

# Predicting Post-Election FAANG Stock Prices: Accurate November Trends and Flat December Projections\*

An XGBoost Model Analysis of Market Behavior and Real-World Implications for  
Meta, Amazon, Apple, Netflix, and Google

Jimin (Jamie) Lee

December 3, 2024

This paper presents a predictive model for FAANG stock prices (Meta, Amazon, Apple, Netflix, and Google) during November and December 2024. Using historical stock price data from Yahoo Finance and an XGBoost regression framework, we estimate adjusted stock prices based on features such as lagged prices, moving averages, and volatility. Our findings indicate strong predictive accuracy for short-term trends in November, with predictions closely aligning with actual prices for stocks like AAPL and GOOGL. However, discrepancies, such as the underestimation of NFLX prices, highlight the impact of unique stock-specific dynamics. Predictions for December show a flat trajectory, potentially influenced by data limitations and post-election market sentiment. This analysis provides valuable insights into FAANG stock behavior during critical periods, equipping investors and policymakers to better understand market responses to external events like elections.

## Table of contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Data</b>	<b>2</b>
2.1	Data Overview . . . . .	2
2.2	Measurement . . . . .	3

---

\*Code and data are available at: <https://github.com/jamiejiminlee/FAANG-Stock-Forecast.git>.

2.3	Variables . . . . .	4
2.3.1	Original Variables . . . . .	4
2.3.2	Constructed Variables . . . . .	4
2.3.3	Outcome Variable . . . . .	5
2.3.4	Predictor Variables . . . . .	5
<b>3</b>	<b>Model</b>	<b>6</b>
3.1	Model Overview . . . . .	6
3.2	Model Assumptions . . . . .	7
3.3	Model Setup . . . . .	8
3.3.1	November Model . . . . .	9
3.3.2	December Model . . . . .	10
3.4	Model Justification . . . . .	10
<b>4</b>	<b>Results</b>	<b>11</b>
4.1	Actual vs Predicted for November 2024 . . . . .	11
4.2	Heatmap of Prediction Errors (November 2024) . . . . .	12
4.3	Average Predicted Stock Price - November . . . . .	12
4.4	Predicted Stock Price Trends - December . . . . .	14
4.5	Average Stock Prices Bar Chart - December . . . . .	15
4.6	Average Predicted Stock Price - December . . . . .	15
4.7	Combined November and December . . . . .	16
4.8	Average Predicted Stock Prices for November and December 2024 . . . . .	16
<b>5</b>	<b>Discussion</b>	<b>18</b>
5.1	Summary of Findings . . . . .	18
5.2	Model Performance and Limitations . . . . .	18
5.2.1	November Model . . . . .	18
5.2.2	December Model . . . . .	19
5.3	US Election Effects . . . . .	19
	<b>Appendix</b>	<b>21</b>
<b>A</b>	<b>Additional Data Details</b>	<b>21</b>
A.1	Data Cleaning . . . . .	21
A.2	Data Limitations and Future Directions . . . . .	21
<b>B</b>	<b>Additional Model Details</b>	<b>22</b>
B.1	Model Justification - Further Details . . . . .	22
B.1.1	Why Separate November and December Models? . . . . .	22
B.1.2	Model Justification and Features . . . . .	22
B.1.3	Hyperparameters and Regularization . . . . .	23
B.2	Model Diagnostics . . . . .	23
B.2.1	Feature Importance . . . . .	23

B.2.2	Residual Analysis . . . . .	24
B.2.3	Error Distribution . . . . .	24
B.2.4	Prediction vs. Actual Values . . . . .	26
B.3	Model Performance and Limitations . . . . .	26
B.3.1	Model Performance . . . . .	26
B.3.2	Model Limitations . . . . .	27
<b>C</b>	<b>Surveys, Sampling, and Observational Data</b>	<b>28</b>
C.1	Nature of Observational Data . . . . .	28
C.2	Sampling Frame and Representativeness . . . . .	28
C.2.1	Strengths . . . . .	29
C.2.2	Limitations . . . . .	29
C.3	Integrating Surveys: A Complementary Perspective . . . . .	29
C.3.1	Proposed Survey Design . . . . .	29
C.3.2	Methodological Considerations . . . . .	29
C.3.3	Strengths and Limitations . . . . .	30
C.4	Parallels Between Observational Data and Surveys . . . . .	30
C.5	Addressing Data Challenges . . . . .	30
C.6	Recommendations for Future Work . . . . .	31
<b>D</b>	<b>References</b>	<b>32</b>

## 1 Introduction

The performance of FAANG stocks - **Meta (META)**, **Amazon (AMZN)**, **Apple (AAPL)**, **Netflix (NFLX)**, and **Google (GOOGL)**, has become a cornerstone of global financial markets, reflecting broader economic trends and investor sentiment. These companies represent diverse sectors, from e-commerce and technology to entertainment and digital advertising, making their stock price movements critical indicators of market health. In the context of the 2024 post-pandemic recovery and a pivotal United States Presidential Election, understanding and predicting FAANG stock price trends is particularly relevant for investors and policymakers alike.

This paper leverages historical stock price data sourced from Yahoo Finance (2020–2024) to develop predictive models for FAANG stock prices in November and December 2024. Utilizing XGBoost regression models, we analyze historical trends and engineered features, including lagged prices, moving averages, and volatility measures, to forecast stock price movements. These predictions provide valuable insights into the behavior of FAANG stocks during critical periods, such as post-election market adjustments, and highlight opportunities and risks for short-term and long-term investment strategies.

The primary estimand of this study is the predicted adjusted stock price for each FAANG company during November and December 2024, derived from features constructed using his-

torical price trends and technical indicators. Our findings reveal that predictions for November align closely with observed trends for most stocks, reflecting the model’s ability to capture short-term dynamics. However, discrepancies, such as the underestimation of NFLX prices, underscore the influence of unique stock-specific factors. For December, predictions indicate a flat trajectory across stocks, likely influenced by data limitations and post-election market sentiment.

The implications of these findings extend beyond price predictions, offering a framework to evaluate market behavior in response to external factors such as elections and economic conditions. The remainder of this paper is organized as follows: Section 2 describes the data sources, variables, and preprocessing steps; Section 3 outlines the modeling approach and feature engineering strategies; Section 4 presents the results and discusses trends in predicted stock prices; and Section 5 addresses the implications and limitations of the findings, with additional model diagnostics detailed in the Appendix [B](#).

## 2 Data

### 2.1 Data Overview

This analysis uses historical stock price data for FAANG companies (Meta, Amazon, Apple, Netflix, and Google) obtained from the Yahoo Finance API via the `tidyquant` library in R on November 30, 2024. The dataset spans from January 1, 2020, to December 31, 2024, providing daily records of stock performance, including key metrics such as opening, high, low, and closing prices, trading volume, and adjusted closing prices normalized for corporate actions. To enhance the predictive power of the data, several derived features were constructed. These include lagged adjusted prices to capture sequential trends, daily returns to measure proportional changes, rolling averages to smooth short- and long-term fluctuations, and a rolling standard deviation of prices to quantify market volatility. These features offer a structured framework for identifying patterns and forecasting future stock movements.

The raw data underwent pre-processing to address missing values, particularly for adjusted prices. Missing entries were filled using forward and backward imputation to ensure data continuity. Placeholder dates for December 2024 were interpolated to prepare for out-of-sample predictions. The cleaned data, prepared using the `tidyverse` of (`tidyverse?`), `lubridate` of (`lubridate?`), `TTR` of (`ttr?`), and `arrow` of (`arrow?`) libraries, provides a consistent and comprehensive foundation for modeling and analysis. Further details on the data preparation process are included in Appendix [A](#).

### 2.2 Measurement

Stock price data encapsulates market dynamics influenced by diverse factors, including macroeconomic conditions, corporate events, and investor sentiment. The raw variables in this

dataset, such as **open**, **high**, **low**, **close**, and **adjusted**, capture key price points and provide a snapshot of market activity for each trading day. However, while these variables accurately reflect daily trading outcomes, they may not fully represent intraday fluctuations or complex trading dynamics.

Constructed variables enhance the analytical depth of the dataset by quantifying sequential trends, price variability, and momentum. For instance, **Lag\_1** captures the adjusted closing price from the previous trading day, providing context for sequential patterns. Variables like **daily\_return** measure proportional daily changes in price, while moving averages such as **Rolling\_Mean\_7**, **sma\_20**, and **sma\_50** smooth short- and long-term price fluctuations, offering a clearer view of overall trends. Additionally, **volatility**, computed as the rolling standard deviation of closing prices, quantifies recent price variability and reflects market uncertainty.

Market activity is further contextualized through **volume**, representing the total shares traded daily, a proxy for investor interest and liquidity. The inclusion of **symbol\_encoded**, a numerical encoding of stock tickers, standardizes categorical data for compatibility with machine learning models. Together, these features create a robust framework for analysis, balancing raw price data with derived metrics that highlight market trends and dynamics.

Despite the strengths of these variables, limitations exist. Raw price points do not account for intraday variations, while constructed metrics may oversimplify complex market behaviors. External shocks, such as economic crises or regulatory changes, may not be fully captured within the variables' scope. Acknowledging these constraints is essential for interpreting the results and understanding the broader context of the analysis. These limitations are further discussed in [Appendix A](#).

