Project Report

Time Series Forecasting : FTSE100 Stock Price and Returns

FTSE 100

Student Names and IDs

Natchapol Chamnankanca 6422782134
Thanapat Thaipakdee 6422782480
Woraseth Limwanich 6422782621
Noppasorn Suwannakul 6422781474

Instructor's Name
Dr. Watthanasak Jeamwatthanachai
Sirindhorn International Institute of Technology
Thammasat University

Table of Contents

1. Overview

1.1) Introduction

The ability to accurately forecast financial time series data is a crucial skill in the field of financial engineering. Time series forecasting models are vital in predicting future price movements, significantly enhancing investment strategies, risk management, and overall financial decision-making. This project provides hands-on experience with various time series forecasting techniques using the FTSE 100 Close price data. The FTSE 100, comprising 100 of the largest publicly listed companies in the United Kingdom, serves as a key benchmark for the overall performance of the UK stock market. Accurately forecasting its closing prices can offer valuable insights to financial analysts and investors.

The dataset used in this study includes daily closing prices of the FTSE 100 index over several years, allowing for the analysis of long-term trends, seasonal patterns, and inherent noise in financial data. By experimenting with different forecasting models, participants will explore these factors and evaluate the models' performance using metrics such as RMSE and MAE.

1.2) Objective of the Project

The primary objective of this project is to provide practical experience in applying various time series forecasting models to the FTSE 100 Close price data. Through working with historical data, participants will engage in exploring different forecasting techniques to understand how they function in real-world financial scenarios. This hands-on approach allows participants to gain insight into the underlying patterns, trends, and seasonality inherent in financial time series data, which are critical for making accurate predictions in financial markets. By experimenting with a range of forecasting models, participants will assess each model's performance using appropriate metrics, such as RMSE and MAE, to measure accuracy and predictability. The project will also encourage a deeper analysis of the strengths and weaknesses of each model, fostering an understanding of when and why certain models perform better under specific conditions.

The key goals of this assignment are to:

- Understanding of time series analysis and forecasting methodologies.
- Gain the ability to identify trends, seasonality, and patterns in financial time series data.
- Assess the effectiveness of various forecasting models using relevant performance metrics.
- Analyze the strengths and limitations of each model and justify their use in financial engineering applications, such as stock price prediction, risk management, and investment strategy formulation.

2. Data Overview

In this section, we emphasize the crucial aspects of data loading and preprocessing, which are foundational steps for any data-driven project. These steps ensure that the dataset is clean, structured, and ready for analysis, thereby enabling effective model development. For this project, we focus on historical data from the FTSE 100 index, particularly its daily closing prices and returns over an extended time frame. This dataset offers a wealth of information to identify market trends, seasonal fluctuations, and other key features that are critical for accurate time series forecasting.

The dataset also includes return data, which measures the percentage change in closing prices from one day to the next. This metric enhances our ability to analyze volatility, price movements, and general market behavior. With both price and return data, we are better equipped to explore and model financial time series in-depth.

2.1) Import Libraries

The initial step involves importing essential libraries such as pandas for data manipulation, numpy for numerical computations, and matplotlib for visualizing the data. Additionally, time series-specific libraries like statsmodels and sklearn are utilized for modeling and evaluation purposes. These libraries provide a robust framework for handling large datasets and performing advanced analyses, making them indispensable for this study.

2.2) Load the dataset

```python
import numpy as np

# Calculate daily returns: Return(t) = (P(t) - P(t-1)) / P(t-1)
ftse_data['Return'] = ftse_data['Close'].pct_change()

# Or calculate log returns (Optional)
ftse_data['Log_Return'] = np.log(ftse_data['Close'] / ftse_data['Close'].shift(1))

# Drop NaN values created by the shift
ftse_data.dropna(inplace=True)

# Save the updated dataset to a new CSV
ftse_data.to_csv('ftse_data_with_returns.csv', index=False)

print("FTSE 100 data with returns saved to 'ftse_data_with_returns.csv'")
```

Figure 2.2 : *Generated return*

The dataset, comprising historical FTSE 100 daily closing prices, is loaded into the environment using Yahoo Finance (yfinance) and numpy generated daily inflation rate. This dataset spans multiple years and captures critical financial information, including the opening, high, low, and adjusted closing prices and return, as well as trading volume. The data is indexed by the trading date, ensuring that all subsequent analyses are time-dependent and accurately reflect the chronological sequence of price movements.

2.3) Exploratory Data Analysis (EDA):

EDA is a pivotal phase where the dataset is examined for basic statistical properties, patterns, and anomalies. Key insights such as the average price, volatility, and range of returns are explored to understand the dataset's underlying structure. Visualization techniques such as line plots, histograms, and boxplots are employed to provide an intuitive understanding of trends and deviations in both price and return data. These visual aids offer a preliminary understanding of potential seasonality or trends that could influence model design.

