Predicting Stock Volatility

Problem Statement and Background

The stock market is an inherently unpredictable game, where anticipating results requires consideration of specific factors. One crucial factor is stock volatility, a measure of the variance in stock prices over a given time period, reflecting the percentage change in a stock's price. While it doesn't prescribe the "best" stock, understanding stock volatility allows us to assess the risks associated with a particular investment.

Predicting fluctuations in stock prices accurately can be challenging. Internal factors such as layoffs, low morale, payroll issues, and unmotivated employees can decrease productivity and impact a company, thereby increasing its stock volatility. External factors like GDP, inflation rates, unemployment, industry performance, and political and regulatory issues can also contribute to this volatility.

However, certain economic factors, like unemployment and interest rates, are beyond a company's control. It's intriguing to explore how these factors influence changes in stock prices and their relationship with volatility. To investigate this, we developed two sets of code based on different models. The first employs standard deviation to calculate stock volatility, while the second utilizes the GARCH model, a time-series statistical tool that forecasts time-varying volatility by incorporating past volatility information.

A stock's volatility can be categorized as high or low, reflecting significant or minimal price changes, respectively. Investors seeking high risk and high rewards may find high-volatility stocks appealing, while those preferring reliable and low-risk returns often choose low-volatility stocks.

Ultimately, understanding a stock's volatility benefits young investors in managing their portfolios, provides stock researchers with insights into a company's well-being, and assists investment bankers in tailoring their services to client needs. The benefits extend beyond mere numbers, offering a comprehensive perspective on the risks and rewards associated with invested stocks.

Data Used

To predict the stock volatility we need three datasets:

1. **Stock Price (in $)** (“Yfinance API”)
2. **Unemployment Rate (in %)** (“Unemployment Rate”)
3. **Interest Rate (in %)** (“Federal Funds Effective Rate”)

Our objective was to acquire the latest data while ensuring a sufficient quantity for analysis. Therefore, we decided to collect data from January 1, 2010, to October 31,

2023. Monthly unemployment and interest rates were recorded and stored in a CSV file with corresponding dates, sorted from oldest to latest.

A similar approach was employed for stock prices, where the closing price of a stock as of the last trading day was collected for the specified dates.Parameters were utilized to customize the dates and extract data from the Yfinance API. This data was then merged with the DataFrame containing unemployment and interest rates to ensure aligned dates for comprehensive analysis.

In gathering data for our analysis, we employed automated tools such as the Yfinance API for stock prices and official governmental databases for unemployment and interest rates. The data collection process was designed to respect privacy and ethical guidelines. Since the information is publicly available and pertains to broad economic indicators and stock market data, there are no direct privacy concerns associated with individuals.

While the data collected provides a comprehensive view of market trends, potential biases may arise from what is included or excluded. For instance, focusing solely on specific economic indicators like unemployment and interest rates may overlook other vital factors influencing stock volatility, such as geopolitical events, sector-specific trends, or technological advancements. Additionally, our analysis primarily centers on the U.S. market, which may not fully capture global market dynamics that can significantly impact stock prices.

Data Science Approaches

We employed two distinct approaches to calculate stock volatility.

Approach 1: Standard Deviation

The initial method involves using standard deviation to quantify stock volatility. In the stock market, manual calculation of standard deviation is common. We used the NumPy library to compute the standard deviation of stock prices, unemployment rates, and interest rates. The inclusion of unemployment data and interest rates is crucial because these common economic indicators significantly influence stock price fluctuations, reflecting external factors that impact daily business operations.

The integration of these external factors involved reading a CSV file containing unemployment and interest rate data into a DataFrame. This DataFrame was then merged with stock prices obtained from the API, enabling insightful connections. To gauge variability in both stock prices and economic indicators, standard deviation is employed. A regression model is created using the merged database, and the train_test_split function is utilized for predictions. These predictions are then used to calculate standard deviation through NumPy's std function. A higher standard deviation implies a more significant degree of fluctuation, indicating increased volatility in either the stock or economic data.

Approach 2: GARCH Model

The second approach utilizes the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, a time series analysis tool for forecasting financial data volatility. This model extends traditional autoregressive models to capture time-varying volatility, employing the Autoregressive Conditional Heteroskedasticity (ARCH) structure to quantify and predict volatility fluctuations.

Initially, the Autoregressive Conditional Heteroskedasticity (ARCH) model is introduced, working by calculating and predicting changes in volatility over time using stock price data (closing price). The model observes past volatility, calculated as stock returns (changes in the closing price), recognizing the tendency for volatility to cluster. For instance, if recent days had high volatility, the model anticipates that the next day might also have high volatility, and vice versa.

Subsequently, the GARCH model is introduced to address conditional heteroskedasticity. This means that the variability of a variable is not constant across all values. The model addresses this by conditioning on past information. As new information becomes available, the model updates its predictions for future volatility. This continuous updating allows the GARCH model to adapt to changing market conditions.