## 2.3 Variables

The dataset used for this analysis includes both original and constructed variables, designed to capture temporal trends, volatility, and momentum in stock price movements. Below is a detailed overview of the variables included in the model:

### 2.3.1 Original Variables

The original variables directly obtained from the raw stock price data are:

- **symbol**: The stock ticker symbol identifying the company (e.g., AAPL for Apple, AMZN for Amazon).
- **date**: The trading date, essential for analyzing time-series trends.
- **open**: The stock's opening price for the trading day, indicating initial market conditions.
- **high**: The highest price reached during the trading day, reflecting intraday volatility.
- **low**: The lowest price during the trading day, reflecting downward market trends.

- **close:** The stock's unadjusted closing price, representing the final value at the end of trading.
- **volume:** The total number of shares traded during the day, reflecting market activity.
- **adjusted:** The closing price adjusted for corporate actions such as splits or dividends, ensuring consistency for trend analysis.

### 2.3.2 Constructed Variables

The constructed variables were engineered to enhance the model's predictive performance by capturing price trends, volatility, and momentum:

- **Lag\_1:** The closing price from the previous trading day, derived using the `lag()` function. This variable captures sequential patterns in stock prices.
- **Rolling\_Mean\_7:** A 7-day moving average of adjusted closing prices, computed using the `rollmean()` function. This variable smooths short-term fluctuations to highlight weekly trends.
- **sma\_20:** The 20-day simple moving average (SMA) of adjusted prices, reflecting medium-term price trends.
- **sma\_50:** The 50-day simple moving average of adjusted prices, providing insights into longer-term trends.
- **volatility:** The 20-day rolling standard deviation of adjusted prices, capturing short-term price variability.
- **daily\_return:** The percentage change in adjusted closing prices relative to the previous day. This value is calculated by dividing the adjusted closing price by its value from the previous day and subtracting 1. This variable normalizes price changes, making them comparable across time and stocks.
- **symbol\_encoded:** A numeric encoding of the stock ticker symbol, allowing the categorical `symbol` variable to be included in the modeling process.

### 2.3.3 Outcome Variable

The **outcome variable** for this analysis is the **adjusted closing price (adjusted)**, which serves as the target for prediction. This represents the stock's closing price adjusted for corporate actions like stock splits or dividends, providing a normalized measure of stock value. The model aims to predict this variable based on historical data and constructed predictors.

Figure 1 illustrates the distribution of the adjusted closing prices (adjusted) for each FAANG stock, faceted by stock symbol. These distributions highlight the variability in stock prices across the companies, with stocks like NFLX showing higher price ranges, whereas GOOGL and AMZN have narrower distributions. The visual emphasizes the diverse price behavior among FAANG stocks, underscoring the need for tailored models that capture these unique

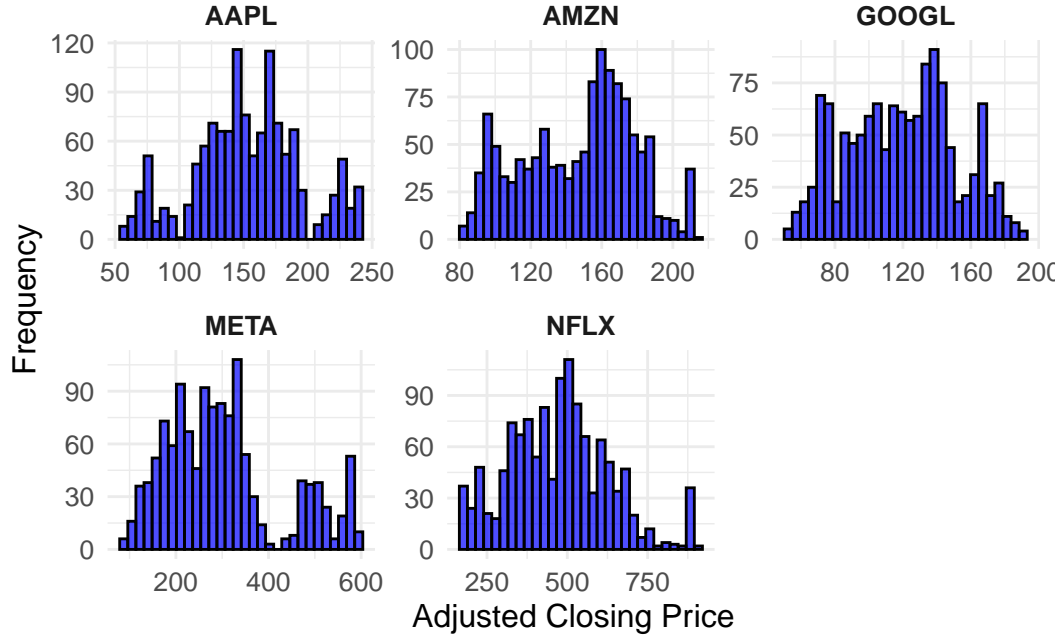


Figure 1: Distribution of Adjusted Closing Prices by Stock

trends. By normalizing for corporate actions like stock splits or dividends, the adjusted closing price serves as a reliable target variable for forecasting and analysis.

### 2.3.4 Predictor Variables

The predictor variables used to model the adjusted closing price are:

- **Lag\_1:** The prior day's adjusted closing price, reflecting momentum.
- **Rolling\_Mean\_7:** A smoothed metric capturing weekly price trends.
- **sma\_20** and **sma\_50:** Simple moving averages representing medium- and long-term price trends.
- **volatility:** The short-term price variability, highlighting potential risks.
- **daily\_return:** The normalized daily percentage price change.
- **symbol\_encoded:** A numerical representation of the stock ticker symbol, capturing company-level differences.

Together, these variables provide a robust framework for understanding and predicting stock price movements over time, leveraging both raw market indicators and derived metrics.

Table 1: Random Sample of Predictive Variables for Modeling Adjusted Closing Price

Lag_1	Rolling Mean (7-Day)	SMA (20-Day)	SMA (50-Day)	Volatility	Daily Return	Symbol (Encoded)
186.33	190.14	185.31	179.13	6.63	-0.01	2
207.89	207.89	205.95	196.42	4.18	0.00	2
125.17	128.87	128.21	133.18	2.43	0.01	2
171.96	163.05	166.76	167.60	6.14	0.00	1
305.92	319.23	331.31	330.76	12.26	0.04	4

Table 1 displays a random selection of five rows from the dataset, showcasing the predictive variables utilized in the modeling process. Each row represents data for a specific stock, including values for lagged adjusted closing prices (Lag\_1), smoothed weekly averages (Rolling Mean (7-Day)), and medium- and long-term moving averages (SMA (20-Day) and SMA (50-Day)). These variables are complemented by measures of price volatility, daily percentage returns (Daily Return), and an encoded numerical identifier for the stock symbol (Symbol (Encoded)). All values are rounded to two decimal places to enhance clarity and precision. This table illustrates the diverse set of historical and technical indicators used to predict future adjusted stock prices.

### 3 Model

#### 3.1 Model Overview

The models in this study forecast the adjusted closing prices of FAANG stocks (Meta, Amazon, Apple, Netflix, and Google) for the months of November and December 2024. Separate models are constructed for each month to capture the time-dependent nature of stock price movements and to account for fluctuations in market conditions leading up to the forecast period. These models utilize a combination of historical stock price data and technical indicators, such as moving averages, volatility, and lagged price features, to estimate future stock prices for each of the five FAANG companies.

To model the changes in stock prices over time, we use features that capture both short-term and long-term market trends. The models incorporate lagged stock prices, 7-day rolling averages, 20-day and 50-day simple moving averages (SMA), daily returns, and volatility indicators. These features are crucial for understanding the patterns and cycles in stock price movements. XGBoost, a powerful machine learning algorithm, is used to build these models, allowing for the capture of non-linear relationships and interactions between the technical indicators and future stock prices.



The training datasets for the November and December models are defined using data up to October 31, 2024, and November 30, 2024, respectively. The models take into account temporal changes in stock prices by using features like lagged prices and rolling averages, which help capture the dynamics of stock price behavior. Additionally, XGBoost’s flexibility allows the models to account for varying levels of volatility and market shifts across the different FAANG stocks.

This modeling approach enables the prediction of future stock prices, considering how past price trends and market indicators evolve. By focusing on the November and December time-frames, the models aim to provide forecasts that reflect recent market conditions and predict likely price movements as we approach the end of 2024. The detailed model specifications, including feature selection and hyperparameter tuning, are further discussed in [Appendix B](#)

## 3.2 Model Assumptions

To ensure the validity and reliability of our stock price prediction models, several key assumptions are made during the modeling process. These assumptions are necessary to capture the complexities of stock price behavior while maintaining a clear and interpretable framework for forecasting:

- **Non-Linearity in Price Movements:** Stock prices often exhibit non-linear patterns due to various market forces, including economic news, company performance, and broader market trends. To account for these non-linearities, we use XGBoost, which allows for the modeling of complex, non-linear relationships between the predictors (such as lagged prices and moving averages) and the target variable (adjusted closing prices). This assumption ensures that the model can effectively capture the changing dynamics of stock price movements over time.
- **Stationarity of Price Features:** The model assumes that the selected technical indicators—such as moving averages, volatility, and daily returns—are stationary over the forecasting period. While stock prices are inherently volatile, we assume that the patterns observed in historical data are stable enough to provide reliable predictions for the future months. This assumption allows us to model stock prices without explicitly accounting for trends or seasonality, assuming that the past behavior is a good representation of future dynamics.
- **Independence of Observations:** Each daily stock price observation is assumed to be independent of the others. While stock prices are influenced by previous values (as captured through lagged features), we assume that each day’s stock price data point does not directly depend on other data points in ways not captured by the lagged variables or technical indicators. This assumption is important for simplifying the modeling process and ensuring that the features can be treated as independent predictors.

