**Final Project Report：Stock Data Visualization**

Helen Zhao, Yihan Luo, Boyang Lin

**Introduction**

Financial illiteracy is a real problem in the modern-day world. Meanwhile, investing is probably one of the most important ways to ensure financial safety. However, since most people are afraid of the stock market and aren’t willing to risk putting their money in it, our project aims to help users gain a better understanding of the stock market overall.

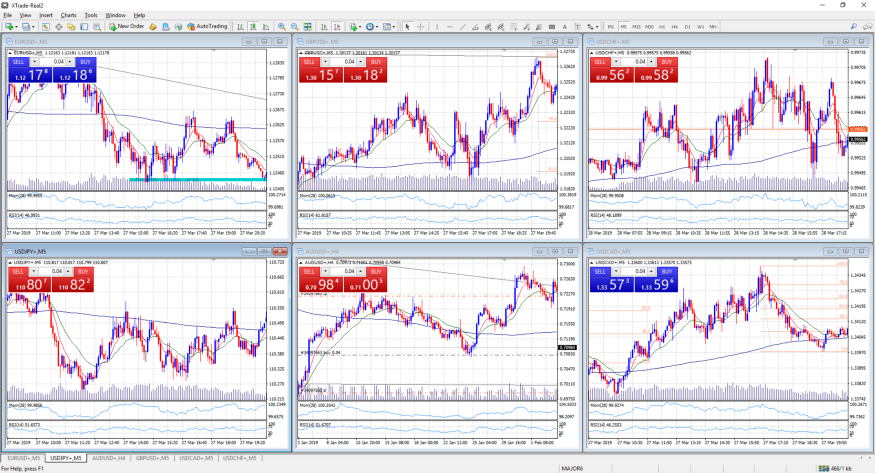
Understanding and comparing stocks is not as easy as it seems. Stocks generate a lot of data every day. Every stock transaction will affect the stock price. Most people are familiar with daily prices. But in fact, this is only the price of the last transaction. However, the ability of digital meters to display data is limited. Paradoxically, the data are very accurate, but the numbers themselves do not really show the relationship and trend between different stocks. Therefore, an important solution is the visualization of stock data. Although this method is not as accurate as the numerical table, it has some strong advantages. In general, visualization can identify the trends and patterns of the represented data faster. Human perception of graphic shape is much better than that of digital table. Not surprisingly, amateur traders often use visual data such as line charts for technical analysis. But visualization offers more advantages. The exchange of data is easier. In addition, there is more information density and less space required. In addition, visualization can also convert data and data relationships into shapes, so that people can understand the underlying data more easily and faster. In addition, visualization can well show the relationship between data sequences. The patterns and fluctuations of time series can be easily represented graphically. Therefore, visualization should be a powerful tool for comparing multivariable data tasks.

In fact, in order to help users better understand the stock market and make investment decisions, the visualization tool needs to include the following aspects:

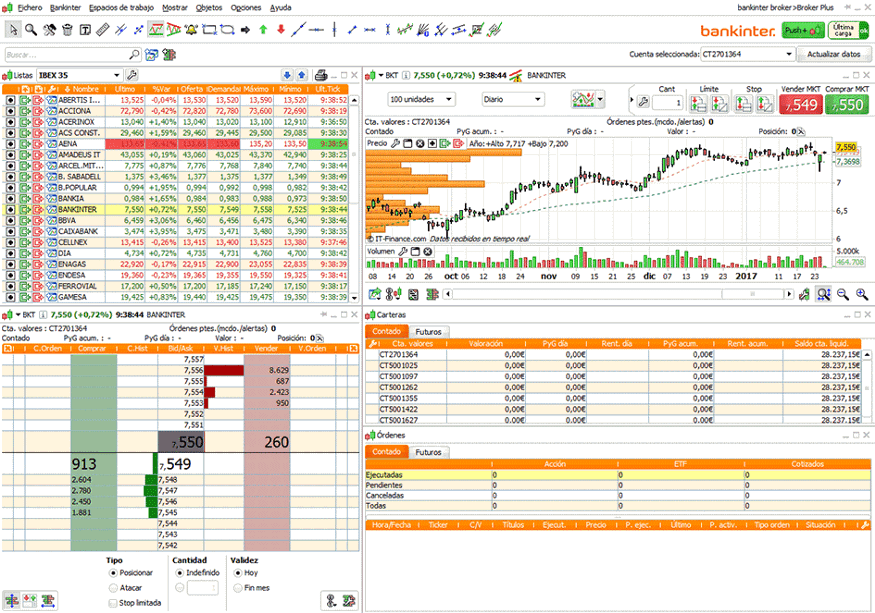
First, the response to the overall stock market situation, so that users can quickly understand the overall situation of each sector of the current stock market. Second, the change of stock data based on time allows users to understand the changes and trends of the stock market in a certain period of time. The third is the comparison between different stocks. By comparing several stocks in a certain period of time, it can better help users make investment decisions.Fourth, the analysis and judgment of their own portfolio. Users should be able to have a clear understanding of their current portfolio and different stocks in the portfolio by visualizing a large number of numerical values.