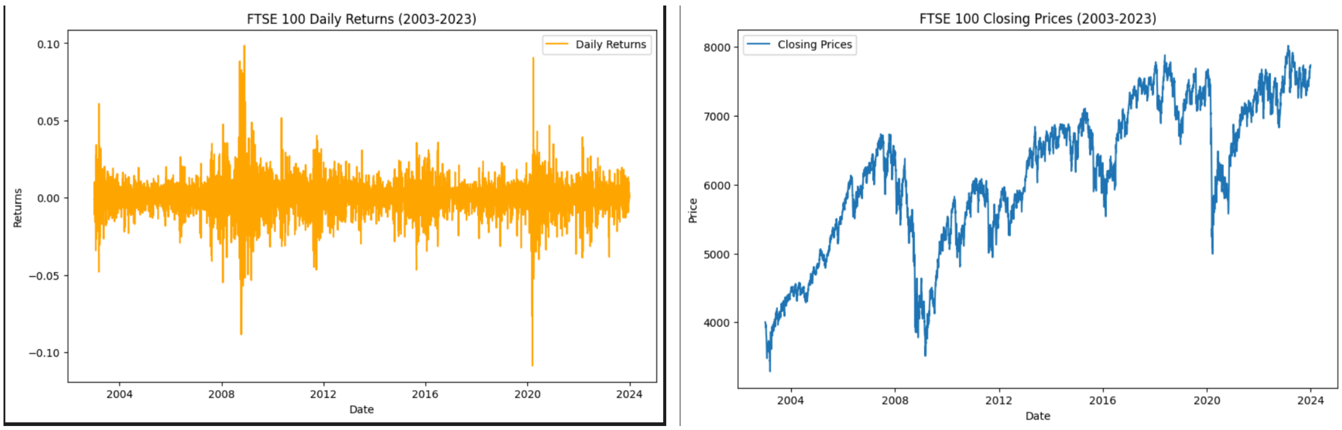

Figure 2.3 : *Daily Returns and Closing Prices*

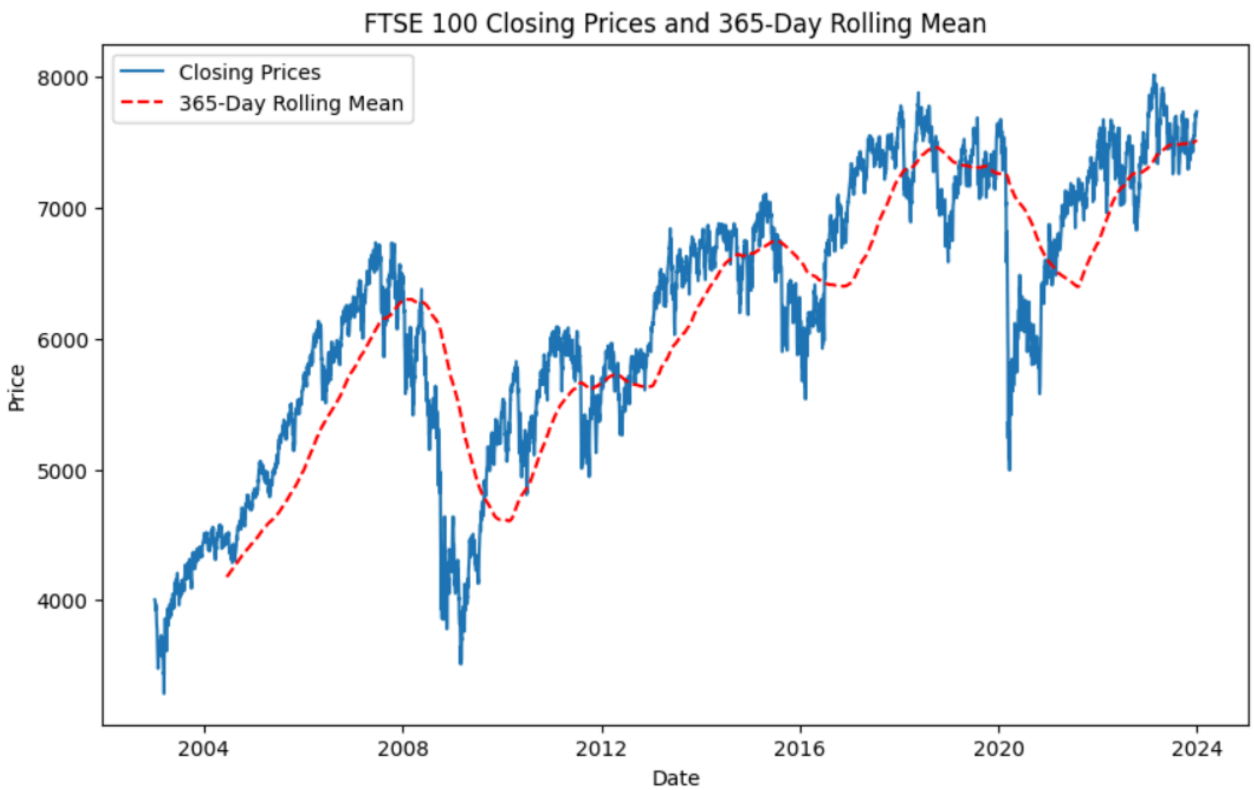

Figure 2.3.1 : *FTSE 100 Closing Prices and 365-Day Rolling Mean*

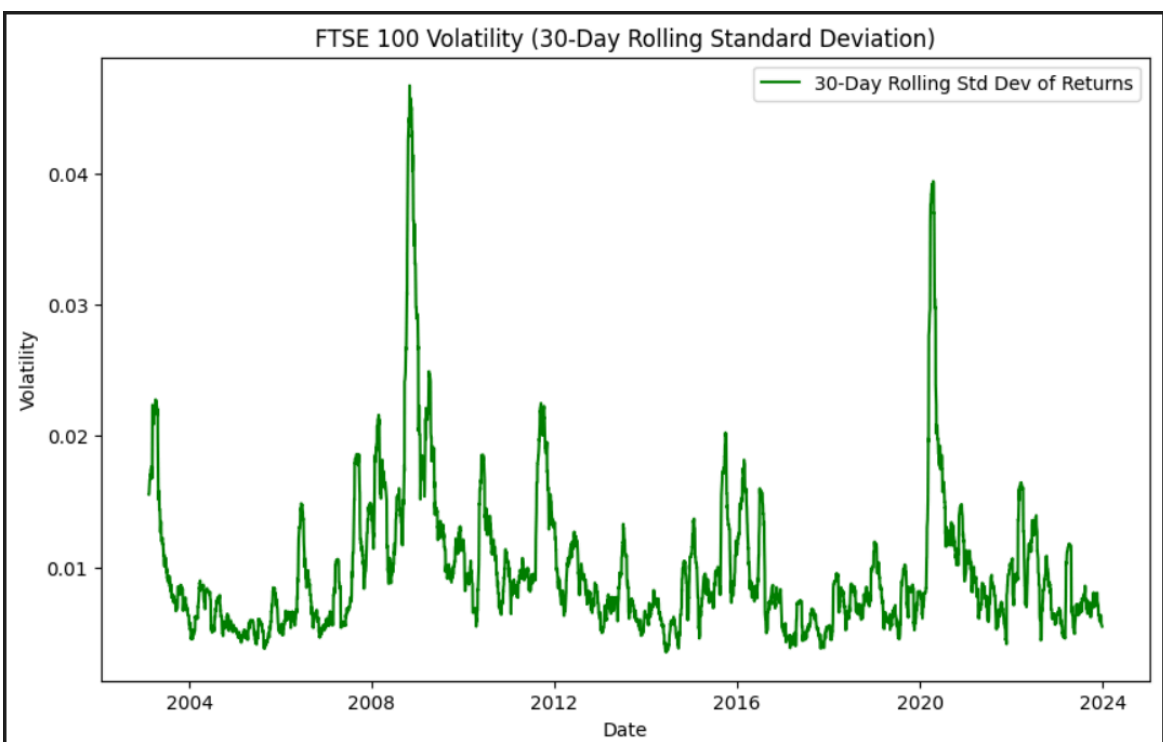

Figure 2.3.2 : *FTSE 100 Volatility*

2.4) Handle with missing data and error

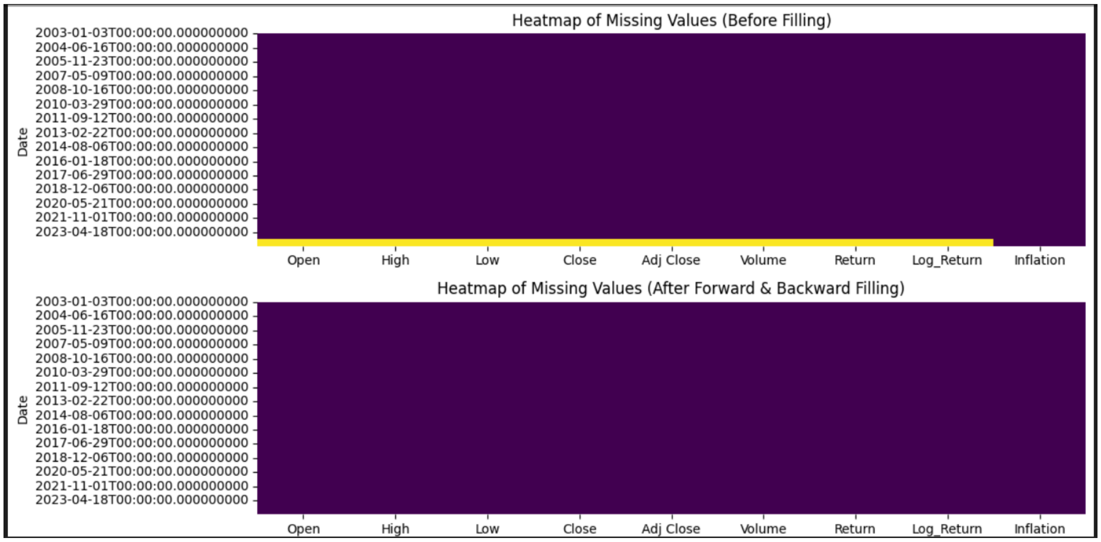

Figure 2.4 : *Heatmap of Missing Values*

Real-world financial datasets often contain missing values or inconsistencies due to market closures or data collection errors. To address this, missing values in the FTSE 100 dataset were handled using forward-fill (ffill) and backward-fill (bfill) techniques, ensuring continuity without introducing artificial patterns. Additionally, outliers and data duplicates were identified and addressed to prevent skewing the results. These preprocessing steps ensure that the dataset is both reliable and reflective of real-world market conditions, thus forming a solid foundation for forecasting and analysis.

3. Model Overview

In this section, we outlined the forecasting models used for analyzing and predicting FTSE 100 trends from 2003 to 2023. Each model is described with its unique characteristics, purpose, and relevance to stock forecasting

3.1) Naive Forecasting Models (Benchmarks)

3.1.1 Naive Forecast : The Naive Forecast model uses the last observed value in the time series to predict all future values. It serves as a baseline to compare the effectiveness of more complex models

- **Formula**

$$\hat{y}_{t+1} = y_t$$

Where $\hat{y}_{t+1}$ is the forecasted value for the next point, and y_t is the most recent observed value

- **Purpose:** Provides a simple benchmark for comparison. Although simplistic, it is often effective for series with little fluctuation

3.1.2 Simple Moving Average (SMA) : The SMA model calculates the average of the closing prices over a specified window, smoothing short-term fluctuations and highlighting longer-term trends.

- **Formula**

$$SMA_t = \frac{1}{N} \sum_{i=t-N+1}^{t} y_i$$

Where N is the window size

- **Purpose:** Helps reveal trends by reducing noise and can be used to predict future values based on recent averages.

3.1.3 Exponentially Weighted Moving Average (EWMA) : Unlike SMA, EWMA assigns exponentially decreasing weights to older observations, It's more responsive to recent data and captures trends with minimal lag.

- **Formula**

$$EWMA_t = \alpha y_t + (1 - \alpha)EWMA_{t-1}$$

Where α is the smoothing factor

- **Purpose:** Useful for series with significant trend or volatility, as it emphasizes recent data

3.2) Holt-Winters Model

The Holt-Winters model, or Triple Exponential Smoothing, extends EWMA by incorporating trends and seasonality.

- **Variants:**
  - **Additive**: For data where seasonal variations are roughly constant over time.
  - **Multiplicative**: For data where seasonal variations grow or shrink with the trend.

- **Formula:**

$$\hat{y}_{t+1} = (L_t + T_t)S_{t+1}$$

Where L_t is the level, T_t is the trend, and S_{t+1} is the seasonality component.

- **Purpose:** Useful for capturing seasonality and trends in time series data, particularly suited for financial indices with seasonal patterns.

3.3) ARIMA Model

ARIMA (AutoRegressive Integrated Moving Average) is a widely used statistical model for stationary time series, combining autoregressive (AR), differencing (I), and moving average (MA) components.

- **Key Components**
  - **Autoregressive (AR)**: Relates the current value to its past values.
  - **Integrated (I)**: Uses differencing to make the series stationary.
  - **Moving Average (MA)**: Relates the current value to past forecast errors.
- **Parameter Selection**
  - The parameters p, d, and q are chosen based on autocorrelation (ACF) and partial autocorrelation (PACF) analysis.
- **Purpose:** ARIMA is effective for capturing trends and patterns in stock prices, allowing us to forecast future values based on past behaviors.

3.4) ARIMAX Model

ARIMAX extends ARIMA by including exogenous variables (predictors) that may influence the time series, like economic indicators.

- **Formula**

$$y_t = \alpha + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \sum_{k=1}^{m} \beta_k X_{t,k} + \epsilon_t$$

Where $X_{t,k}$ are the exogenous variables.

- **Purpose:** Incorporates external factors that affect the FTSE 100, such as economic indicators, allowing a more comprehensive forecast.

3.5) SARIMA Model

Seasonal ARIMA (SARIMA) incorporates seasonal differencing and seasonal terms in ARIMA, which is useful when a series exhibits recurring patterns.

- **Formula**

$$SARIMA(p, d, q)(P, D, Q)_s$$

Where s is the seasonal period, and (P, D, Q) are seasonal counterparts to ARIMA parameters.

- **Purpose:** Effective for time series with both trend and seasonality, allowing seasonal cycles to be explicitly modeled.

3.6) SARIMAX Model

SARIMAX combines SARIMA with exogenous variables, capturing both seasonal patterns and external influences.

- **Formula**

$$y_t = \alpha + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \sum_{k=1}^{m} \beta_k X_{t,k} + \epsilon_t$$

- **Purpose:** Suitable for capturing complex financial series by accounting for both seasonal patterns and influential external variables.

4. Model comparison

In this section, we provide a comprehensive comparison of various forecasting models applied to the FTSE 100 dataset. Each model was evaluated for its performance in predicting both the closing prices and the returns of the FTSE 100 index. The models compared include the Naive model, Simple Moving Average (SMA), Exponentially Weighted Moving Average (EWMA), Holt-Winters Additive model, ARIMA, ARIMAX, SARIMA, and SARIMAX. To ensure an accurate assessment, we measured performance using Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) metrics. This analysis helps determine which model best captures the trends and seasonal patterns in the dataset and offers insights into their effectiveness in different scenarios.

