# CAUSAL INFERENCE IN FINANCIAL MARKETS: A GENERATIVE AI-POWERED WEB APPLICATION FOR ANALYZING MACROECONOMIC INDICATORS AND STOCK MARKET DATA IN THE US

Zixuan Zhou<sup>1</sup>, Carlos Gonzalez<sup>2</sup>

# <sup>1</sup> Brandeis University, 415 South St, Waltham, MA 02453 <sup>2</sup> Computer Science Department, California State Polytechnic University, Pomona, CA 91768

#### ABSTRACT

We aim to address the challenge of understanding stock market behavior during economic uncertainty by integrating S&P 500 stock data with macroeconomic indicators and leveraging Generative AI [6]. Traditional approaches often focus on technical or fundamental analysis, but few incorporate real-time data or AI-driven insights. Our solution combines Python-based data analysis, OpenAI's API, and interactive visualizations to create a user-friendly platform for exploring stock trends and generating financial models [7]. Key challenges included ensuring AI accuracy and balancing functionality with simplicity, which we addressed through feedback mechanisms and modular design. The platform allows users to input queries (e.g., "Tesla stock performance") and receive real-time insights, including text-based responses and visualizations. Preliminary results highlight the platform's potential to democratize access to financial knowledge and improve decision-making. By combining modern AI technologies with traditional financial analysis, our project offers a versatile tool for investors, researchers, and policymakers, making it a valuable resource for navigating complex financial markets [8].

#### **KEYWORDS**

Web Application, Generative AI, Macroeconomic Indicators, Causal Inference, Financial Data

### **1. INTRODUCTION**

The financial markets are complex systems influenced by a myriad of factors, including macroeconomic indicators, investor sentiment, and global economic events. Understanding the relationship between these factors and stock prices is critical for investors, policymakers, and researchers. However, the interplay between macroeconomic variables (e.g., inflation, GDP, unemployment rates) and stock market performance remains poorly understood, especially during periods of economic instability such as the COVID-19 pandemic. For instance, during the 2020 economic shutdown, the S&P 500 experienced unprecedented volatility, dropping by 34% in March 2020 before recovering sharply (Yardeni & Johnson, 2020). This volatility highlighted the need for a deeper understanding of how macroeconomic factors influence stock prices.

David C. Wyld et al. (Eds): ADCO, CMCA, NeCoM, NLPI, CSITEC, BDBS, DMML, SP – 2025 pp. 167-177, 2025. CS & IT - CSCP 2025 DOI: 10.5121/csit.2025.150514

168

The problem lies in the lack of comprehensive tools that integrate both macroeconomic indicators and stock market data to provide actionable insights. Traditional approaches often focus on either technical analysis (e.g., price trends) or fundamental analysis (e.g., company financials), but few studies incorporate macroeconomic variables or leverage modern technologies like Generative AI to enhance accessibility and usability (Fama & French, 1992). This gap limits the ability of investors and researchers to make informed decisions, particularly during economic downturns.

This problem is important because it affects a wide range of stakeholders, including individual investors, institutional investors, and policymakers. For example, individual investors rely on accurate and accessible tools to make informed decisions, while researchers need robust platforms to explore complex financial relationships. Studies have shown that global stock markets, including the S&P 500, are significantly influenced by pandemic-related uncertainty, inflation, and investor sentiment, creating feedback loops that amplify volatility (Dash & Maitra, 2022; Dong et al., 2022). In the long run, solving this problem could lead to more stable financial markets and better-informed economic policies. By integrating Generative AI into financial analysis, we can democratize access to financial insights and empower users to explore market data in real time.

Dash and Maitra (2022) focus on understanding the impact of COVID-19 uncertainty on global equity markets using time-frequency co-movement analysis. While their approach effectively captures interactions among macroeconomic variables, it lacks real-time analysis and AI integration. Our project improves on this by incorporating Generative AI to provide dynamic, user-driven insights and real-time data analysis.