- **Stationarity of Model Parameters:** We assume that the relationships between the stock prices and the chosen technical indicators (e.g., moving averages, volatility) remain relatively stable over the forecast horizon. While stock prices can be influenced by many unpredictable factors, this assumption allows the model to generalize well to the forecasting periods (November and December 2024).
- **Feature Importance:** We assume that the selected features—lagged prices, moving averages, volatility, and daily returns—capture the most important aspects of stock price movements for the FAANG companies. While other factors (such as macroeconomic events, news sentiment, or geopolitical developments) may influence stock prices, these features are chosen for their relevance to technical analysis and their ability to capture short-term and long-term market trends.
- **Use of XGBoost’s Flexibility:** The XGBoost algorithm is assumed to be an appropriate method for this analysis due to its ability to handle large datasets, capture non-linear relationships, and deal with complex interactions between features. We assume that its regularization techniques will help prevent overfitting while still capturing the intricate patterns present in the stock price data.

These assumptions form the foundation of the models and guide the interpretation of the results. While they allow for effective forecasting, it is important to acknowledge that stock price prediction is inherently uncertain, and the accuracy of predictions may vary depending on unforeseen market events and fluctuations.

### 3.3 Model Setup

To accommodate differences in data availability and prediction requirements, the modeling process was split into two separate models for November and December. The November model was trained and validated using historical data up to October 31, 2024, enabling predictions that could be compared to actual observed values for November. In contrast, the December model was trained on data up to November 30, 2024, and relied solely on historical trends and engineered features for prediction, as actual December data was unavailable. This separation allowed the analysis to leverage observed data for evaluating short-term accuracy in November while extending the prediction framework to December despite the lack of real data.

Modeling processes are conducted by employing the following packages: `tidyquant` package of (`tidyquant?`) for downloading and managing financial data, `xgboost` package of (`xgboost?`) for implementing the machine learning model, `dplyr` package of (`dplyr?`) for data manipulation, and `lubridate` package of (`lubridate?`) for handling dates and times.

### 3.3.1 November Model

The primary goal of the November model is to predict the adjusted closing prices of FAANG stocks for the month of November 2024 based on historical stock price data. This model is structured as a regression task using the XGBoost algorithm to estimate the future adjusted closing prices for each of the five FAANG companies.

The model setup can be described as follows:

$$\hat{y}_i = f(X_i)$$

Where:

- $\hat{y}_i$  is the predicted adjusted closing price for the stock on day  $i$ .
- $X_i$  represents the feature set for day  $i$ , which includes:
  - **Lag\_1**: The adjusted closing price of the previous day.
  - **Rolling\_Mean\_7**: The 7-day rolling average of the adjusted closing price.
  - **sma\_20**: The 20-day simple moving average.
  - **sma\_50**: The 50-day simple moving average.
  - **volatility**: The 14-day rolling standard deviation of daily returns.
  - **daily\_return**: The daily percentage change in the adjusted closing price.
  - **symbol\_encoded**: A numeric encoding for the FAANG companies (Meta, Amazon, Apple, Netflix, Google).

The dataset for the November model is prepared by extracting the relevant data from January 2020 to October 31, 2024. Features like moving averages and volatility are calculated based on this historical data, and the model is trained to predict the adjusted closing prices for November 2024.

The model was implemented using the **xgboost** package in R, which is optimized for performance in regression tasks:

$$y_i = \text{XGBoost}(X_i, \theta)$$

Where  $y_i$  represents the actual adjusted closing price for day  $i$  and  $\theta$  represents the model parameters to be learned during training. The model is trained with the following key parameters: - **Max Depth**: 6 - **Learning Rate**: 0.1 - **Number of Rounds**: 100 - **Objective Function**: Regression with squared error

### 3.3.2 December Model

The December model follows a similar approach to the November model, but is trained using data up to November 30, 2024. This ensures that the model accounts for the most recent stock price movements leading up to December, providing a more accurate prediction for the final month of 2024.

The setup for the December model mirrors that of the November model, with the same features and target variable, but applied to the updated dataset. The model structure remains the same, and the XGBoost algorithm is again used for regression:

$$\hat{y}_i = f(X_i)$$

Where:

- $\hat{y}_i$  is the predicted adjusted closing price for the stock on day  $i$  in December 2024.
- $X_i$  includes the same features as the November model, but with data up to November 30, 2024.

The December model is trained using the same XGBoost parameters as the November model. Once trained, the model will generate predictions for December 2024, leveraging the latest market conditions reflected in the training data.

The assumptions and model design are consistent with the November model, with the key differences being the updated data and the forecast period. By training separate models for November and December, we can capture any evolving trends or shifts in stock price behavior as we approach the end of 2024.

## 3.4 Model Justification

The XGBoost model is the most suitable approach for forecasting FAANG stock prices in November and December 2024. Our primary outcome variable is the adjusted closing price of each stock, which is continuous and unbounded. XGBoost, a gradient boosting algorithm, excels in handling regression tasks with large datasets, capturing complex, non-linear relationships between features such as lagged prices, moving averages, volatility, and daily returns. To model these relationships effectively, XGBoost's ability to handle feature interactions and non-linearities is key. The inclusion of rolling averages and volatility allows the model to account for both short-term fluctuations and longer-term trends in stock prices, which are often influenced by broader market dynamics. Additionally, the model's flexibility helps capture the intricacies of stock price behavior over time, ensuring that past market patterns are reflected in future predictions.

While simpler models like linear regression could be used, they would fail to capture the complexity and non-linearity inherent in financial data. Machine learning techniques like random

forests or support vector machines could also provide robust predictions, but XGBoost’s combination of accuracy, speed, and interpretability makes it the best choice for this analysis. Ultimately, XGBoost’s ability to learn from the data and adjust to market conditions enhances its suitability for predicting FAANG stock prices in an uncertain and volatile market environment.

## 4 Results

Using the models trained on historical data, we applied the XGBoost algorithm to predict the adjusted closing prices for each of the FAANG stocks during these two months. For November, we filtered the data from November 1st to 30th, 2024, and prepared the relevant features for prediction, including lagged prices, rolling means, and volatility indicators. Since we have the actual adjusted closing prices for November, we compare the predicted values with the actual data to evaluate model performance. For December, the same approach was applied to the data from December 1st to 31st, 2024, but since the actual prices for this month are unavailable, the results for December focus solely on the predicted values.

### 4.1 Actual vs Predicted for November 2024

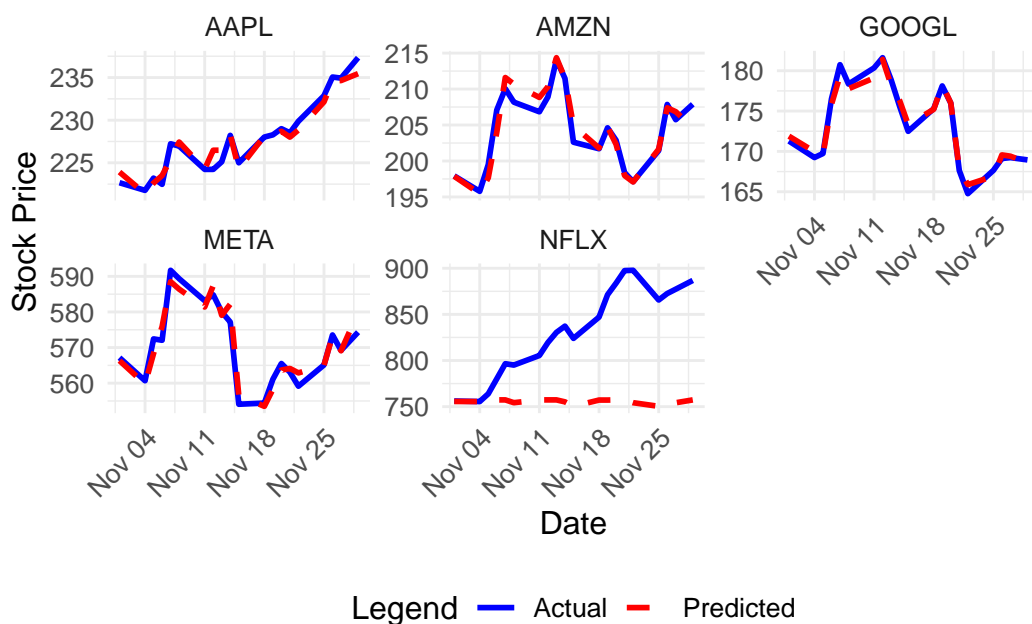


Figure 2: Actual vs Predicted Stock Prices for November 2024

Figure 2 presents the comparison of actual versus predicted stock prices for the FAANG companies during November 2024. For each company—Apple (AAPL), Amazon (AMZN), Meta (META), Netflix (NFLX), and Google (GOOGL)—we plot the actual adjusted closing prices (in blue) alongside the predicted values (in red). The predictions, derived from our XGBoost models, demonstrate a generally strong alignment with the observed stock price movements, particularly for AAPL and META, where the predicted prices closely follow the actual trends. However, for companies like NFLX, the predicted values show a less consistent match, especially later in the month. This discrepancy suggests that while the model captures major trends effectively, there may be some volatility or market events that were not fully accounted for in the prediction process. These visualizations help assess the performance of our models and highlight areas where the predictions are robust as well as where there is room for improvement.