4.1) Naive Forecast

The Naive Forecast model uses the most recent actual value as the forecast for the next period. While simplistic, this model serves as a baseline for comparison. The model's RMSE for predicting closing prices was 62.35, and the RMSE for returns was 0.016. Despite its simplicity, the Naive Forecast model provides a reasonable baseline, but it lacks the sophistication required to handle the complexities of financial data trends.

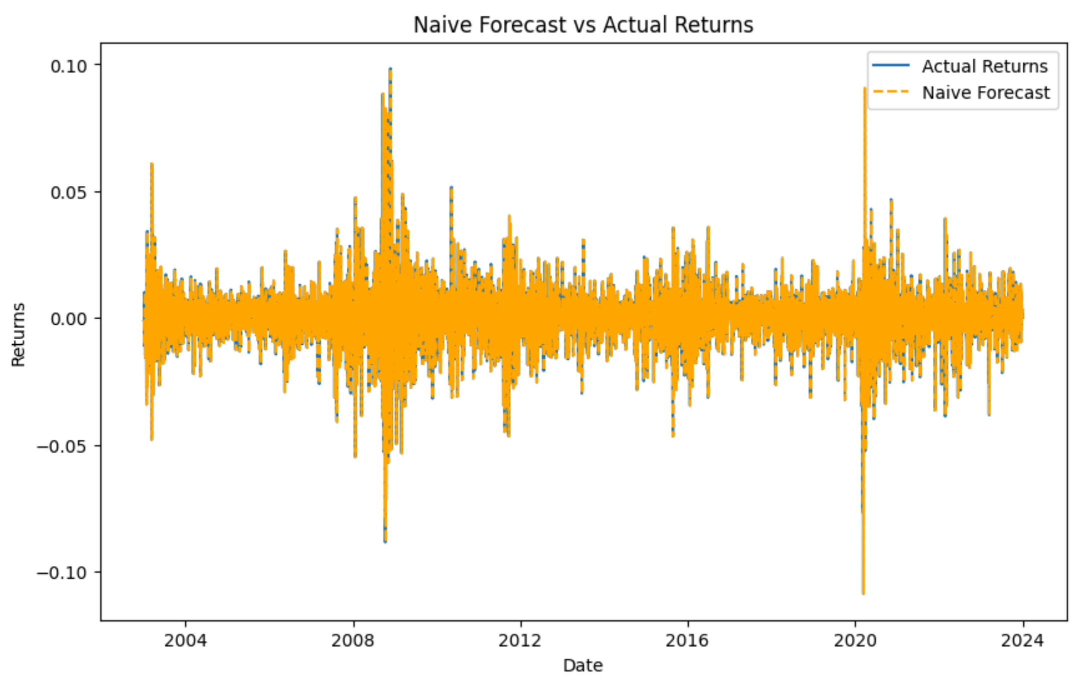

Figure 4.1.1 : *Naive Forecast vs Actual Returns*

This figure compares the performance of the Naive Forecast model with actual returns of the FTSE 100 index. The Naive Forecast uses the value from the previous day as the forecast for the next day, resulting in a close overlap of the two series, reflecting the model's simple assumption of market behavior.

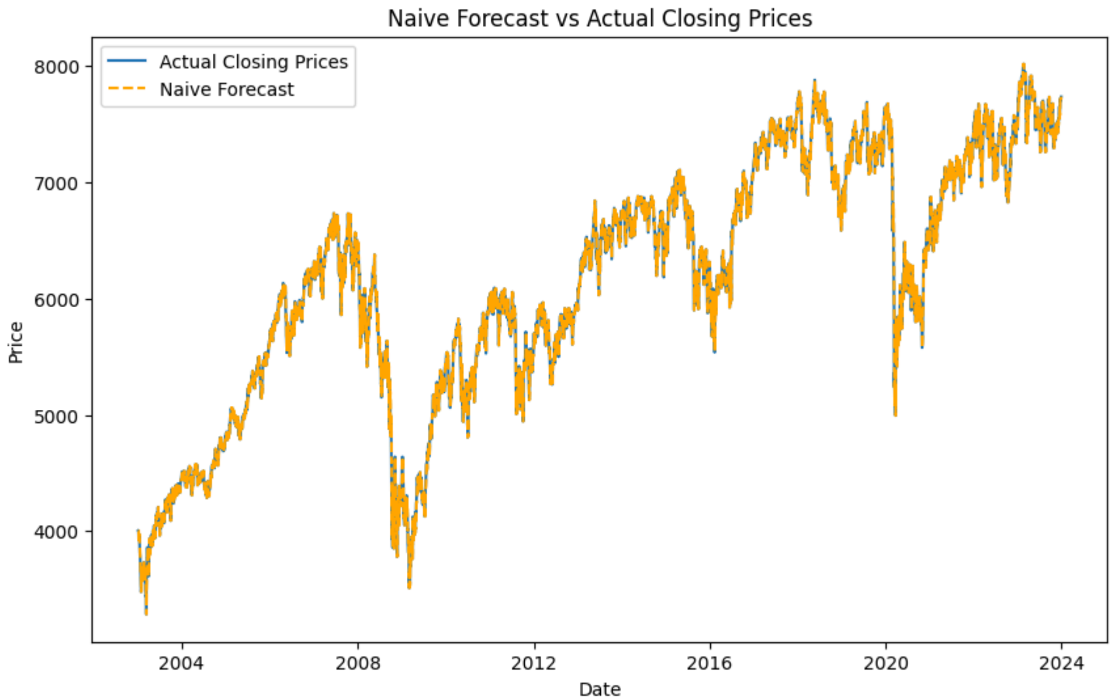

Figure 4.1.2: *Naive Forecast vs Actual Closing Prices*

This figure illustrates the comparison between the Naive Forecast model and actual closing prices of the FTSE 100 index. The Naive model assumes that the previous closing price will be the same as the forecasted price for the following day, which results in a lagging effect, especially during volatile periods.

4.2) Simple Moving Average (SMA)

The SMA model averages the closing prices (or returns) over a defined rolling window (in this case, 30 periods). While SMA smooths out short-term fluctuations, it struggles with fast-changing data and sudden market shifts. The RMSE for closing prices was notably higher than other models at 173.73, and for returns, it achieved an RMSE of 0.01095. This suggests that while SMA can smooth trends, it lacks the ability to capture sudden market movements.

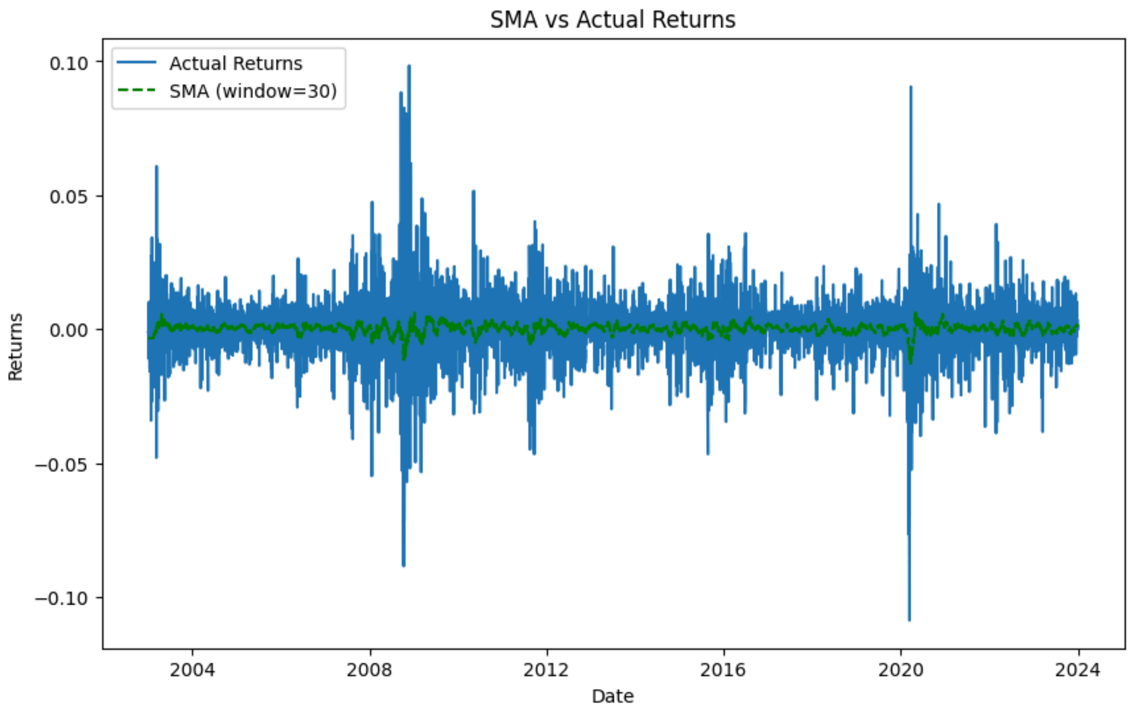

Figure 4.2.1 : *SMA vs Actual Returns*

This figure shows the comparison between the Simple Moving Average (SMA) model and actual returns. The SMA model averages the returns over a specified window (e.g., 30 days), which smooths out short-term fluctuations and helps highlight longer-term trends in returns.