**Related work**

There are many famous chart platforms and software packages that can be used to display and analyze stock market charts.

Metatrader is a lightweight and simple tool. Despite it has certain limitations and the fact that it is only being used in OTC markets it has become one of the most popular tools among retail active investors; this is probably due to the fact that it has been adopted by a large number of retail OTC brokers around the globe.

Pro RealTime is another widely used tool among small and retail investors. Capabilities are functionalities are much better than Metatrader and it is used for regulated markets. It has more functionalities such as order book depth, volume profile and many analysis tools.



Although these software can well analyze recent events - or not if you plot a larger time period - it is usually a very bad tool to systematically observe what is happening in the disk over a longer period of time. Understanding what happened on any day two years ago is not as easy as people think, and it is likely to be impossible at all, because such chart software is designed to be used as a trading platform rather than a historical and strategic analysis platform. There are a large number of indicators, trend lines and auxiliary lines in the visualization of these software, but for users, the more complex is not the better.

**Methods**

The most representative stocks are probably the S&P 500, a list of stocks of the 500 largest corporations by market capitalization. Most of the big companies we know are in it, like amazon, google, Facebook, etc. We collected the prices of these stocks from the past five years. It is worth noticing that the stock market has been extremely volatile in these years because of many global events, like COVID 19. We hope this data can help users familiarize themselves with the overall trend.

The first function is rather straight forward, we generated a line graph and a bar graph that displays the daily change in the S&P 500 index. Users can specify the timeframe for which they wish to view the graphs. For example, they can choose the time during Olympics, or covid. The trend should be very clear and gives users an idea of the stock market’s overall trend.

The second function is a treemap representation. The size of each block is proportional to the company’s market capitalization, while the color represents the increase or decrease in price. The companies are grouped on the treemap by sector. To make the data more interesting, we added a time element. Users can see the overall color/sector change based on time, which is both straightforward and impactful. This function gives user the opportunity to explore how world events influence the stock market in different ways. For example, different sectors might behave differently. (health care sector during covid?)

The third function is the line graph. For users who want to know more about specific stocks, we provide them with a customizable tool that can cover the stock prices of companies selected by users and compare the differences of stocks of different companies in an intuitive way. Even among the 500 largest companies, there are huge differences in share prices. For example, we can select different stocks from the same industry and compare their performance in the same time range. In addition, we can compare stocks in different industries. We have added an interactive mode in this figure so that users can know the specific quantitative differences between two stocks at a certain time node on the basis of understanding the qualitative differences of different stocks, which is very helpful for making investment decisions at a certain time node.

The first three functions focus on users' rapid understanding of the stock market situation. On this basis, users may also need to control their existing or future portfolio. We hope to provide a better visualization tool beyond simple numbers, so that users can have an overall understanding of their portfolio. Especially when users have a large number of types and complex investment situations, the visual information can help users well.

Our last feature is a stock portfolio simulator. Users will be able to test out their long-term strategy here. This tool will show users the large potential for returns they stand to accrue from investing in the stock market over time. They can type in the stock ticker name (such as: AAPL, FB, GOOGL…) and the number of shares they want to add to their portfolio. We provide two different forms of visualization. The doughnut chart is a good way to visualize the component parts of users’ stock portfolio. We also provide users with a bubble chart. It’s more fun and gives more information. Each stock is represented by a bubble, where its x label tells users how much their share position has changed for this stock, and its y label represents the percentage change. Size of the bubble is proportional to users’ current share position of the stock. Different colors of bubbles represent different sectors, which shows portfolio diversity.