## 4.2 Heatmap of Prediction Errors (November 2024)

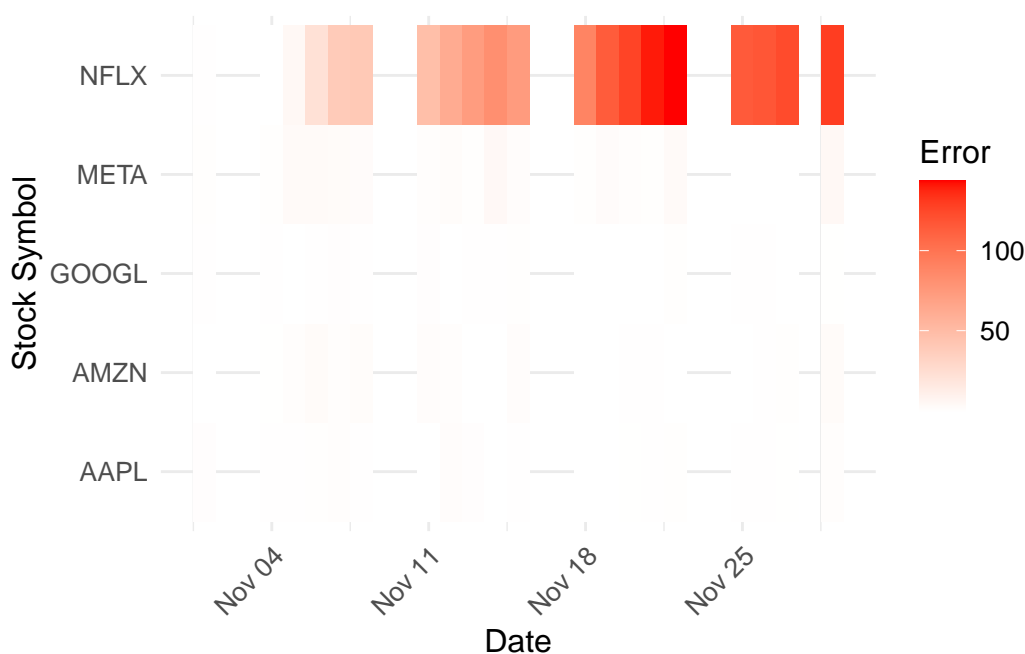


Figure 3: Heatmap of Prediction Errors for November 2024

Figure 3 highlights the discrepancies between the predicted and actual stock prices for each FAANG company on a daily basis. The color intensity, ranging from light pink to deep red, indicates the magnitude of the errors, with darker shades representing higher errors. Notably, the Netflix (NFLX) stock shows consistently higher errors throughout November, particularly on November 4th and November 11th, where the error values reach above 100, as indicated

by the deeper red color. Meta (META) and Google (GOOGL) exhibit moderate errors, with values peaking around November 11th but staying relatively consistent compared to NFLX. In contrast, Amazon (AMZN) and Apple (AAPL) show smaller error values, especially in the first half of the month, where the colors are much lighter. This heatmap provides a clear view of which stocks and dates the model struggled with the most, offering insights into potential areas for further model improvement.

### 4.3 Average Predicted Stock Price - November

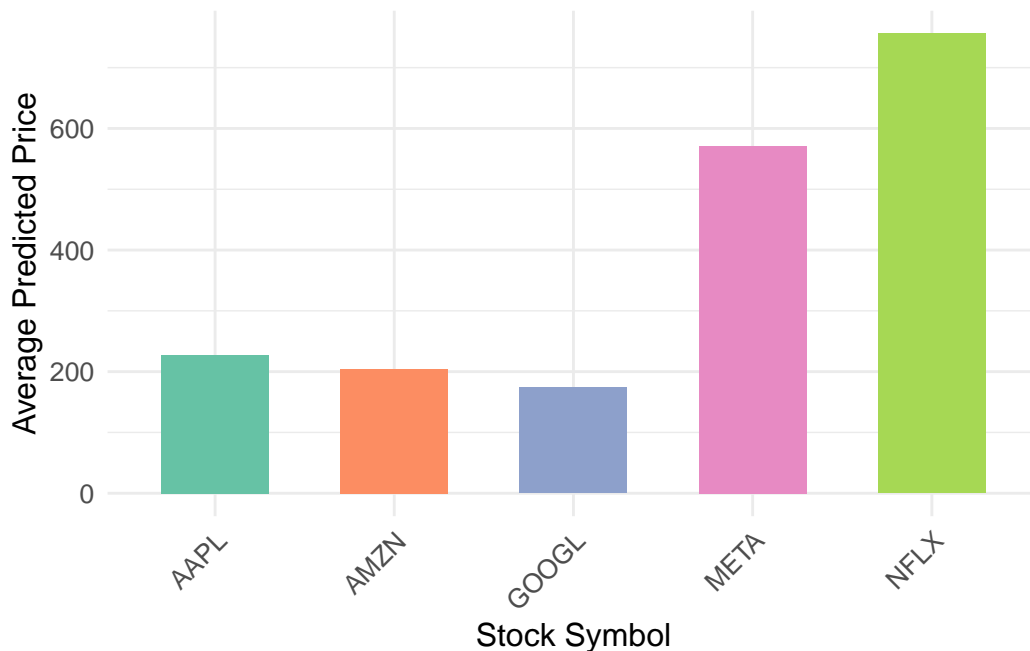


Figure 4: Average Predicted Stock Prices for November 2024

Figure 4 illustrates the average predicted stock prices for each FAANG company in November 2024. The predictions reveal significant variations across the companies, with Netflix (NFLX) showing the highest predicted average price, approximately **\$600**, followed by Meta (**'META'**) with an average just above **\$500**. In contrast, Apple (**'AAPL'**), Amazon (**'AMZN'**), and Google (**'GOOGL'**) exhibit considerably lower average predicted prices, with AAPL at around **\$225**, AMZN at **\$210**, and GOOGL at approximately **\$180**. This visualization highlights the distinct predicted performance of each company, suggesting that Netflix is expected to outperform the others in terms of stock price in November 2024, based on our model's predictions. These results provide a clear comparative overview of expected stock movements and serve as a foundation for further analysis of market trends in the upcoming months.

#### 4.4 Predicted Stock Price Trends - December

The line chart above presents the predicted stock prices for each FAANG company throughout December 2024. The predicted prices for each stock exhibit a strikingly consistent slope, with no significant fluctuations or adjustments throughout the month. For example, Netflix (NFLX) shows a steady increase, while Apple (AAPL), Amazon (AMZN), and Google (GOOGL) display relatively flat predicted price trends. This consistent pattern suggests potential limitations in the model, as real-world stock prices are typically subject to more volatility and market dynamics that are not captured by this prediction.

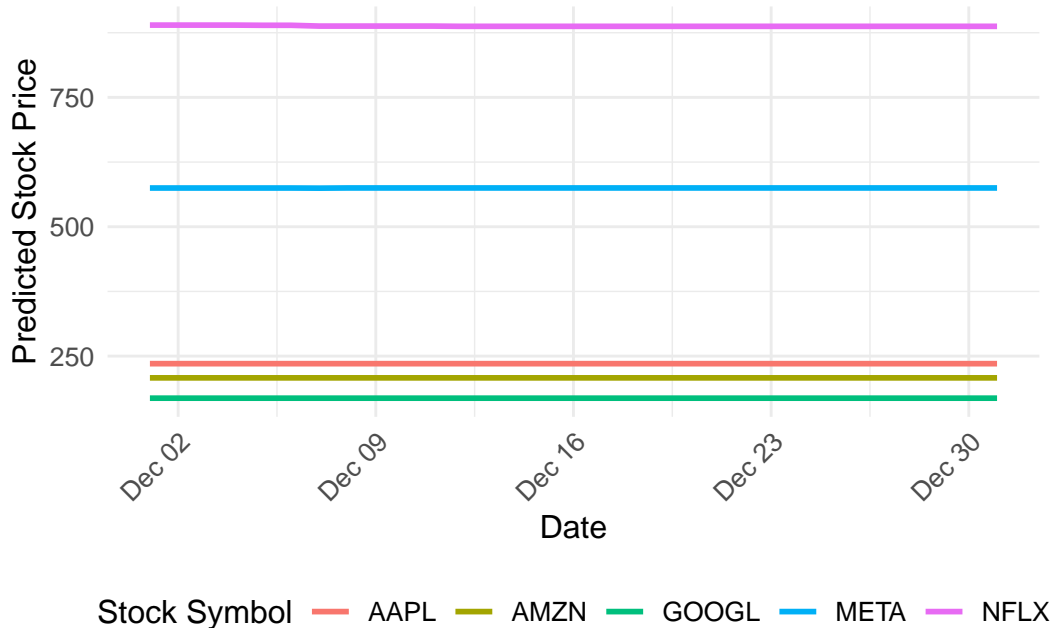


Figure 5: Predicted Stock Price Trends for December 2024

The lack of variability in the predictions may indicate that the model fails to account for potential market shifts, such as reactions to economic news, company announcements, or other external factors. Additionally, the absence of any sharp price movements could suggest that the model overly smooths the data, possibly due to the features used or the nature of the training data. These limitations highlight the importance of refining the model to incorporate broader market influences and improve its responsiveness to sudden changes in the stock market environment.