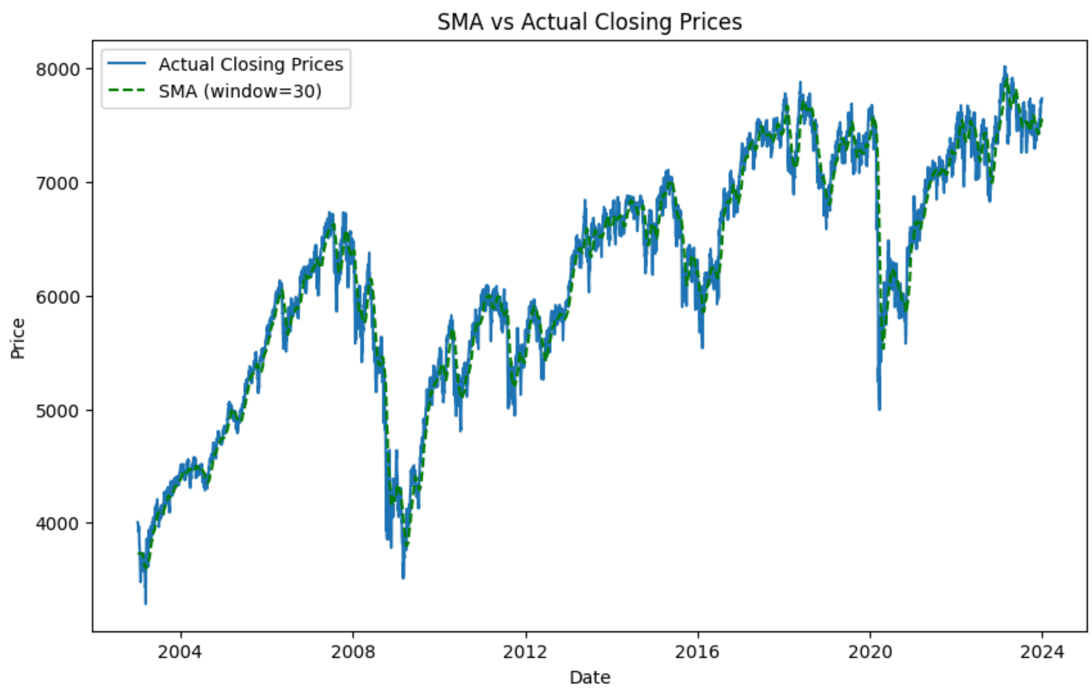

Figure 4.2.2 : *SMA vs Actual Closing Prices*

The comparison between the Simple Moving Average (SMA) model and actual closing prices is depicted in this figure. The SMA is useful for identifying trends by averaging the closing prices over a fixed window, but it can lag behind sudden market movements.

4.3) Exponentially Weighted Moving Average (EWMA)

The EWMA model applies exponentially decreasing weights to older observations, allowing the model to react more quickly to recent changes in data. This feature gives it a competitive edge over SMA. The RMSE for EWMA in predicting closing prices was 146.51, with a return RMSE of 0.01065, showing that the model balances trend-smoothing and adaptability to recent price movements better than SMA.

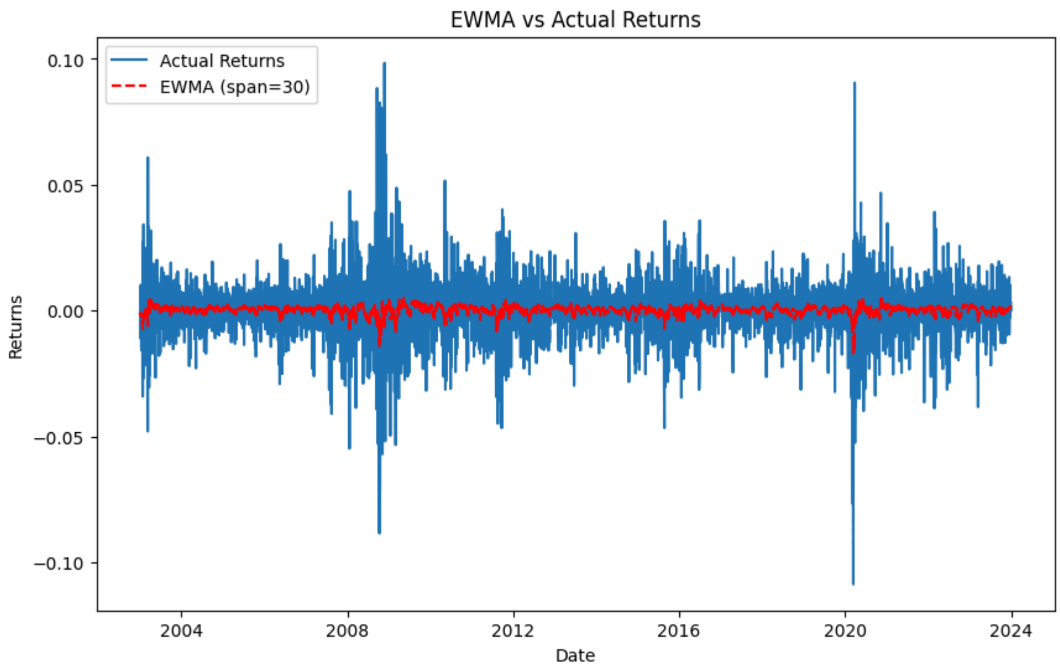

Figure 4.3.1 : *EWMA vs Actual Returns*

This figure shows the Exponentially Weighted Moving Average (EWMA) model compared to actual returns. The EWMA model places more weight on recent observations, making it more responsive to sudden changes in returns than the SMA model.

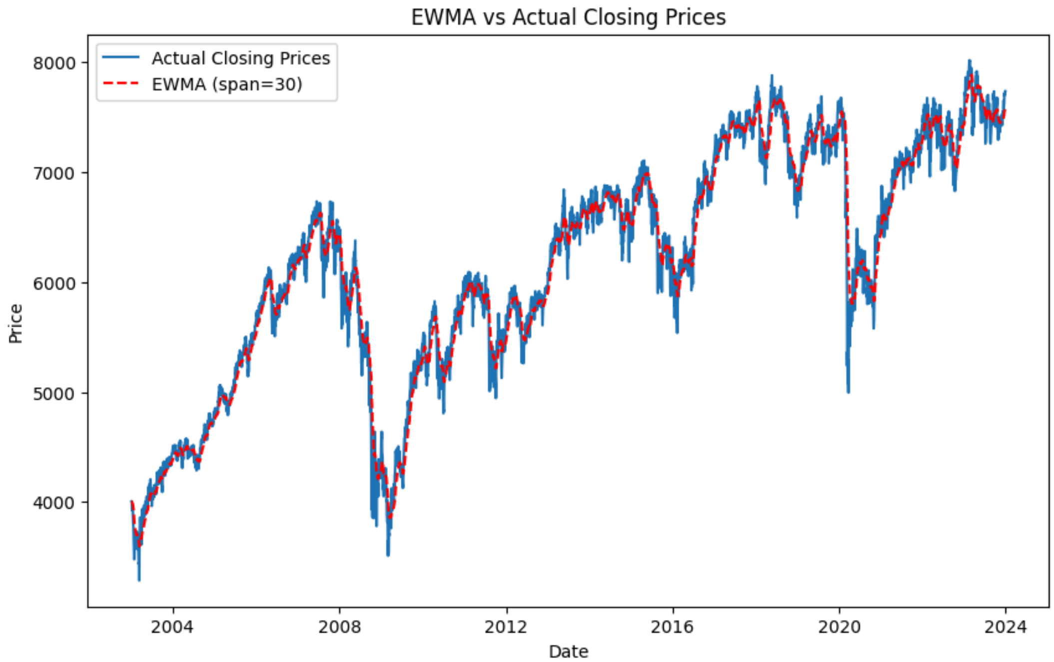

Figure 4.3.2 : *EWMA vs Actual Closing Prices*

This figure compares the EWMA model and actual closing prices. By weighting more recent observations more heavily, the EWMA model captures more timely price movements, providing a more responsive forecast than the SMA.

4.4) Holt-Winters Additive Model

The Holt-Winters Additive model is designed to account for trends and seasonality in the dataset. The model struggled to accurately predict returns, resulting in an RMSE of 0.0364 for returns, which was the highest across all models. However, for closing prices, it performed competitively with an RMSE of 62.41, demonstrating its ability to model seasonal components in the data.

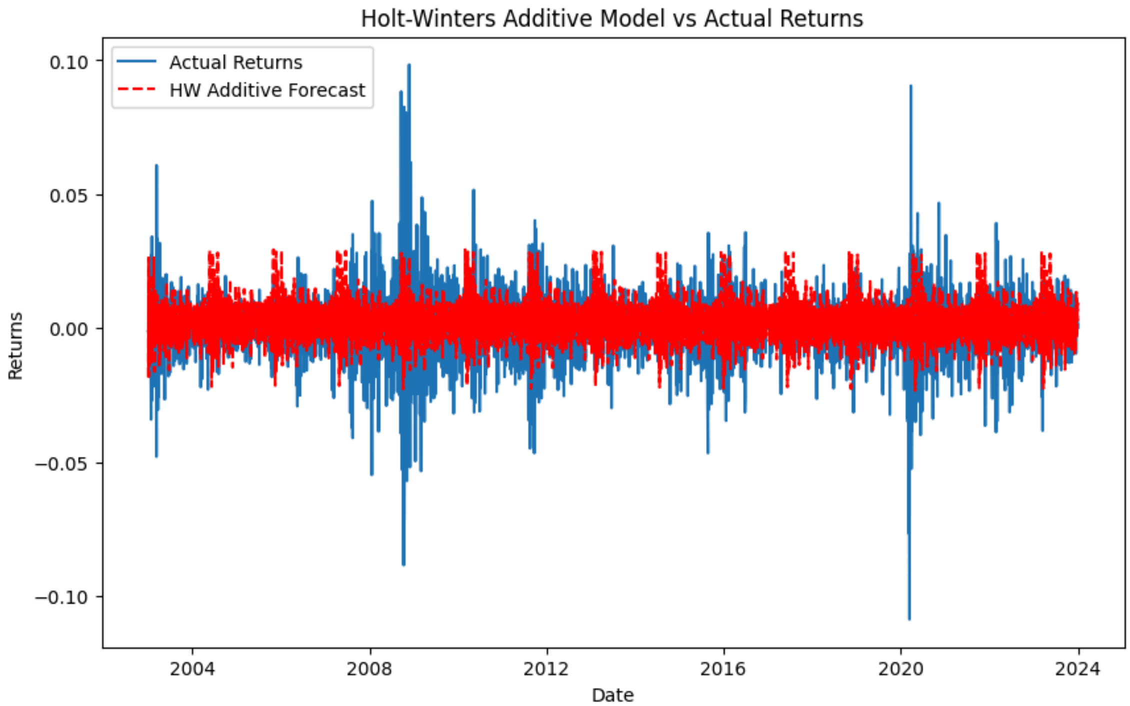

Figure 4.4.1 : *Holt-Winters Additive Model vs Actual Returns*

This figure compares the Holt-Winters Additive model's performance against actual returns. The Holt-Winters method considers level, trend, and seasonality in the returns, making it suitable for datasets with consistent seasonal fluctuations, as reflected in the FTSE 100 index returns.