Dong et al. (2022) use a TVP-VAR model to analyze sectoral volatility connectedness during the pandemic. Their methodology provides valuable sector-specific insights but lacks user-friendly tools and real-time capabilities. Our platform addresses this by offering interactive visualizations and AI-generated financial models, making it more accessible and actionable.

Uddin et al. (2021) explore the role of economic resilience in mitigating pandemic-induced stock market volatility. While their study offers a broad perspective, it overlooks individual stock performance and real-time analysis. Our solution integrates real-time data and AI tools to provide tailored insights, improving adaptability and user engagement.

Our proposed solution is to develop an interactive platform that combines S&P 500 stock data with macroeconomic indicators and leverages Generative AI to provide real-time financial insights and analysis [9]. This solution addresses the problem by creating a user-friendly interface where users can explore stock market trends, generate financial models, and receive AI-driven insights tailored to their queries.

Our approach is effective because it integrates modern AI technologies with traditional financial analysis, making complex data accessible to a broader audience [10]. For example, we use OpenAI's API to power a chatbot feature that allows users to ask questions about specific companies, stock performance, or macroeconomic trends. The AI provides text-based responses for general inquiries and visualizations (e.g., moving averages, stock performance charts) for more detailed analysis. This dual functionality enhances user engagement and makes financial research more intuitive.

This solution is better than traditional methods because it bridges the gap between technical financial analysis and user accessibility. While traditional tools often require specialized knowledge to interpret data, our platform simplifies the process by leveraging Generative AI to guide users through their analyses. For instance, users can input queries like "Tesla stock

performance" or "Apple company overview" and receive instant, actionable insights. Additionally, the platform's ability to generate visualizations and downloadable models (e.g., discounted cash flow models) makes it a versatile tool for both casual investors and professionals. Studies have shown that integrating macroeconomic variables and investor sentiment into financial models can improve predictive accuracy, especially during periods of high uncertainty (Dash & Maitra, 2022; Uddin et al., 2021). By combining financial data with AI, we create a dynamic and user-centric solution that adapts to the needs of its users.

The experiment aimed to test the causal relationship between macroeconomic indicators (IORB and UNEMPLOYMENT) and the S&P 500 stock price. Using historical data from sources like FRED and Yahoo Finance, the experiment was set up to preprocess the data (e.g., differencing for stationarity, checking for multicollinearity) and analyze it using OLS and robust regression models. The most significant finding was that IORB (Interest on Reserve Balances) had a strong positive effect on the S&P 500 stock price (p-value = 0.001), suggesting that higher interest rates on reserve balances are associated with higher stock prices, likely due to increased investor confidence. In contrast, UNEMPLOYMENT had a marginally significant negative effect (p-value = 0.063), indicating a weak relationship. The OLS model explained 36.8% of the variance in stock prices, while the robust regression improved the significance of IORB. The results suggest that monetary policy (IORB) plays a crucial role in influencing stock prices, while the effect of unemployment remains less clear, possibly due to confounding factors or limited data.

# **2.** CHALLENGES

In order to build the project, a few challenges have been identified as follows.

#### 2.1. Ensuring the Accuracy and Relevance

A critical component of our solution is the integration of Generative AI to handle user queries and generate insights. A major challenge here is ensuring the accuracy and relevance of AIgenerated responses, especially when dealing with complex financial data. Research has shown that investor sentiment and macroeconomic variables, such as inflation and GDP, significantly influence stock market volatility, and these factors must be accurately reflected in AI responses (Dash & Maitra, 2022; Dong et al., 2022) [11]. To address this, we could implement a feedback mechanism where users can rate the quality of AI responses, allowing the system to learn and improve over time. Additionally, we could fine-tune the AI model with domain-specific financial data to enhance its understanding of stock market concepts and terminology. This would ensure that the AI provides reliable and actionable insights tailored to user needs.