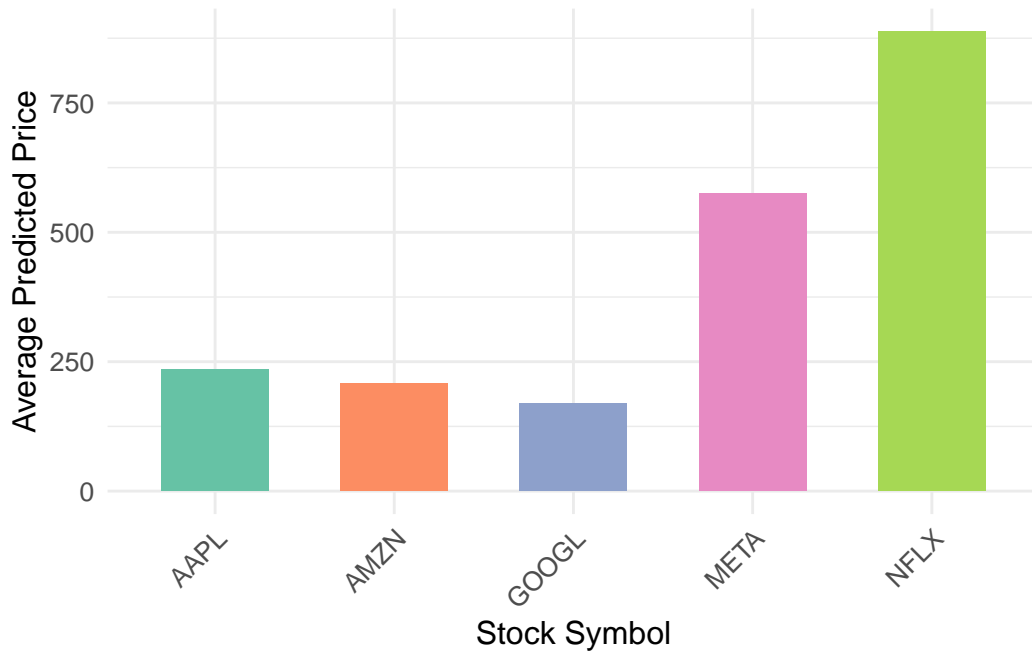


Figure 6: Average Predicted Stock Prices for December 2024

#### 4.5 Average Stock Prices Bar Chart - December

The bar chart above displays the average predicted stock prices for each FAANG company in December 2024. As shown, Netflix (NFLX) is expected to have the highest average predicted stock price, just above **\$750**, followed by Meta (META) at approximately **\$600**. In contrast, Apple (AAPL), Amazon (AMZN), and Google (GOOGL) show significantly lower predicted prices, with AAPL around **\$250**, AMZN slightly lower, and GOOGL at just above **\$200**. This suggests that, according to the model, Netflix is projected to continue outperforming other FAANG companies in terms of stock price. These predictions reflect market trends modeled from historical data, with a noticeable discrepancy between the higher-performing stocks and the others. The chart provides a comparative view of how each stock is expected to fare during December 2024, contributing to our understanding of potential market movements for the month.

#### 4.6 Average Predicted Stock Price - December

Table 2: Average Predicted Stock Prices for December 2024

Table 2: Average Predicted Stock Prices for December 2024

Stock Symbol	Average Predicted Price
NFLX	\$887.84
META	\$574.92
AAPL	\$235.37
AMZN	\$207.93
GOOGL	\$168.69

Table 2 presents the average predicted stock prices for each FAANG company in December 2024. According to the predictions, Netflix (NFLX) is expected to have the highest average stock price at **\$887.84**, followed by Meta (META) with an average of **\$574.92**. Apple (AAPL) is predicted to have an average stock price of **\$235.37**, while Amazon (AMZN) and Google (GOOGL) are expected to have lower average prices, at **\$207.93** and **\$168.69**, respectively. These predicted values highlight significant variations across the FAANG companies, with Netflix forecasted to substantially outperform the others. This table offers a clear numerical summary of the expected market performance for each company, providing insights into potential market trends for December 2024.

#### 4.7 Combined November and December

Figure 7 above illustrates the predicted stock prices for each FAANG company across November and December 2024. The predictions are shown for both months, with the November predictions represented by the teal lines and the December predictions shown in red. For companies like Apple (AAPL) and Meta (META), the model predicts a notable increase in stock prices during November, followed by a flat trend throughout December, indicating minimal movement in the predicted prices during the second month. In contrast, for Netflix (NFLX), the predicted stock price remains almost flat through both months, showing limited variation. Amazon (AMZN) and Google (GOOGL) exhibit more fluctuations in November, with predicted prices stabilizing into flat trends for December. This pattern highlights a potential limitation of the model, as it fails to capture significant market dynamics or volatility in December, suggesting that the model may not fully account for market influences or external factors that could impact stock prices at the end of the year.

#### 4.8 Average Predicted Stock Prices for November and December 2024

Figure 8 compares the average predicted stock prices for each FAANG company in November and December 2024. The chart clearly shows the predicted stock prices for November

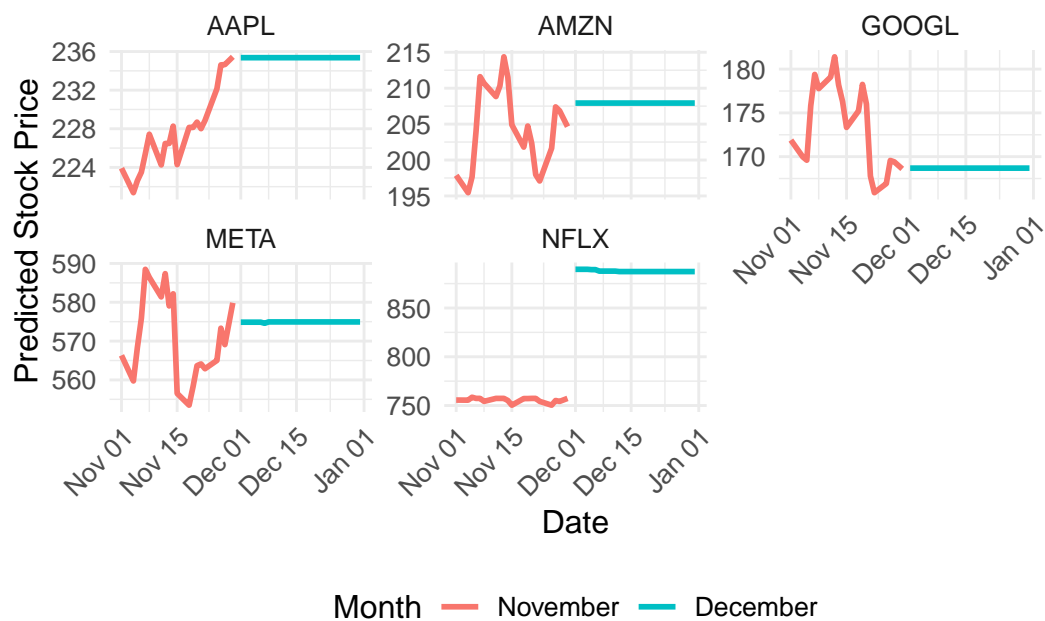


Figure 7: Predicted Stock Prices for November and December 2024

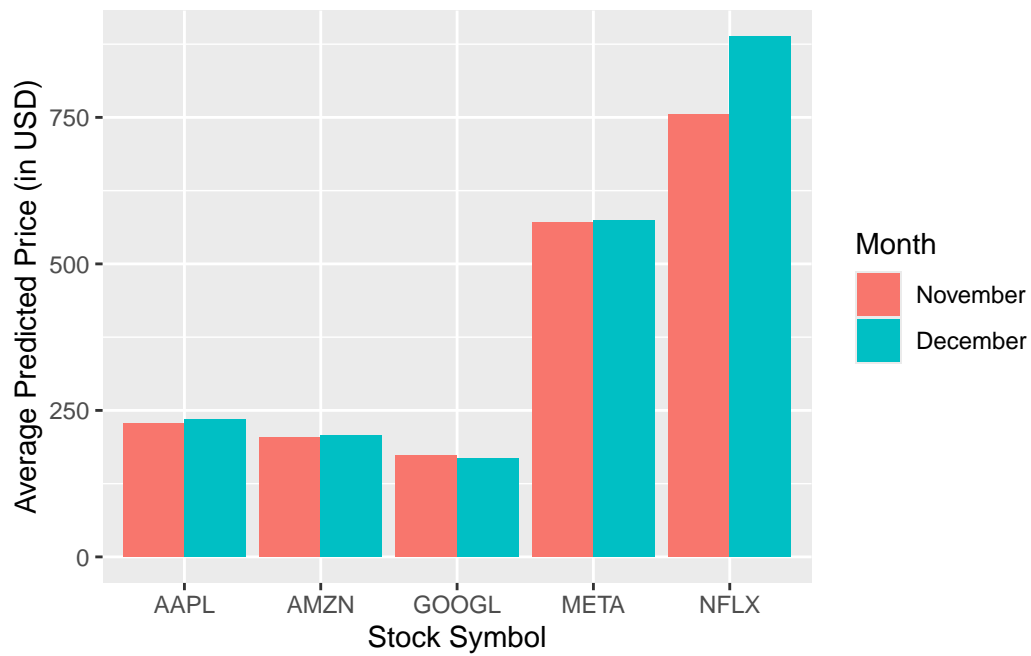


Figure 8: Average Predicted Stock Prices for November and December 2024

(red bars) and December (teal bars). Notably, Netflix (NFLX) is expected to experience the highest increase in stock price, with a marked difference between November and December. Meta (META) follows with a similar pattern, although the difference between months is less pronounced. In contrast, stocks like Apple (AAPL), Amazon (AMZN), and Google (GOOGL) show relatively stable predicted prices between the two months, with only slight variations. These results suggest that while some stocks, like Netflix, are predicted to have larger fluctuations, others show minimal change in their predicted prices, indicating potential stability or lack of significant market events influencing these stocks during December 2024.