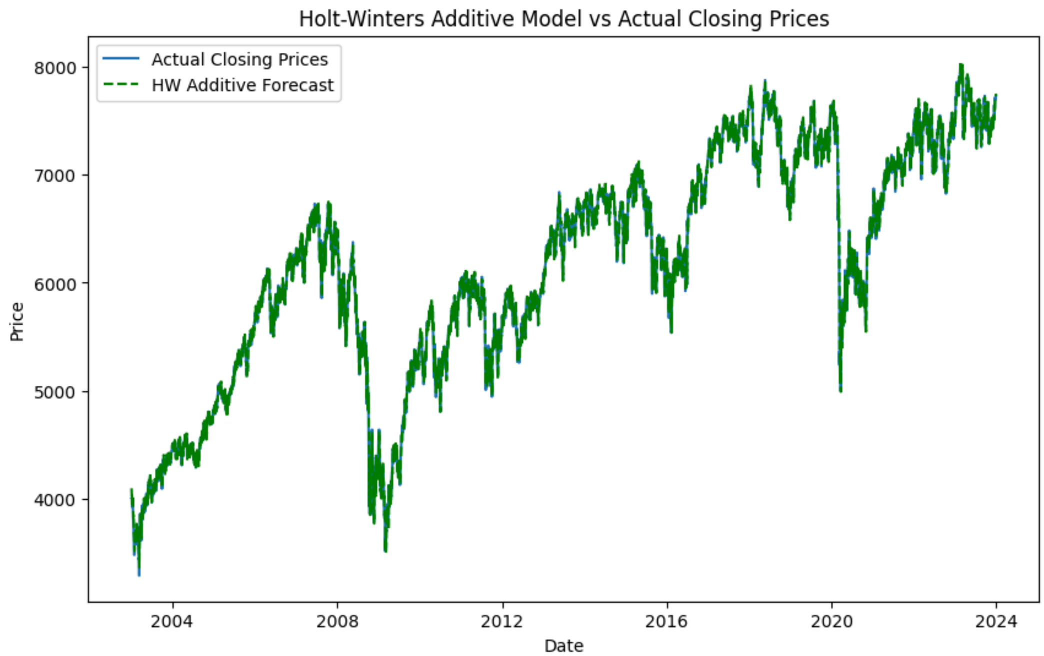

Figure 4.4.2 : *Holt-Winters Additive Model vs Actual Closing Prices*

This figure illustrates the performance of the Holt-Winters Additive model compared to actual closing prices. The model is designed to account for seasonal patterns, which are evident in the historical FTSE 100 closing prices.

4.5) ARIMA (AutoRegressive Integrated Moving Average)

The ARIMA model, commonly used in time series forecasting, combines autoregressive and moving average components with differencing to make the data stationary. ARIMA performed well, achieving an RMSE of 69.03 for closing prices and 0.01114 for returns. It shows the model's capability in forecasting non-seasonal data patterns, but it still underperformed compared to more complex models.

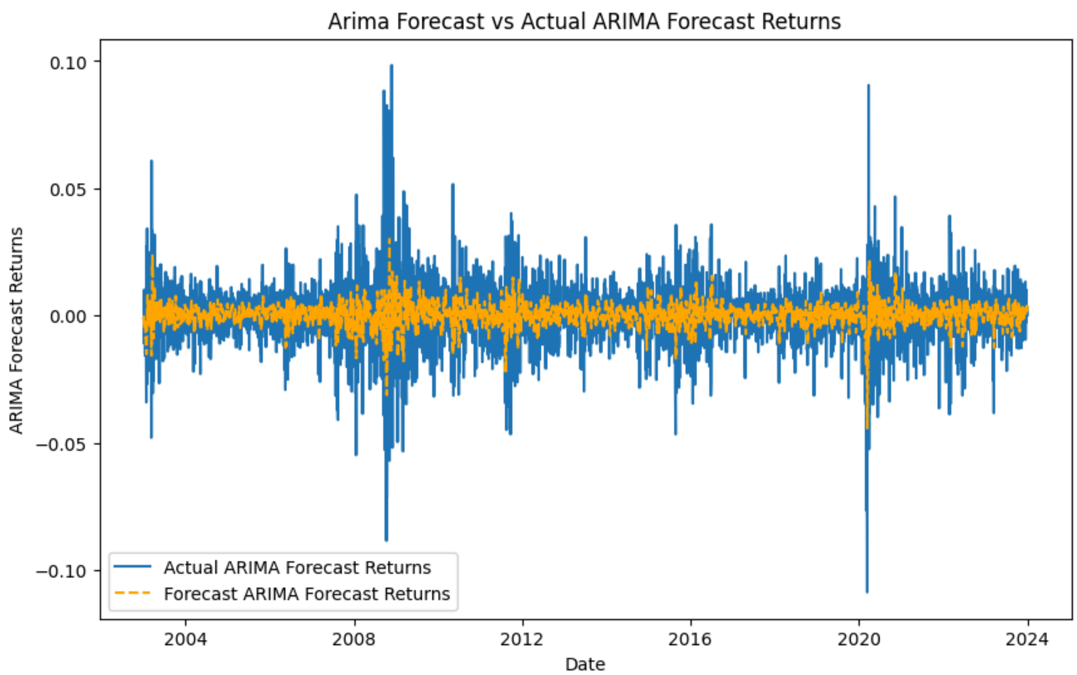

Figure 4.5.1 : *ARIMA Forecast vs Actual ARIMA Forecast Returns*

This figure compares the ARIMA model's forecast for returns with the actual ARIMA forecasted returns. It shows how closely the ARIMA model follows the actual behavior of the forecasted returns, with some deviation in periods of high volatility. The comparison highlights the ARIMA model's effectiveness in capturing general trends and patterns in the returns, though it may struggle with sharp fluctuations.

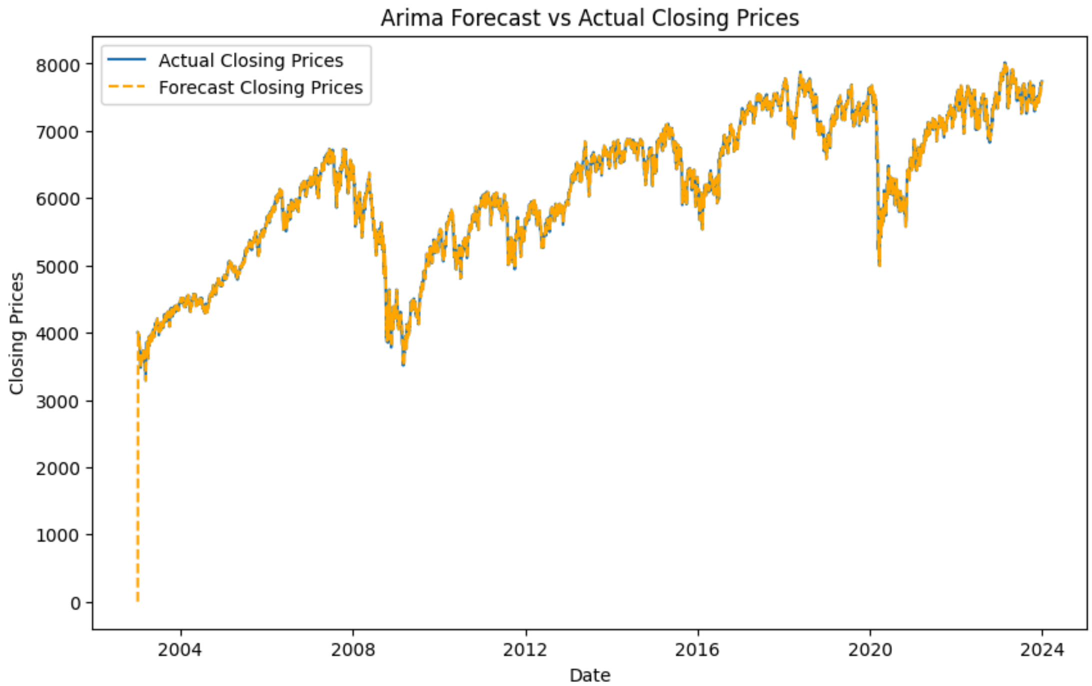

Figure 4.5.2 : *ARIMA Forecast vs Actual Closing Prices*

This figure shows the performance of the ARIMA model for forecasting closing prices. The model adjusts for both trend and seasonal effects, making it more robust compared to simpler models like Naive and SMA, particularly over long periods.

4.6) ARIMAX (ARIMA with Exogenous Variables)

The ARIMAX model builds on ARIMA by incorporating exogenous variables, in this case, a lagged closing price. The addition of external information can potentially improve forecasting. The model's RMSE for closing prices was 69.02, and for returns, it was 0.01113, slightly better than ARIMA. However, the improvement was marginal, suggesting that the chosen exogenous variable did not add substantial value.