Results and Conclusions

Our analysis using both the standard deviation and GARCH model approaches yielded insightful results about stock market volatility. While the integration of unemployment and interest rates with stock prices revealed a nuanced understanding of external economic factors on market behavior.

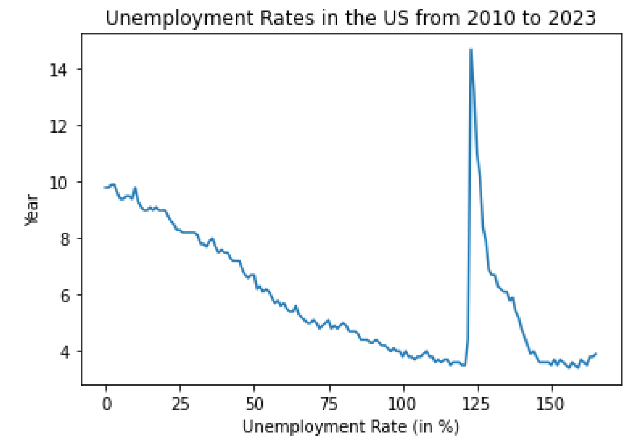
1. Standard Deviation Approach: The standard deviation calculations highlighted periods of increased volatility in correlation with fluctuations in unemployment and interest rate data. Higher standard deviation values were observed during economic downturns, indicating increased market volatility during these periods.
2. GARCH Model: The GARCH model provided a more dynamic and comprehensive view of volatility. It successfully captured the conditional heteroskedasticity in the data, showing that volatility is not uniform but varies over time. The model's predictions were particularly effective in identifying periods of clustered volatility, aligning with major economic events.

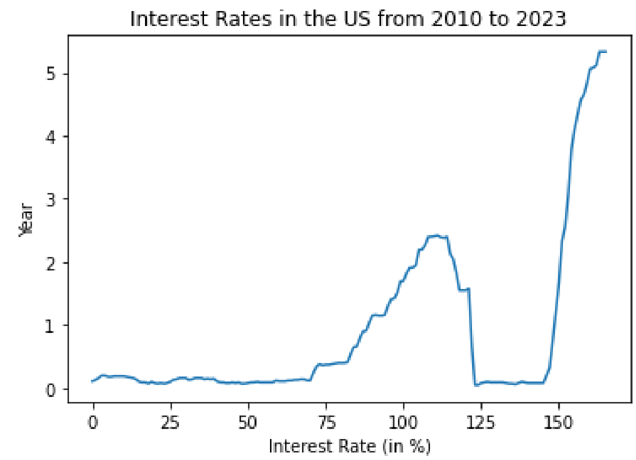
Through this project, we gained a deeper understanding of stock market volatility and its driving factors. Our findings underscore the importance of considering both internal company dynamics and external economic indicators in predicting stock volatility. The application of both standard deviation and GARCH models has demonstrated the complex nature of the stock market, where volatility is influenced by a myriad of factors.

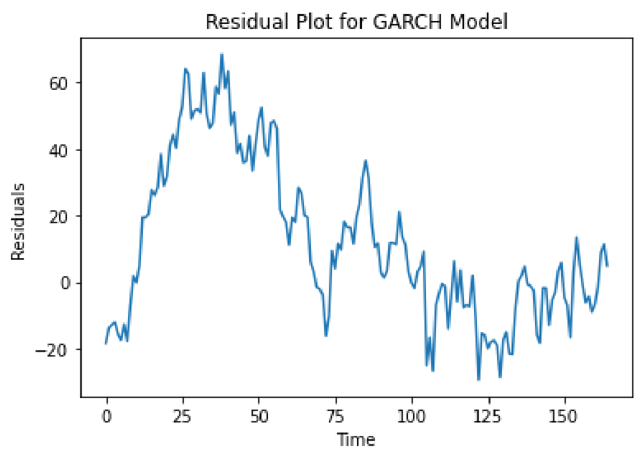
The significance of this study extends to various stakeholders in the financial market. Investors, both individual and institutional, can leverage these insights for better risk management and investment decision-making. Additionally, the findings contribute to the broader field of financial analysis, offering a foundation for future research in stock market behavior and volatility prediction.