## **5 Discussion**

### **5.1 Summary of Findings**

The predicted FAANG stock prices for November and December 2024 highlight several trends with real-world implications. In November, the predicted prices closely align with actual stock prices for AAPL, AMZN, and GOOGL, suggesting consistent market behavior across these stocks during the month. However, NFLX predictions were significantly lower than actual prices, likely reflecting unique market dynamics specific to the stock. These dynamics could include sector-specific factors, such as increased investor interest in streaming services, or company-specific announcements, such as major content releases or subscriber growth milestones, that were not captured in the historical data. Additionally, NFLX may be more sensitive to post-election market sentiment compared to other FAANG stocks, contributing to its divergence. Further analysis is needed to fully understand the market behavior of NFLX, particularly in response to macroeconomic and company-specific events during the period.

For December, the steady slope in predicted prices across all stocks, with NFLX and META leading in average value, suggests an expectation of stability or consistent growth. This could reflect broader investor confidence post-election. However, this trend might also stem from the analysis data extending only until November, limiting the model’s ability to capture variations within December. These observations suggest that while the predictions provide actionable insights for short-term planning, unique stock-specific dynamics and the influence of incomplete datasets warrant further investigation to fully understand long-term trends.

### **5.2 Model Performance and Limitations**

#### **5.2.1 November Model**

The November model demonstrated strong performance in predicting short-term stock price trends. Predictions for AAPL and GOOGL closely aligned with actual prices, indicating the model’s ability to effectively leverage historical data and features like moving averages and lagged prices. This suggests that the model captured stable market behaviors for these stocks

during November’s relatively steady market conditions. However, the model significantly underestimated NFLX prices, pointing to its inability to account for stock-specific dynamics such as sector-specific investor sentiment or company-specific announcements. This limitation arises from the reliance on quantitative features and the exclusion of qualitative factors, such as news events or broader macroeconomic trends. The model’s dependence on historical data also makes it less responsive to sudden or emerging trends, which are particularly impactful for stocks like NFLX.

### 5.2.2 December Model

The December model’s predictions revealed a flat trajectory across all stocks, highlighting its inability to capture intra-month variations and post-election volatility. This limitation is likely due to the analysis data only extending through November, restricting the model’s ability to learn patterns or fluctuations specific to December. Even with complete, constructed December data, the reliance on technical indicators and short-term historical trends may have limited the model’s ability to reflect dynamic market movements. Stocks like NFLX and META, which showed consistently higher predicted averages, may reflect overfitting to earlier trends, preventing the model from adapting to potential changes in market conditions.

Further model diagnostics, including feature importance evaluations and residual analysis, are provided in the Appendix [B.2](#).

## 5.3 US Election Effects

Table 3 illustrates the stock prices of FAANG companies one trading day before and after the elections for the years 2016, 2020, and 2024. The post-election period often shows noticeable fluctuations in stock prices, with some stocks experiencing a sharp increase and others a decrease. For example, AAPL saw a slight increase in 2016 and 2020, while AMZN and NFLX both exhibited declines in 2016. In contrast, the 2024 data reveals an overall upward trend across all FAANG stocks, with particularly strong gains for NFLX, which rose from \$755.51 before the election to \$780.21 after the election, reflecting a surge in stock prices.

Table 3: FAANG Stock Prices for One Trading Day Before and After the Election (2016, 2020, and 2024)

year	symbol	Before Election	After Election	Price Change
2016	AAPL	25.56	25.67	↑
2016	AMZN	39.25	38.59	↓
2016	GOOGL	40.00	40.18	↑
2016	META	121.78	122.81	↑
2016	NFLX	124.58	122.19	↓

Table 3: FAANG Stock Prices for One Trading Day Before and After the Election (2016, 2020, and 2024)

year	symbol	Before Election	After Election	Price Change
2020	AAPL	106.18	112.21	↑
2020	AMZN	150.22	162.06	↑
2020	GOOGL	81.02	87.08	↑
2020	META	260.58	286.52	↑
2020	NFLX	484.12	496.95	↑
2024	AAPL	221.77	222.48	↑
2024	AMZN	195.78	207.09	↑
2024	GOOGL	169.24	176.51	↑
2024	META	560.68	572.05	↑
2024	NFLX	755.51	780.21	↑

This post-election stock behavior suggests that broader market reactions to election outcomes may have impacted stock prices significantly, particularly towards the end of the election year. Given this, it is likely that the increase in stock prices in late 2024 contributed to the predicted stock price trends in our model. As our model leverages historical data and trends starting from 2020, including fluctuations in stock prices towards the end of 2024, this surge likely influenced the final predictions, making them more optimistic and reflecting the strong market response post-election. This potential influence should be considered when interpreting the results, as it emphasizes the role of market sentiment and external factors that may cause sharp fluctuations in stock prices, beyond the intrinsic factors captured by the model.

## Appendix

### A Additional Data Details

#### A.1 Data Cleaning

The data cleaning process transformed raw FAANG stock price data from Yahoo Finance into a structured format suitable for predictive modeling. The dataset spans daily stock prices for Meta (META), Amazon (AMZN), Apple (AAPL), Netflix (NFLX), and Alphabet (GOOGL) from January 2020 to November 2024. To prepare for December 2024 predictions, placeholder dates were added to ensure a continuous sequence of daily observations for each stock, accounting for non-trading days such as weekends and holidays. Missing values in the adjusted closing price column were handled through forward and backward interpolation, ensuring the continuity required for calculating rolling statistics and lagged features.

Feature engineering played a central role in enriching the dataset. Key features included **Lag\_1** (adjusted closing price from the previous day), **Rolling\_Mean\_7** (7-day rolling average), **sma\_20** and **sma\_50** (20-day and 50-day simple moving averages), **volatility** (20-day rolling standard deviation), and **daily\_return** (percentage change in adjusted closing price). Stock symbols were numerically encoded to allow the model to recognize patterns across multiple stocks. Rigorous data validation ensured that no critical missing values remained, all feature values fell within logical ranges, and the dataset was consistent. The cleaned and feature-rich dataset was saved in Parquet format for efficient processing and compatibility with the machine learning workflow.

#### A.2 Data Limitations and Future Directions

While the cleaned dataset is robust and suitable for modeling, there are notable limitations. First, external factors such as macroeconomic indicators, earnings announcements, or industry-specific news are not included, despite their potential to significantly impact stock prices. Second, the interpolation of missing values assumes smooth price continuity, which may not fully account for abrupt market changes or high-impact events during non-trading days. The analysis is also limited to FAANG stocks, which, while influential, may not reflect broader market dynamics or trends that could indirectly affect these stocks. Expanding the dataset to include other major stocks or sectors could provide a more comprehensive view of market behavior. Furthermore, the models are designed for short-term forecasting, specifically for December 2024, and are not optimized for capturing longer-term trends or cyclical patterns.

Lastly, while the engineered features effectively summarize historical trends, they may miss more complex interactions or latent patterns in stock price movements. Future analyses could incorporate additional data sources, such as macroeconomic indicators or sector-specific variables, to enhance predictive accuracy. Advanced feature engineering techniques, like seasonal

decomposition or non-linear transformations, could also improve the models' ability to capture nuanced market behaviors. Addressing these limitations would strengthen the robustness and generalizability of the analysis, offering deeper insights into FAANG stock dynamics.

## **B Additional Model Details**

### **B.1 Model Justification - Further Details**

The XGBoost regression model was chosen for predicting FAANG stock prices in December 2024 due to its ability to model complex non-linear relationships, which are common in financial time series data. Given the dynamic and time-sensitive nature of stock prices, the November and December models were trained separately to ensure that predictions account for evolving market trends and feature importance within each specific period.

#### **B.1.1 Why Separate November and December Models?**

Separate models for November and December were developed to capture the unique characteristics of stock price movements during these months. By training the November model using data up to October 31, 2024, and the December model using data up to November 30, 2024, we ensure that each model incorporates the most recent market information available for its respective forecasting period. This separation minimizes over-fitting to earlier trends and allows the December model to adapt to any shifts in market behavior or volatility observed in November.

#### **B.1.2 Model Justification and Features**

XGBoost was selected due to its ability to handle non-linear interactions and temporal dependencies, which are crucial for stock price prediction. Key engineered features include:

- **Lag\_1**: Represents the adjusted closing price from the previous day, capturing short-term price dependency.
- **Rolling\_Mean\_7, sma\_20, sma\_50**: Reflect short- and medium-term trends, offering insights into market behavior over recent periods.
- **Volatility**: Captures price variability using the 20-day rolling standard deviation, a critical factor in understanding market risk.
- **Daily\_Return**: Quantifies day-to-day price changes, highlighting rate-of-change trends.
- **Symbol\_Encoded**: Encodes stock-specific identifiers numerically, enabling the model to generalize patterns across multiple FAANG stocks.



By incorporating these features, the model is able to account for short-term fluctuations and medium-term market dynamics, which are critical for accurate stock price forecasting.

### B.1.3 Hyperparameters and Regularization

The November and December models share key hyperparameters optimized to balance complexity and generalizability: - **Max depth**: Set to 6, allowing the model to capture significant interactions without overfitting. - **Learning rate**: Set to 0.1, ensuring gradual refinement of predictions. - **Boosting rounds**: Fixed at 100 to provide sufficient iterations for model convergence. XGBoost's built-in regularization techniques mitigate overfitting, especially important given the high-dimensional engineered features and the presence of missing values due to holidays and weekends.