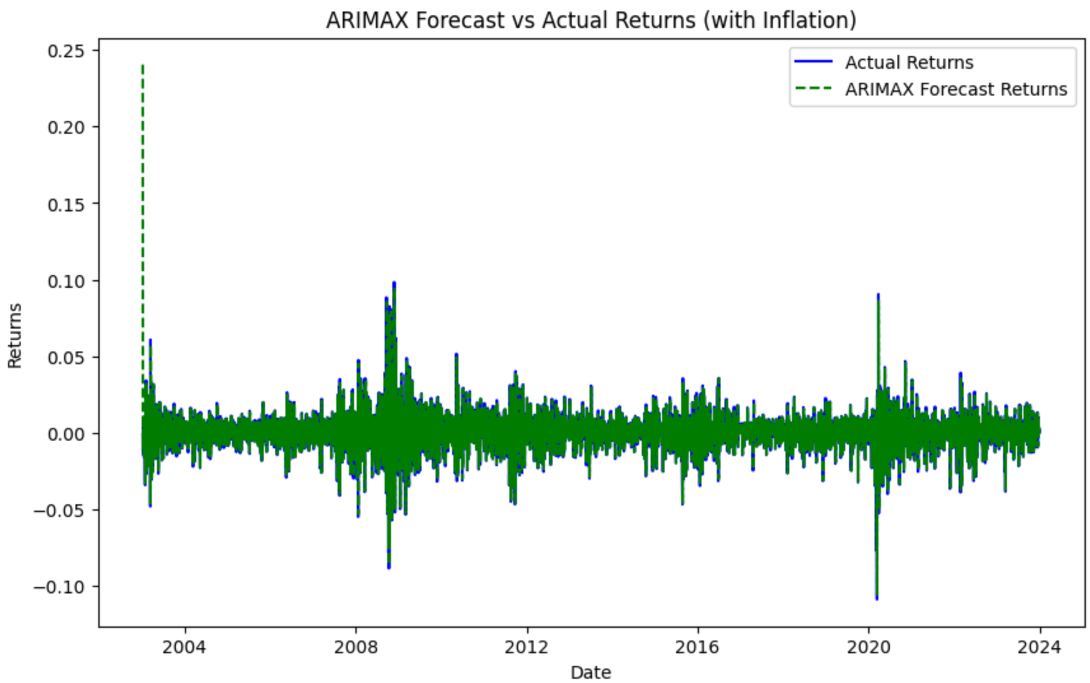

Figure 4.6.1 : *ARIMAX Forecast vs Actual Returns (with Inflation)*

This figure compares the ARIMAX (ARIMA with Exogenous Variables) model's forecast with actual returns, incorporating inflation as an exogenous variable. The inclusion of inflation provides additional context for price movements, offering a more comprehensive forecast of returns.

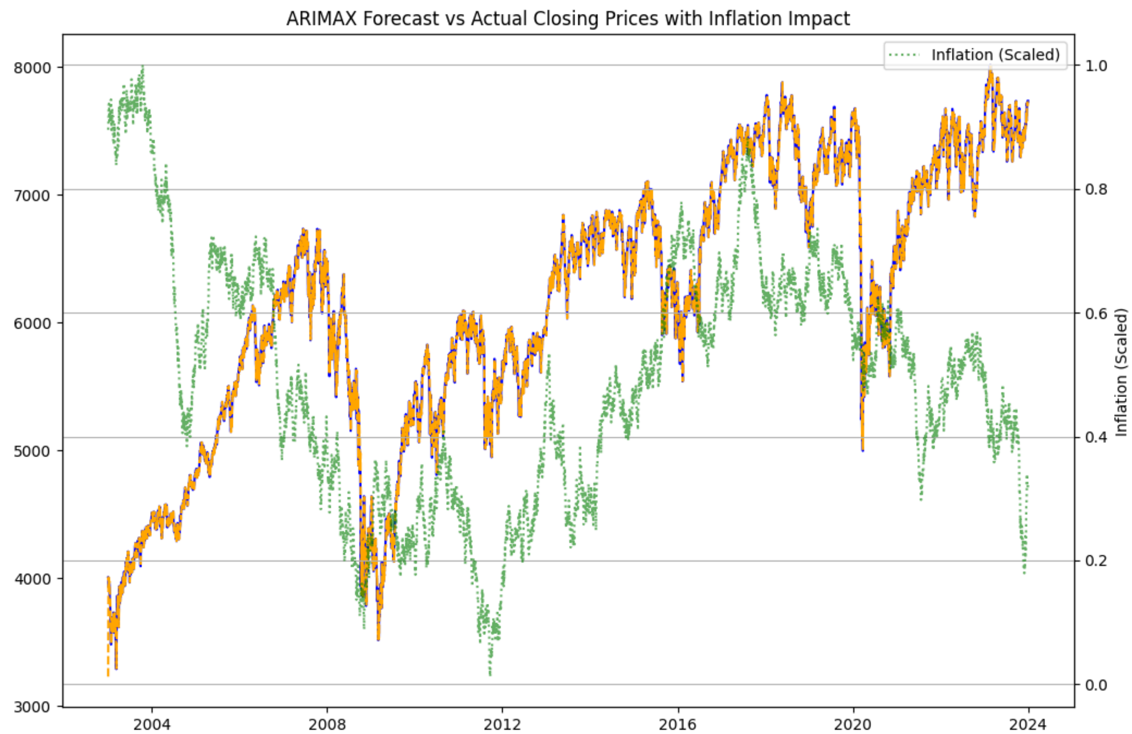

***Figure 4.6.2 :** ARIMAX Forecast vs Actual Closing Prices with Inflation Impact*

This figure illustrates the ARIMAX model's forecast of closing prices alongside inflation. The exogenous variable of inflation is scaled and plotted to show its correlation with FTSE 100 price movements, highlighting how external economic factors affect the index. However, the reason why ARIMAX=ARIMA is because of low inflation and we try to normalize it but the raw data shown as *figure 4.6.3*. So, it made our forecast not different from ARIMA.

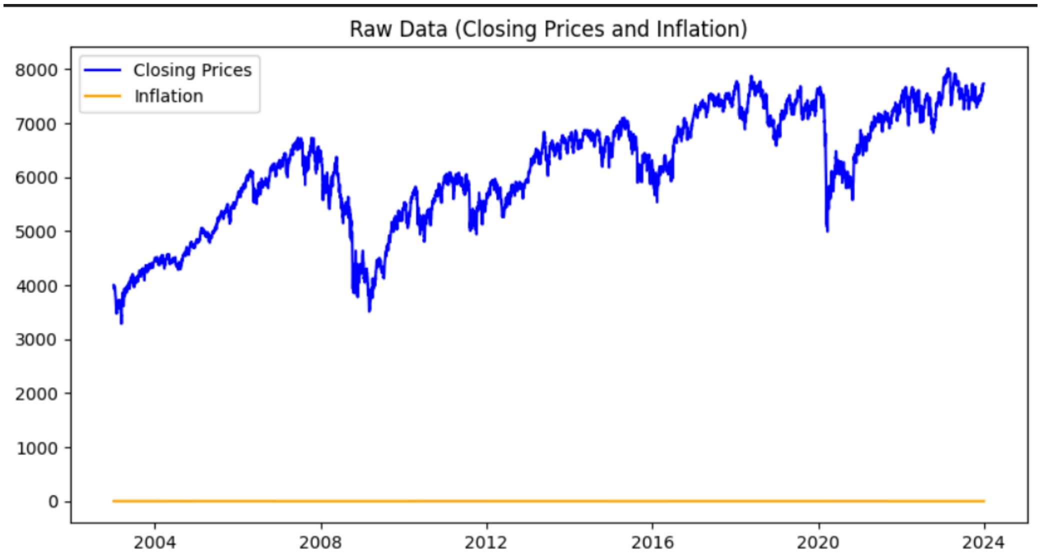

***Figure 4.6.3 :** Raw Data*

4.7) SARIMA (Seasonal ARIMA)

SARIMA, which incorporates seasonality into the ARIMA framework, performed moderately well. Its RMSE for closing prices was 97.37, higher than ARIMA, but the model showed comparable performance to ARIMA for returns with an RMSE of 0.01120. This indicates SARIMA's ability to capture both trends and seasonality, though the seasonal component might not have contributed significantly to the improvement for this dataset.

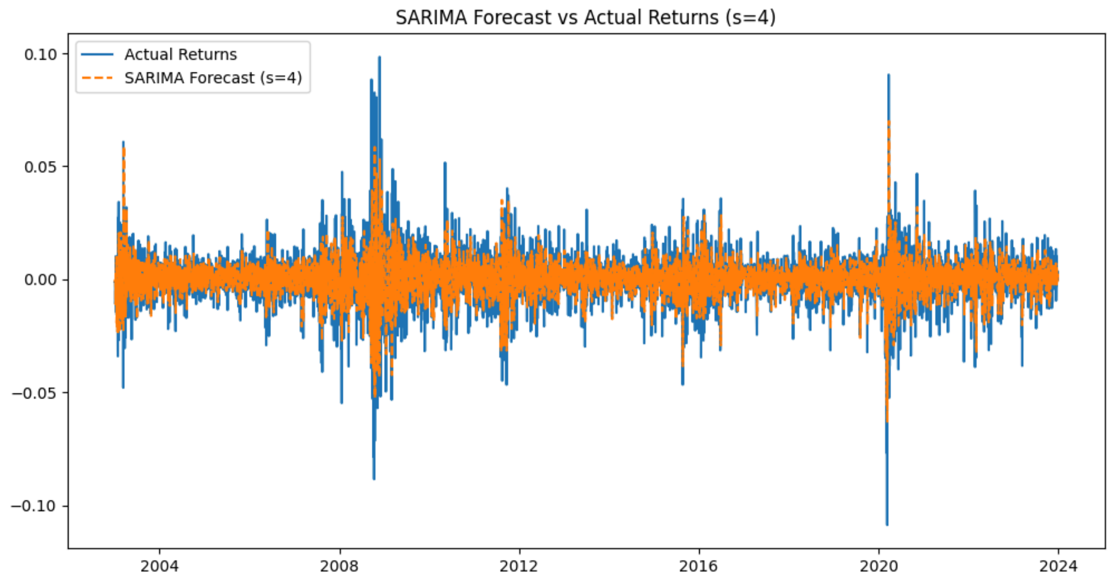

Figure 4.7.1 : *SARIMA Forecast vs Actual Closing Prices (s=4)*

This figure depicts the comparison between the SARIMA (Seasonal ARIMA) model and actual closing prices. The SARIMA model is particularly suited for time series with seasonal behavior, using a seasonal period of 4, which captures quarterly patterns in the data.