Hopefully users will feel empowered by this real stock market and discover that the stock market isn’t so scary.

**Results and Analysis**

Based on the historical data of the stock market we collected and the visualization tools we developed, we can well understand the stock situation at a certain stage or time node.

For example, we want to analyze the changes in the stock market during the epidemic. We can use the first tool to select the time range to understand the overall changes of the stock market; Through the second tool, by pulling the timeline, we can clearly see the response of different plates to the epidemic at different stages of epidemic development in chronological order: in the tree view on January 31, 2020, each plate shows a large area of red, and then each plate shows an overall alternating rise and fall trend.

**Treemap:**

It provides a clear overview of the whole market and displays complex hierarchy: reflect the relationship between stocks in different sectors. It adds additional context to the tooltip (displayed when we hover over a specific stock)。

The red and green color palette, which seems to be widely accepted in the financial data world, makes this dashboard unavailable to color blind people.

It is difficult to make an accurate comparison according to the color and irregular box size.

**Line Graph:**

It reflects the evolution of time well. Although the line chart is very simple, we should focus on how to increase the interaction between it and users. Its complexity comes from the built-in interaction. Here we can set the starting point and ending point of time, and determine the time scale we want to observe. We can observe the changes of the stock market within a week, a month, a year and five years.

**Discussion**

In the whole process of visual design and production, we have focused on the following issues:

In the whole process of visual design and production, we have focused on the following issues:

1. How can users make better decisions based on the overall stock market change trend? The online stock market data visualization is mostly the overall trend of the stock market and the line chart of a single stock. On this basis, we think about how users need more visualization to help make decisions. On the one hand, users tend to focus on a small number of stocks when making final specific decisions. At this time, the overall trend of the stock market and the line chart of a single stock can provide very limited help. Therefore, our visualization tool provides a more targeted comparative analysis of a small number of stocks. On the other hand, the change trend of users' own portfolio seems to be more meaningful than the overall change trend of the stock market. Through the visualization of stock holding data, it is helpful for users to classify and further analyze stocks.

2 different interaction modes for selecting time range. In our visualization tool, we have two different time range selection methods: drop-down menu selection and drag bar selection. We consider the advantages and disadvantages of these two different interaction modes in the design process. The drop-down menu selection is more accurate and faster. Especially when the user has a clear purpose of analyzing stock data in a certain time range, it can lock the time period accurate to the day in the least operation. Compared with the drop-down menu, the drag bar selection is not accurate enough. Users often need to make careful adjustments several times before locking at a time node. However, unlike the drop-down menu, the interactive mode of drag is continuous, that is, our visual image has time continuity without interruption in the user's operation process, which plays an effect similar to animation. Compared with the purposeful analysis mentioned earlier, this effect can enable users to make independent discovery in the process of interaction. Therefore, using different visualization methods in different diagrams can have different effects.

3. Combination of overall trend and specific data. The amount of data in the stock market is huge. The overall trend reflected by the huge amount of data and the specific value of a node are very important to users. Under different visualization scenarios, we need to consider the detail of the displayed data, and judge whether it is more important to convey the information of the overall trend or to display more detailed and specific data, Or combine the two through some interactive way. For example, in the tree chart, although we have specific increase information of each stock, we try to convey the overall trend to users. Therefore, by transforming specific data into visual red-green comparison, we can transmit information more quickly and effectively than specific numbers. We also provide users with an interactive channel to understand a specific data through mouseover interaction. The same method is also applicable to the trend comparison of multiple stocks.

**Future Work**

In the current version, our four pages are slightly independent. If we have more time, we should strengthen the connection and interaction between the four visual charts. For example, when we click a stock in the portfolio, we can jump to the trend chart of the stock. In addition, in the process of the project, we used API as the data source. In this way, we can turn the project into a visual application that can reflect the stock market in real time.

In fact, as for accessibility, the common color matching of red and green in the stock market may be a disaster for color blind people. If this is fully considered, we should consider using other colors (such as red and blue).