Unlike traditional time-series models such as ARIMA, which assume stationarity and often require extensive preprocessing, XGBoost directly handles non-stationary data. This flexibility makes it an ideal choice for stock price forecasting, where trends and volatility can vary significantly over time. Furthermore, XGBoost automatically handles missing data, making it robust for real-world applications with gaps in trading days. The decision to train separate models for November and December, coupled with XGBoost's ability to handle complex non-linearities, ensures that the forecasts reflect both temporal dependencies and evolving market dynamics. By balancing model flexibility and interpretability, the November and December XGBoost regression models provide a robust and reliable framework for predicting FAANG stock prices, tailored to the unique characteristics of each time period.

## B.2 Model Diagnostics

### B.2.1 Feature Importance

Figure 9 provides a clear visualization of the relative contributions of the predictive variables used in the XGBoost model. The variable `Lag_1`, which represents the previous day's adjusted price, demonstrates the highest importance, significantly outweighing all other features. This indicates that the previous day's price is a critical factor in predicting the current adjusted price, aligning with the strong autocorrelation often observed in time series data like stock prices. `Rolling_Mean_7`, the 7-day rolling mean of adjusted prices, also contributes meaningfully, suggesting that short-term trends influence the model's predictions. The remaining features, including `daily_return`, `symbol_encoded`, and various moving averages (`sma_20`, `sma_50`) and volatility, show minimal importance in this model. This disparity suggests that the model heavily relies on recent past prices and short-term trends while assigning relatively less weight to longer-term trends or encoded categorical variables. Future iterations of the model could explore whether these less important features can be refined or excluded to simplify the model further without impacting its performance.

## Feature Importance for November Model

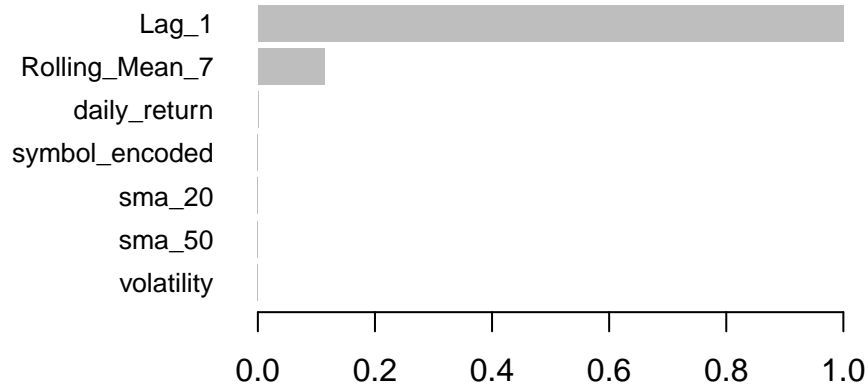


Figure 9: Feature Importance

### B.2.2 Residual Analysis

Figure 10 evaluates the model's performance by analyzing the residuals (errors) against the predicted (fitted) values. The residuals appear randomly scattered around the red dashed line at zero, which is a good indication that the model does not suffer from significant bias. However, the spread of residuals does not remain constant, with some visible heteroscedasticity (changing variance) at higher fitted values. This suggests that the model's prediction errors may increase for certain ranges of the target variable, indicating a potential need to address variance issues or reevaluate the model's assumptions.

### B.2.3 Error Distribution

Figure 11 shows the frequency of residuals (errors) from the model's predictions. The distribution is approximately bell-shaped and centered around zero, suggesting that the residuals are fairly normally distributed. This is a good indication that the model does not exhibit significant bias and aligns with the assumption of normally distributed errors in regression. However, there are slight tails on both sides, indicating occasional larger prediction errors, which may warrant further investigation to determine if these are due to outliers or systematic issues in the model.

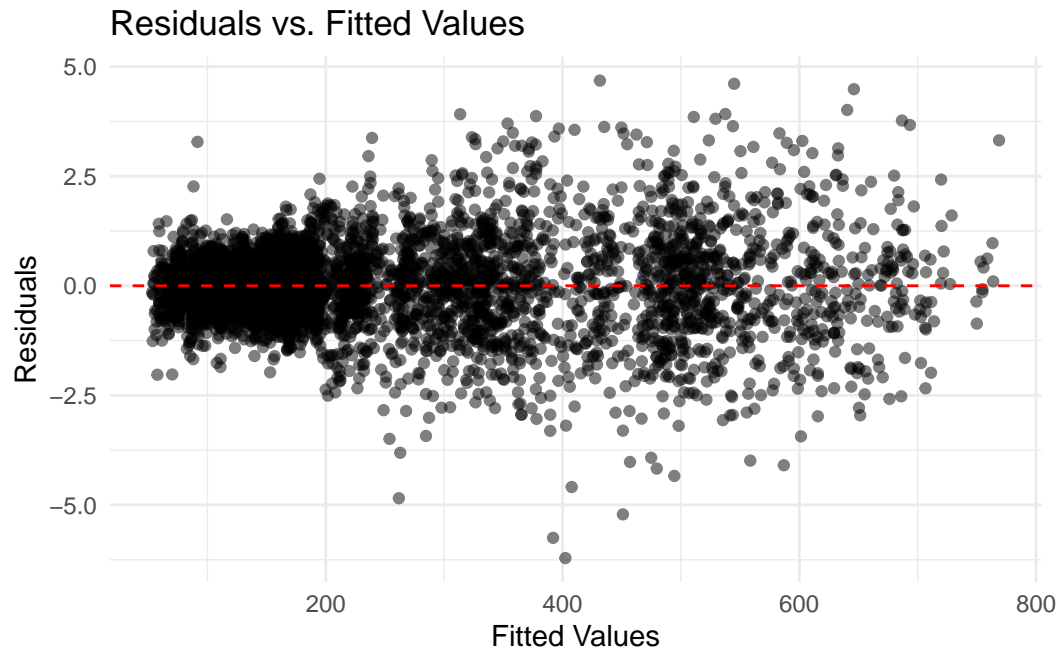


Figure 10: Residual vs. Fitted Values

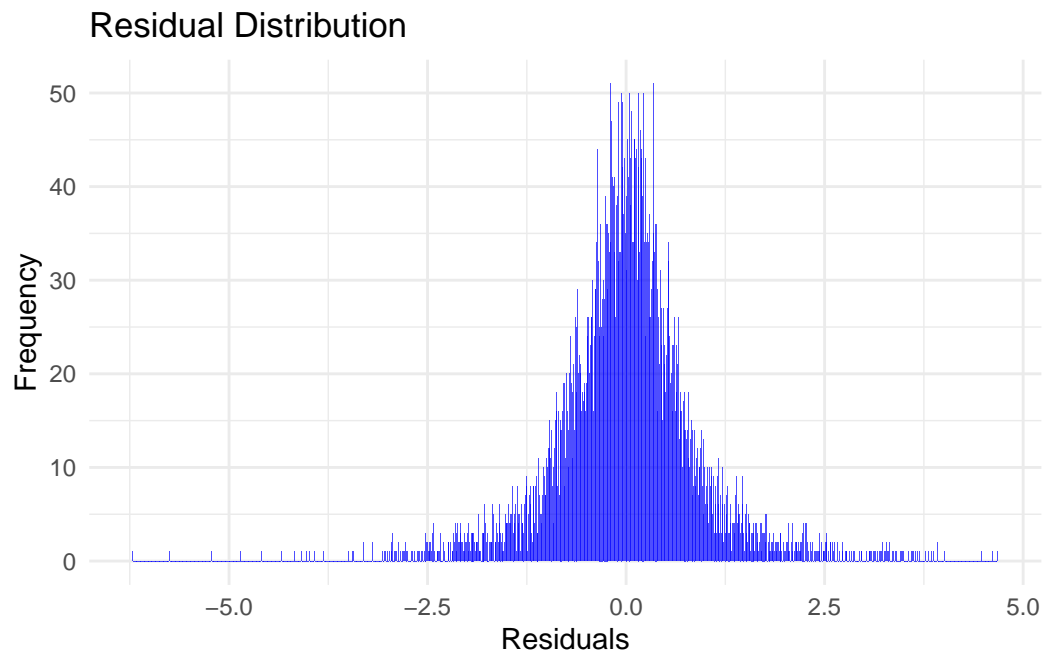


Figure 11: Residual Distribution

### B.2.4 Prediction vs. Actual Values

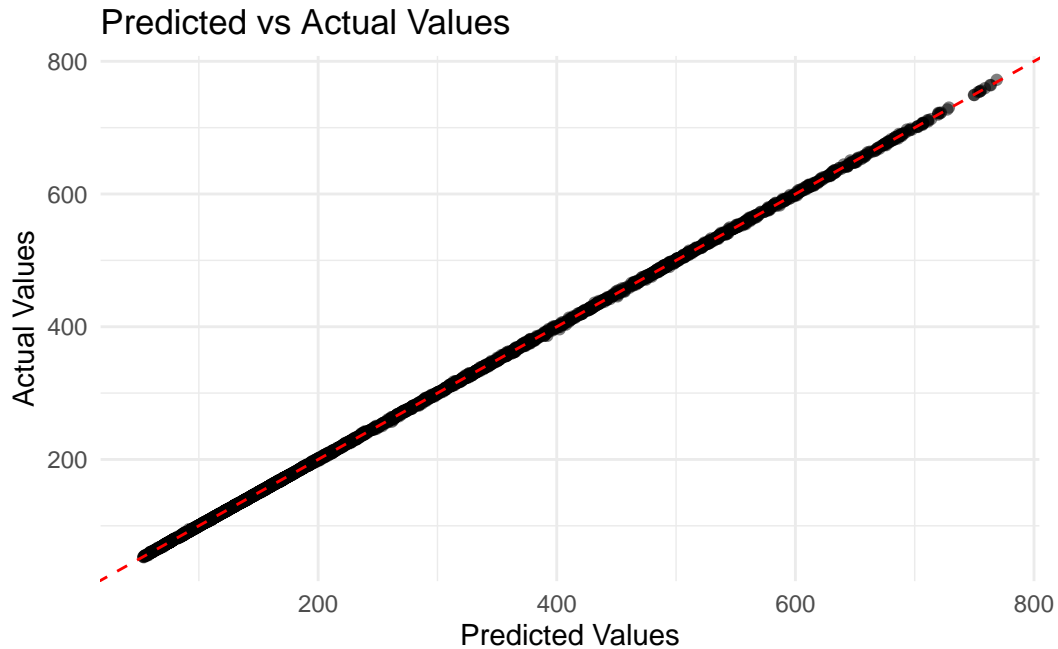


Figure 12: Predicted vs. Actual Values

Figure 12 demonstrates the model's accuracy by comparing its predictions with the true target values. The points align closely along the red dashed diagonal line, indicating a high level of agreement between the predicted and actual values. This suggests that the model performs well in predicting the adjusted stock prices with minimal error. The tight clustering along the diagonal further supports the reliability of the model's output, with no significant deviations or systematic bias observed in the predictions.