#### 2.2. Designing a User Interface

Another challenge is designing a user interface that balances functionality and simplicity. Users with varying levels of financial expertise must be able to navigate the platform effortlessly while accessing advanced features. Studies have shown that during periods of economic uncertainty, such as the COVID-19 pandemic, investors seek tools that are both intuitive and comprehensive (Uddin et al., 2021). To address this, we could adopt a modular design, offering basic functionalities (e.g., stock performance visualizations) on the main interface and advanced tools (e.g., financial model generation) in a separate section. Additionally, we could incorporate tooltips, tutorials, and an AI-guided onboarding process to help users familiarize themselves with the platform. This approach would ensure that the platform is accessible to beginners while still meeting the needs of advanced users.

### 2.3. Ensuring the Models' Accuracy

A major component of our program is the ability to generate financial models (e.g., discounted cash flow models) using Generative AI. A potential challenge is ensuring that these models are accurate and customizable to user inputs. Research has demonstrated that macroeconomic variables, such as GDP and inflation, play a critical role in shaping stock market performance, and these factors must be incorporated into financial models to ensure accuracy (Dash & Maitra, 2022; Dong et al., 2022). To address this, we could implement a template-based system where users can select from pre-built models and customize parameters (e.g., timeframes, growth rates). The AI would then generate the model based on these inputs and provide downloadable outputs in formats like Excel or CSV. Additionally, we could include validation checks to ensure that user inputs are within reasonable ranges, reducing the risk of errors in model outputs.

# **3. SOLUTION**

Our tool is designed to act as an interactive platform, leveraging API's to provide users with realtime financial insights into various different company performance metrics.

The system begins with a user input query, where users can ask questions about specific companies, industries, or sectors. The system firstly extracts the company information from the query using OpenAI's chat completion API.

Once the company is identified, the system collects inferential data, such as the company overview, news and sentiment data, as well as financial stock data, from the Alpha Vantage API [12]. This aims to provide more context on market perception and trends in data.

This retrieved data is then compiled into a comprehensive financial overview, including various different company performance metrics, news sentiment analysis, and integrates these components into a unified response, combining textual-based insights with visualizations.

The user then receives the response, presenting in an intuitive format for investors and researchers to explore financial market trends more dynamically and efficiently.

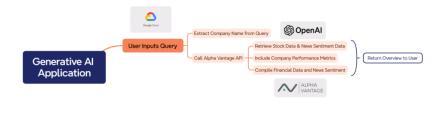


Figure 1. Overview of the solution

The system begins with the user typing in a query into a text field, and once submitted, is sent to the backend server to begin preprocessing and inference. Since our system expects a large variety of questions to be asked, we streamline this task using OpenAI to filter messages based on the desired instruction. For example, if the user would solely like textual-based responses in their output, we consider a different preprocessing approach, compared to other tasks such as diagram generation.

Decision Generator Enter your query below to generate Al-powered insights.	
Enter your query here	Generate Response
Generated Graph	

Figure 2. Decision generator



Figure 3. Screenshot of code 1

In this example, we can see an OpenAI chatcompletion object being created to filter based on desired outcome. We primarily filter with the condition that the response be textual or visual. If the user aims to receive a textual-based response, we utilize different endpoints and preprocessing strategies to compile a formalized output, which is different in our diagram-visualizing approach. Depending on which function is being used will determine the structure of the output, which is vital in determining the best needs for the user based on their initial query. This section provides financial data, especially stock data, by calling the Alpha Vantage API endpoint using Python. The data is collected, preprocessed, and formatted for user analysis. Using Matplotlib, the processed data is visualized as a graph, displaying key stock metrics like open, high, low, close prices, and trading volume for better financial insights [13]. Users can instantly get insights into a company's stock performance through the trends, volatility, and trading patterns using this function.

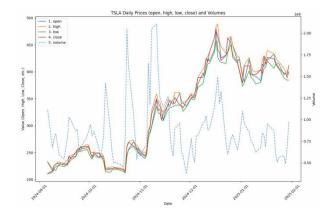


Figure 4. TSLA daily prices and volumes



Figure 5. Screenshot of code 2

The json object received from the Alpha Vantage endpoint remains consistent among various different requests, so we aim to streamline this process by calling matplotlib and generating a graph based on the numerical data presented to us. This involves scaling the 'volume' prices to the 'open,' 'close,' 'high,' and 'low' prices, as well as presenting all required information in the graph. The users can get a brief overview of the performance of a company, based on their specific request. For example, if the user would like an annual snapshot on the performance of the company Tesla, the system can filter based on the specifications of the query.

