trends over time as effectively as a line chart.

In visualizing Company Revenue, we used a multifaceted approach instead of a stacked bar plot. While our approach offers a diverse perspective, a stacked bar plot could provide a more straightforward comparison of revenues across sectors, though it might struggle to display the detailed breakdowns our current method offers.

For News and Events, we used sentiment analysis integrated with stock price movements, instead of a quantitative visualization like the VIX index. Our method provides a

nuanced understanding of news impact but might lack the straightforward volatility indication that a VIX index visualization would offer.

Each chosen method has strengths in certain areas but also limitations compared to the alternatives, underscoring the need for careful consideration of the target audience and the specific insights we aim to convey.

Summary

In summary, our group's endeavor in this project is to explore the multifaceted nature of stock data, extending our analysis to several key factors that influence stock market dynamics. Building on our initial concept from Milestone 1, we have utilized various tools and knowledge acquired in our coursework, including D3 and Shiny App, to create four distinct interactive visualizations. Each visualization is dedicated to one of the four identified factors: International Influence, Company Factors, News and Events, and Macro-economic variables. The methods each possess advantages but also have certain trade-offs compared with other possible methods suggested in Milestone 1. In later approaches, we plan to combine all separate visualization of different factors with stock data to provide a comprehensive and inspiring overview of stock trend and its influencing factors.