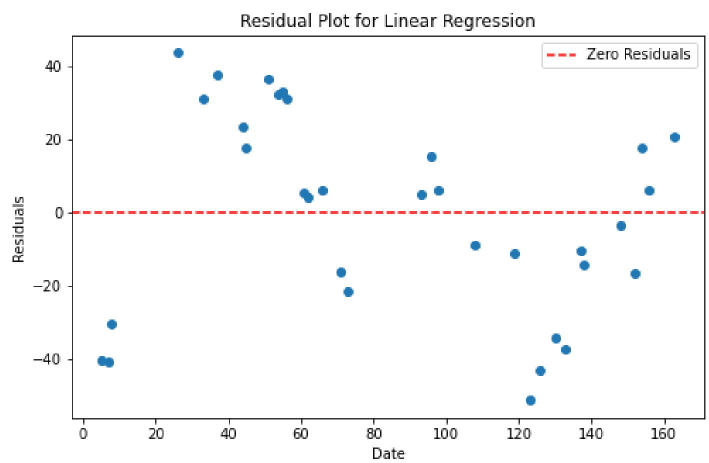
Ultimately, our project underscores the dynamic and interconnected nature of financial markets, where understanding and predicting volatility is key to successful investment and economic analysis. The integration of advanced statistical models with economic data has proven to be a valuable approach in this endeavor.

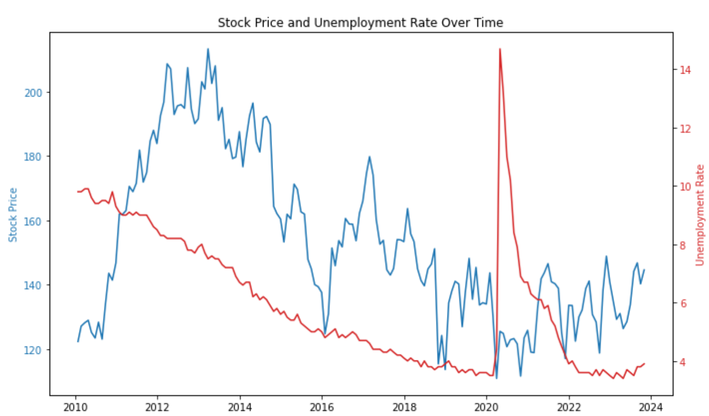
Visualization and Graphs

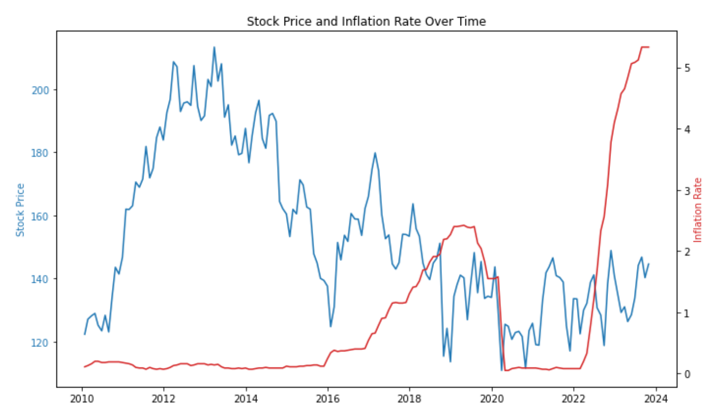












Future Work

The results of our project on predicting stock market volatility using the GARCH model and standard deviation approach open several avenues for future exploration. Here are potential next steps to further delve into this topic:

Expanding Economic Indicators: Our current model primarily incorporates unemployment and interest rates. Future work could include a broader range of economic indicators such as GDP growth rate, consumer price index (CPI), political stability indices, and global economic events. This expansion would allow for a more holistic view of the factors influencing stock volatility.

Global Market Analysis: Extending the analysis beyond the U.S. stock market to include global markets could provide insights into how international events and economic policies impact volatility. This could be particularly relevant in understanding the interconnectedness of global financial markets.

Bibliography

“Calculating Volatility: A Simplified Approach.” *Investopedia*, 2023, www.investopedia.com/articles/basics/09/simplified-measuring-interpreting-volatility.asp. Accessed 3 Dec. 2023.

“Federal Funds Effective Rate.” *Stlouisfed.org*, 2023, fred.stlouisfed.org/series/FEDFUNDS. Accessed 3 Dec. 2023.

Khuong. “How to Predict Stock Volatility Using the GARCH Model in Python.” *Medium*, DataDrivenInvestor, 17 Oct. 2022, medium.datadriveninvestor.com/how-to-predict-stock-volatility-using-garch-model-in-python-df5ba46bae35. Accessed 3 Dec. 2023.

“Unemployment Rate.” *Stlouisfed.org*, 2023, fred.stlouisfed.org/series/UNRATE. Accessed 3 Dec. 2023.

“Volatility: Meaning in Finance and How It Works with Stocks.” *Investopedia*, 2023, www.investopedia.com/terms/v/volatility.asp. Accessed 3 Dec. 2023.

“Yfinance.” *PyPI*, 18 Nov. 2023, pypi.org/project/yfinance/. Accessed 3 Dec. 2023.

Note: *We uploaded a full scale version of the graphs to avoid adding too many pages to this document.*