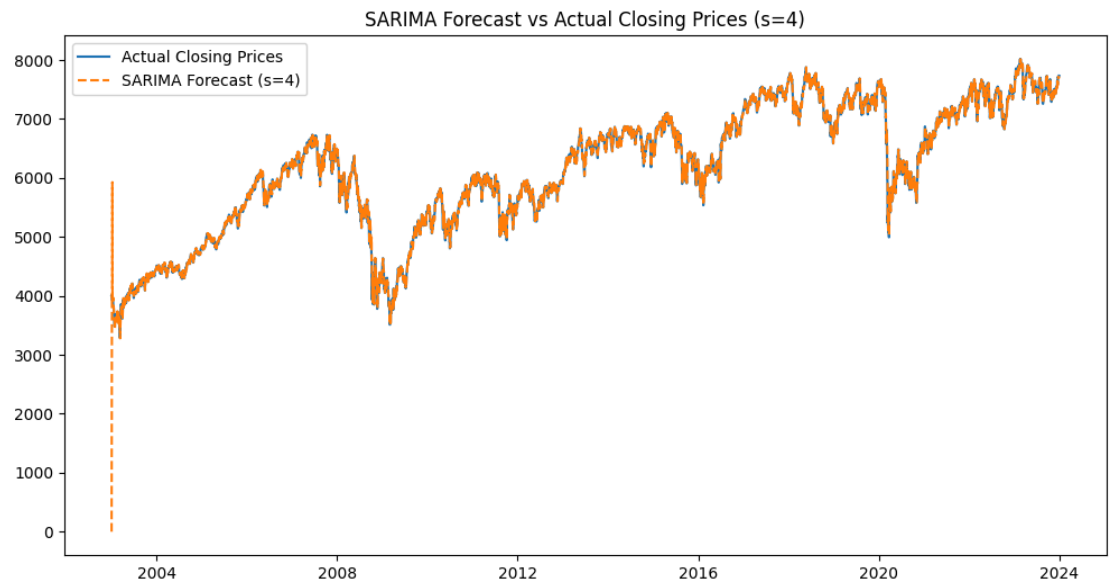

Figure 4.7.2 : *SARIMA Forecast vs Actual Returns (s=4)*

In this figure, the SARIMA model's forecast is compared to actual returns. The model captures both the seasonal and non-seasonal aspects of the time series, accounting for quarterly variations in the return data.

4.7.1) SARIMA Model : Plot ACF and PACF Graphs

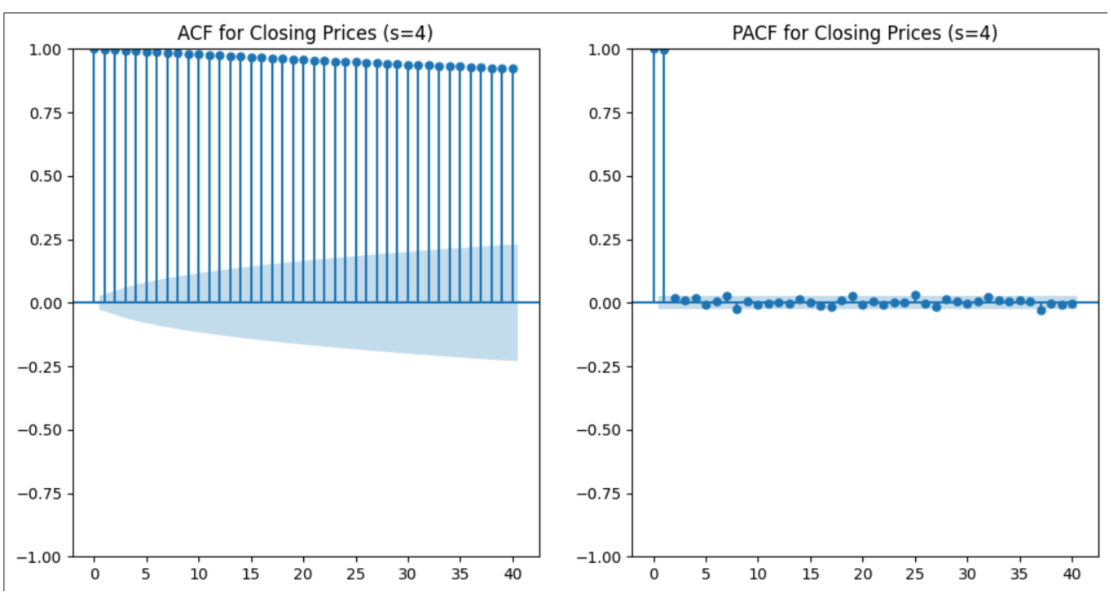

Figure 4.7.3 : *ACF and PACF for Closing Prices*

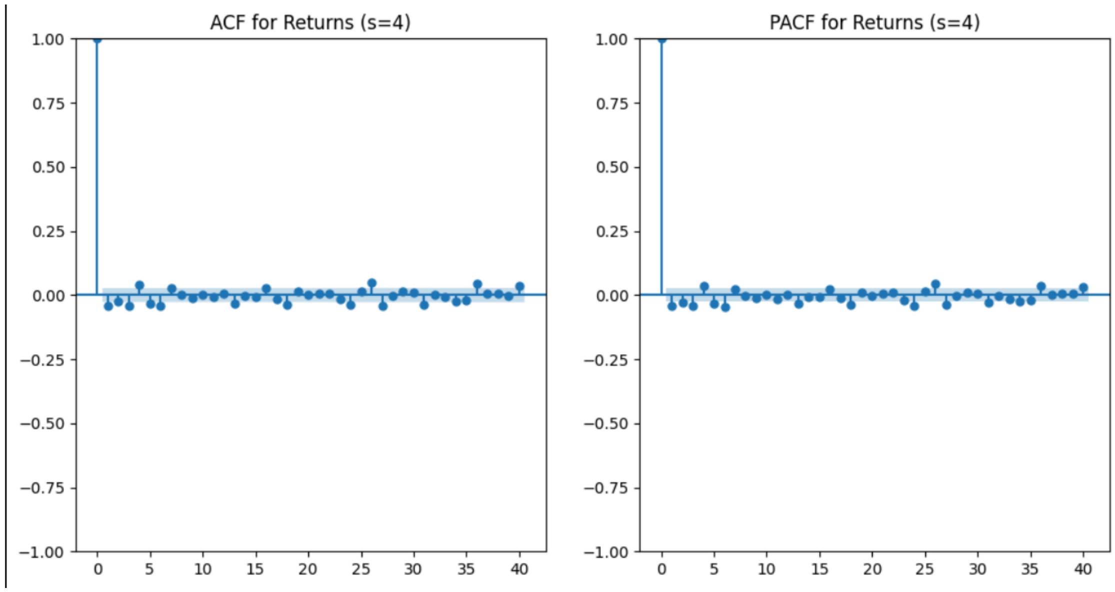

Figure 4.7.4 : *ACF and PACF for Returns*

4.8) SARIMAX (SARIMA with Exogenous Variables)

SARIMAX extends SARIMA by integrating external variables. Similar to ARIMAX, this model aimed to leverage additional data (e.g., inflation) for more accurate predictions. While the RMSE for closing prices was 97.36, close to SARIMA, the return RMSE was slightly lower at 0.0112. Despite these results, the exogenous variable did not dramatically improve performance.

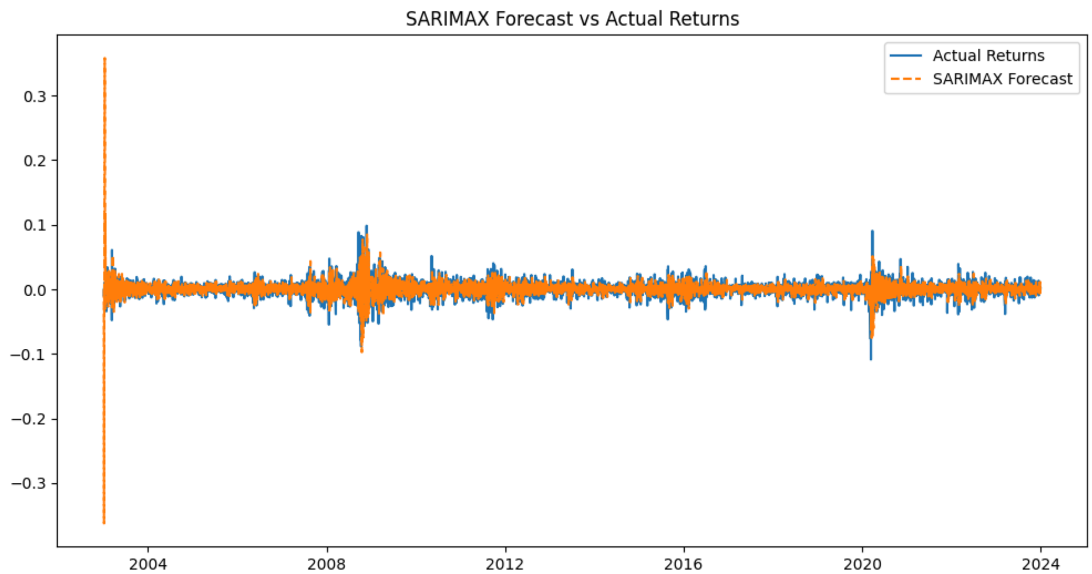

Figure 4.8.1 : *SARIMAX Forecast vs Actual Returns*

This figure compares the SARIMAX model's forecast with actual returns. The SARIMAX model extends SARIMA by incorporating exogenous variables, providing a more comprehensive view of returns in relation to external factors such as inflation.