Data collection for monitoring performance is vital in this process, as we're continually aiming to improve the results presented to users [14]. As such, we compile and store financial stock visualization data into a firebase database, which is synced with our client, to continue monitoring the strength of the data visualization task.

Name	Size	Туре
IiZD0pQdoVFi.png		image/png
3Hcv8gE3MtiH.png		image/png
4dR5zVDkvfnr.png		image/png
ATHwcjMh0Uzy.png		image/png
FxPOWZU8Guwa.png		image/png
MEmfY46tbGGW.png		image/png
IluukiRa3FQr.png		image/png
UUFSC73jbLEr.png		image/png
V01NvIOFrY87.png		image/png

Figure 6. Screenshot of the images



Figure 7. Screenshot of code 3

To properly anonymize the diagram data, we firstly create an encoded name for each of the image files. We then store the relative paths for the images created in the server's instance, initialize the

firebase storage bucket, and load & save the image to the provided bucket. We run this process upon each instance to ensure that the data is being properly saved and stored for future use. We aim to further categorize these graphs by assigning labels and creating a structured metadata file to further compartmentalize and analyze the data more deeply.

# 4. EXPERIMENT

A possible blind spot in the program is the causal relationship between macroeconomic indicators (e.g., IORB and UNEMPLOYMENT) and the S&P 500 stock price. It is important to test this because understanding how these variables influence stock prices can help investors and policymakers make informed decisions. If the program fails to accurately capture these relationships, it could lead to incorrect conclusions and poor decision-making.

The experiment is designed to test the causal relationship between macroeconomic indicators (IORB and UNEMPLOYMENT) and the S&P 500 stock price using causal inference techniques.

The experiment is set up as follows:

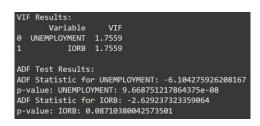
Data Collection: Historical data for the S&P 500 stock price, IORB (Interest on Reserve Balances), and UNEMPLOYMENT rates are collected from reliable sources such as the Federal Reserve Economic Data (FRED) and Yahoo Finance.

Data Preprocessing: The data is cleaned, and non-stationary variables are different to ensure stationarity. Multicollinearity is checked using Variance Inflation Factor (VIF). We propose the following models:

Ordinary Least Squares (OLS) Regression: To estimate the linear relationship between the variables [15].

Robust Regression: To handle outliers and improve the robustness of the results.

The control data is sourced from the same datasets, ensuring consistency and reliability. The experiment is set up this way to isolate the effects of IORB and UNEMPLOYMENT on the S&P 500 stock price while controlling for other factors. By using both OLS and robust regression, the experiment ensures that the results are reliable and not skewed by outliers or multicollinearity.



Computer Science & Information Technology (CS & IT)

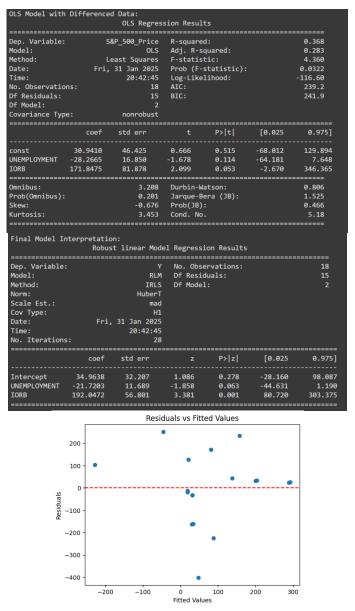


Figure 8. Figure of the experiment

The analysis of the experiment reveals the following key insights via causal inference: The robust regression model shows that IORB has a significant positive effect on the S&P 500 stock price (coefficient = 192.0472, p-value = 0.001). This suggests that higher interest rates on reserve balances are associated with higher stock prices, possibly due to increased investor confidence. The effect of UNEMPLOYMENT is marginally significant in the robust regression model (coefficient = -21.7203, p-value = 0.063). This indicates that higher unemployment rates may have a slight negative impact on stock prices, though the effect is not strong. The OLS model has an R-squared value of 0.368, meaning it explains 36.8% of the variance in the S&P 500 stock price. While this is moderate, it suggests that other factors not included in the model may also influence stock prices.