## B.3 Model Performance and Limitations

### B.3.1 Model Performance

Based on the evaluation metrics and diagnostic checks performed on both the November and December XGBoost models, we can conclude that the models demonstrate strong performance and reliability in forecasting FAANG stock prices. The close alignment between the observed stock prices and the predicted values, as evidenced by the residual analysis and predicted vs. actual plots, indicates that the models are well-calibrated. This calibration ensures that the models accurately capture both the central trends and variability in stock prices across the

FAANG group. Such alignment is critical for financial forecasting, where precise predictions are required to make informed investment decisions.

The residual distribution further supports the robustness of the models, with residuals centered around zero and following an approximately normal distribution. This suggests that the models do not suffer from significant bias and handle the inherent noise in financial time-series data effectively. The use of engineered features, including lagged values, rolling averages, and volatility measures, has allowed the models to capture both short- and medium-term trends in stock price movements, contributing to their high predictive accuracy.

The decision to develop separate models for November and December has further enhanced performance by tailoring the forecasts to the unique market dynamics of each period. The November model, trained on data up to October 31, 2024, successfully captures trends leading into November, while the December model, trained on data up to November 30, 2024, incorporates additional market shifts and variability observed during November trading. This approach ensures that each model reflects the most recent data available, improving the reliability of their respective forecasts.

Additionally, feature importance analysis highlights the predictive power of key variables such as **Lag\_1**, **Rolling\_Mean\_7**, and **volatility**, which play critical roles in explaining stock price movements. The robustness of the models is further validated by their ability to generalize across multiple FAANG stocks, as demonstrated by their performance on unseen data.

### B.3.2 Model Limitations

While the November and December XGBoost models demonstrate strong performance, several limitations should be noted:

1. **Assumption of Stationarity:** Although XGBoost can handle non-stationary data, the models rely on historical data patterns to forecast future trends. Sudden market shocks, such as regulatory changes or geopolitical events, may lead to deviations from these patterns, reducing predictive accuracy.
2. **Feature Representation:** Despite the inclusion of lagged values, rolling averages, and volatility measures, the models may not fully capture all relevant factors influencing stock prices, such as macroeconomic indicators, earnings reports, or sector-specific news.
3. **Overfitting Risk:** Although regularization techniques were applied, there remains a risk of overfitting due to the high-dimensional feature space and temporal correlations in the data. This could result in reduced generalizability, particularly for unseen market conditions.
4. **Limited Interpretability:** While feature importance provides some insights, XGBoost models are inherently black-box in nature. This limits their ability to provide clear

explanations for predictions, which may be important for stakeholders seeking more interpretive models.

5. **Data Gaps and Extrapolation:** The interpolation of missing data, particularly for weekends and holidays, assumes continuity in stock price movements. This may introduce inaccuracies, especially if significant events occur during non-trading days.
6. **Short-Term Forecast Horizon:** The models are designed to forecast stock prices for December 2024. Their utility diminishes for longer-term predictions, where additional dynamics and external variables may play a larger role.

Overall, the November and December XGBoost models exhibit strong predictive performance and reliability, underpinned by thorough calibration, effective feature engineering, and tailored training periods. However, their performance is subject to certain limitations, including sensitivity to market shocks and the inherent challenges of modeling complex financial data. Despite these constraints, the models provide meaningful insights into FAANG stock dynamics and serve as a valuable tool for understanding and predicting market trends during these critical months.

## C Surveys, Sampling, and Observational Data

### C.1 Nature of Observational Data

The data used in this study is observational, comprising naturally occurring financial transactions within the stock market. Unlike experimental or survey data, observational data in financial markets is influenced by external factors beyond experimental control, such as macroeconomic policies, geopolitical events, and investor sentiment. This introduces complexities in distinguishing causality from correlation in stock price movements, as unmeasured confounding factors may bias interpretations. Observational data provides a high-frequency, granular perspective of market dynamics. However, its reliance on historical events as proxies for predictive features limits its ability to account for unprecedented external shocks or changes in market behavior, such as those induced by elections or global crises.

### C.2 Sampling Frame and Representativeness

The sampling frame for this analysis includes all trading days from January 2020 through November 2024, capturing a comprehensive record of FAANG stock activity. This time period reflects critical phases of economic recovery, technological advancement, and significant political events, such as the 2024 U.S. Presidential Election. While this dataset ensures comprehensive coverage of FAANG stocks, it remains constrained by sector-specific focus.

### C.2.1 Strengths

- **High-Frequency Coverage:** Daily data provides detailed insights into short-term fluctuations and longer-term trends.
- **Sector Focus:** FAANG stocks, as market leaders, act as strong indicators of broader economic conditions and investor sentiment.

### C.2.2 Limitations

- **Sector Bias:** Focusing solely on FAANG stocks limits the generalizability of findings to other industries or market segments.
- **Data Gaps:** Placeholder prices for December 2024, generated through imputation, may underestimate volatility or fail to capture post-election market adjustments accurately.

## C.3 Integrating Surveys: A Complementary Perspective

While stock prices offer a quantitative view of market behavior, investor sentiment surveys add a qualitative dimension, revealing expectations and perceptions that often precede observed price movements. By integrating survey methodologies, the analysis can better contextualize and interpret stock price dynamics during critical events like elections.

### C.3.1 Proposed Survey Design

To complement the observational stock data, investor sentiment surveys could include questions such as: - **Market Confidence:** “Do you expect FAANG stock prices to increase, decrease, or remain stable over the next month?” - **Election Impact:** “How has the 2024 U.S. Presidential Election influenced your investment decisions?” - **Risk Appetite:** “What is your likelihood of investing in technology stocks in the current market climate?”

### C.3.2 Methodological Considerations

- **Sampling and Representativeness:** Stratify respondents by demographics, portfolio size, and investment experience to ensure diverse market perspectives.
- **Weighting:** Use post-survey weighting to align survey demographics with actual investor populations, similar to post-stratification in traditional surveys.

### C.3.3 Strengths and Limitations

#### C.3.3.1 Strengths:

- **Forward-Looking Insights:** Surveys capture investor sentiment that may not yet be reflected in stock prices.
- **Diverse Perspectives:** Stratified sampling ensures representation across institutional and retail investors.

#### C.3.3.2 Limitations:

- **Response Bias:** Survey respondents may overstate optimism or pessimism, especially during volatile periods.
- **Sampling Complexity:** Ensuring representativeness across various investor demographics is resource-intensive.

### C.4 Parallels Between Observational Data and Surveys

Observational stock data shares several structural similarities with survey methodologies: - **Stratification:** The division of data by stock symbol mirrors stratified sampling in surveys, allowing for stock-specific analysis while maintaining aggregate insights. - **Imputation:** Just as survey non-responses require careful imputation to avoid bias, missing stock data necessitates robust methods to estimate gaps, such as forward/backward filling or time-series modeling. - **Weighting:** Future models could incorporate macroeconomic indicators (e.g., GDP, interest rates) as weights, akin to post-stratification in survey data, to adjust for external market influences.

### C.5 Addressing Data Challenges

- **Temporal Dependence:** Stock prices exhibit temporal dependence, where current values are influenced by past trends. To mitigate this, predictive features such as lagged prices (`Lag_1`) and rolling averages (`Rolling_Mean_7`) are included, capturing both momentum and smoothed trends.
- **Imputation for Missing Data:** Imputation was used to fill missing December 2024 values, ensuring data continuity. However, this approach risks underestimating potential volatility or masking significant market responses to external events, such as post-election dynamics. Future work could leverage time-series forecasting models to improve imputation accuracy.



- **Observational Bias:** External confounders, such as Federal Reserve policies, geopolitical tensions, or earnings reports, can obscure causal relationships. Including predictors that account for such factors in future models could mitigate these biases.

## C.6 Recommendations for Future Work

1. **Survey Integration:** Combine investor sentiment surveys with stock data to enhance interpretability and predictive power.
2. **Expanding Scope:** Extend the analysis to non-FAANG sectors for broader market generalizability.
3. **Improved Imputation:** Apply advanced time-series models to better estimate missing values and capture expected volatility.

By addressing these challenges and integrating survey methodologies, future analyses can provide a richer understanding of market dynamics, improving the robustness and applicability of findings.

## D References