The robust regression model improves the significance of IORB, indicating that it is better at handling outliers and providing reliable estimates. The weak effect of UNEMPLOYMENT was

surprising, as higher unemployment is typically associated with economic downturns and lower stock prices. This could be due to the limited sample size or the specific time period analyzed. The strong effect of IORB was unexpected, as higher interest rates are often thought to reduce stock prices. This result may reflect unique economic conditions during the study period. The biggest effect on the results comes from IORB, which is highly significant in the robust regression model. This suggests that monetary policy (as reflected in IORB) plays a crucial role in influencing stock prices. The limited effect of UNEMPLOYMENT may be due to its indirect relationship with stock prices or the presence of confounding variables not included in the model. As such, we can see that the experiment highlights the importance of IORB as a driver of stock prices, while the effect of UNEMPLOYMENT remains less clear. Further research with additional variables and a larger dataset can seek to better understand these relationships.

### **5. RELATED WORK**

One scholarly source that addresses the problem of understanding stock market behavior during economic uncertainty is the study by Dash and Maitra (2022), which examines the impact of COVID-19 pandemic uncertainty on global equity markets [1]. Their methodology uses time-frequency co-movement analysis to explore interactions among GDP, CPI, investor sentiment, and market liquidity. While this approach effectively captures the interconnectedness of macroeconomic variables and stock market volatility, it has limitations. For instance, it relies heavily on historical data and does not incorporate real-time analysis or advanced AI tools. Additionally, the study focuses on G7 countries, which may limit its applicability to other markets.

Our project improves on this by integrating real-time data analysis and Generative AI, allowing users to explore stock market trends dynamically. Unlike Dash and Maitra's static approach, our platform provides interactive tools for users to generate financial models and receive AI-driven insights tailored to their queries. This makes our solution more accessible and actionable for a broader audience.

Another relevant study is by Dong et al. (2022), which investigates the impact of COVID-19 on global stock sectors using a time-varying parameter vector autoregressive (TVP-VAR) model [2]. Their methodology focuses on sectoral volatility connectedness and asymmetric responses to pandemic intensity. While this approach provides valuable insights into sector-specific trends, it has limitations. For example, it does not incorporate user-friendly tools for real-time analysis or leverage modern technologies like Generative AI. Additionally, the study's reliance on quantile-on-quantile regression (QQR) may limit its ability to capture complex, non-linear relationships. Our project addresses these limitations by combining traditional econometric methods with AI-driven tools. Unlike Dong et al.'s sector-specific focus, our platform offers a holistic view of the stock market, integrating macroeconomic indicators and real-time data. By leveraging Generative AI, we provide users with interactive visualizations and customizable financial models, making our solution more versatile and user-centric.

A third study by Uddin et al. (2021) explores the effect of COVID-19 on global stock market volatility and examines whether economic resilience factors (e.g., financial development, corporate governance) can mitigate pandemic-induced uncertainty [3]. Their methodology uses panel data analysis across 34 countries, providing a broad perspective on stock market behavior. However, this approach has limitations. It does not incorporate real-time data analysis or advanced AI tools, and its focus on country-level economic resilience may overlook individual stock performance.

Our project improves on this by integrating real-time stock data and macroeconomic indicators into a unified platform. Unlike Uddin et al.'s static analysis, our solution leverages Generative AI to provide dynamic, user-driven insights. For example, users can generate financial models (e.g., discounted cash flow) and receive AI-generated responses to specific queries. This makes our platform more adaptable and actionable for both individual investors and researchers.

# **6.** CONCLUSIONS

While our project successfully integrates S&P 500 stock data with macroeconomic indicators and leverages Generative AI to provide real-time financial insights, it has several limitations. First, the accuracy of AI-generated responses depends on the quality and scope of the training data. For instance, the AI may struggle with highly specific or niche queries due to limited domain-specific fine-tuning [4]. Second, the platform's reliance on APIs for real-time data introduces potential latency and data availability issues, which could affect user experience. Third, the current version of the platform does not fully account for confounding variables, such as geopolitical events or sector-specific shocks, which could influence stock market performance.

To address these limitations, we propose several improvements. First, we could expand the training dataset for the AI model to include more diverse financial data, such as earnings reports and industry-specific metrics. Second, we could implement caching mechanisms to reduce API latency and ensure smoother data retrieval [5]. These enhancements would make the platform more robust, reliable, and adaptable to changing market conditions.

We aim to bridge the gap between traditional financial analysis and modern AI technologies, offering a user-friendly platform for exploring stock market trends and generating actionable insights. The current implementation shows the strengths of integrating Generative AI and real-time data analysis, a significant step forward in democratizing access to financial knowledge.

## REFERENCES

- [1] Dash, Saumya Ranjan, and Debasish Maitra. "The COVID-19 pandemic uncertainty, investor sentiment, and global equity markets: Evidence from the time-frequency co-movements." The North American Journal of Economics and Finance 62 (2022): 101712.
- [2] Dong, Zibing, et al. "Impacts of COVID-19 on global stock sectors: Evidence from time-varying connectedness and asymmetric nexus analysis." The North American Journal of Economics and Finance 62 (2022): 101753.
- [3] Uddin, Moshfique, et al. "The effect of COVID–19 pandemic on global stock market volatility: Can economic strength help to manage the uncertainty?." Journal of Business Research 128 (2021): 31.
- [4] Gushendri, Adzra Helga Engrasia, and IrniYunita. "Determinants of Stock Liquidity during COVID-19." International Journal of Science, Technology & Management 5.6 (2024): 1298-1303.
- [5] Fama, Eugene F., and Kenneth R. French. "The cross-section of expected stock returns." the Journal of Finance 47.2 (1992): 427-465.
- [6] Su, Zhi, Tong Fang, and Libo Yin. "Understanding stock market volatility: What is the role of US uncertainty?." The North American Journal of Economics and Finance 48 (2019): 582-590.
- [7] Sauter, Nicholas K., et al. "New Python-based methods for data processing." Acta Crystallographica Section D: Biological Crystallography 69.7 (2013): 1274-1282.
- [8] Kaufmann, Roger, Andreas Gadmer, and Ralf Klett. "Introduction to dynamic financial analysis." ASTIN Bulletin: The Journal of the IAA 31.1 (2001): 213-249.
- [9] Feuerriegel, Stefan, et al. "Generative ai." Business & Information Systems Engineering 66.1 (2024): 111-126.
- [10] Burlakov, Vyacheslav V., et al. "The modern trends of development of AI technologies." Artificial Intelligence: Anthropogenic Nature vs. Social Origin (2020): 374-383.
- [11] Landefeld, J. Steven, Eugene P. Seskin, and Barbara M. Fraumeni. "Taking the pulse of the economy: Measuring GDP." Journal of Economic Perspectives 22.2 (2008): 193-216.

- [12] Raut, Kunal, et al. "Stock Market Prediction using Alpha Vantage API and Machine Learning Algorithm." (2022).
- [13] Bisong, Ekaba, and EkabaBisong. "Matplotlib and seaborn." Building machine learning and deep learning models on google cloud platform: A comprehensive guide for beginners (2019): 151-165.
- [14] Willson, Stephanie, and Kristen Miller. "Data collection." Cognitive interviewing methodology (2014): 15-33.
- [15] Craven, B. D., and Sardar MN Islam. "Ordinary least-squares regression." The SAGE dictionary of quantitative management research 1 (2011): 224-228.

© 2025 By AIRCC Publishing Corporation. This article is published under the Creative Common Attribution (CC BY) license.