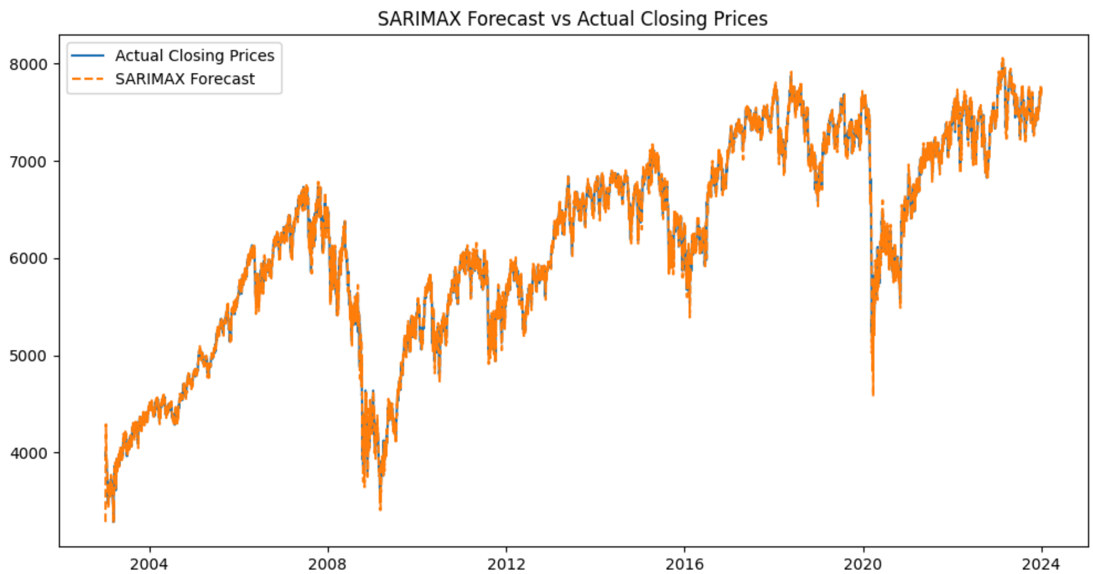

Figure 4.8.2 : *SARIMAX Forecast vs Actual Closing Prices*

This figure illustrates the SARIMAX model's forecast for closing prices, which accounts for seasonal patterns and external factors like inflation. The model captures price movements more accurately than non-seasonal models, making it suitable for long-term forecasting.

4.8.1) SARIMAX Model : Plot ACF and PACF Graphs

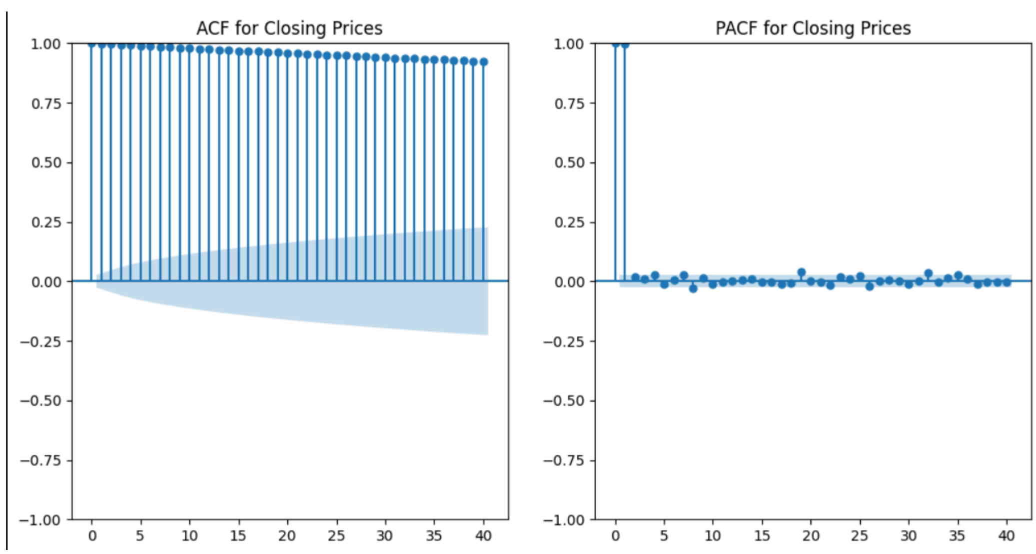

Figure 4.8.1 : *ACF and PACF for Closing Prices*

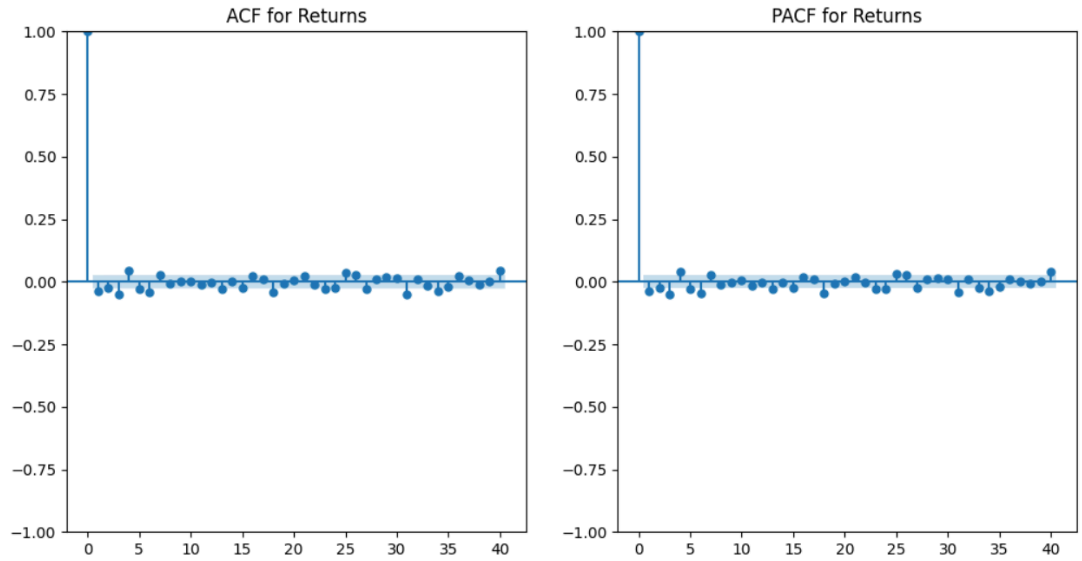

Figure 4.8.2 : *ACF and PACF for Returns*

4.9) Model Performance Comparison

The final step in our analysis was to visualize the RMSE performance of all models side by side, enabling a direct comparison. The bar charts in Figure 18 RMSE Comparison of Models for Returns and Figure 19 RMSE Comparison of Models for Closing Prices present a clear overview of each model's forecasting accuracy. From the visual analysis, it is evident that while no single model excelled across both metrics, simpler models like EWMA performed well for returns, and more complex models like ARIMA and SARIMAX offered competitive performance for closing prices.

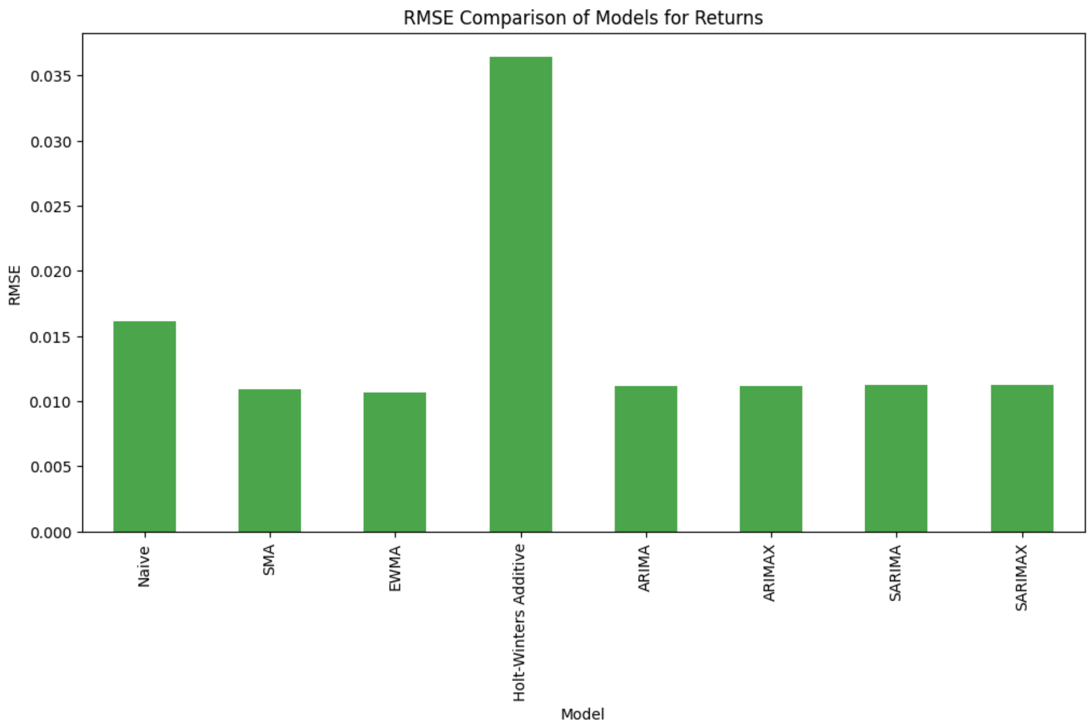

Figure 4.9.1 : *RMSE Comparison of Models for Returns*

This bar chart compares the Root Mean Squared Error (RMSE) of various models for forecasting returns. Lower RMSE values indicate better model performance. The figure highlights that the Holt-Winters Additive model performs poorly, while models like ARIMA, ARIMAX, and SARIMAX provide more accurate return forecasts.

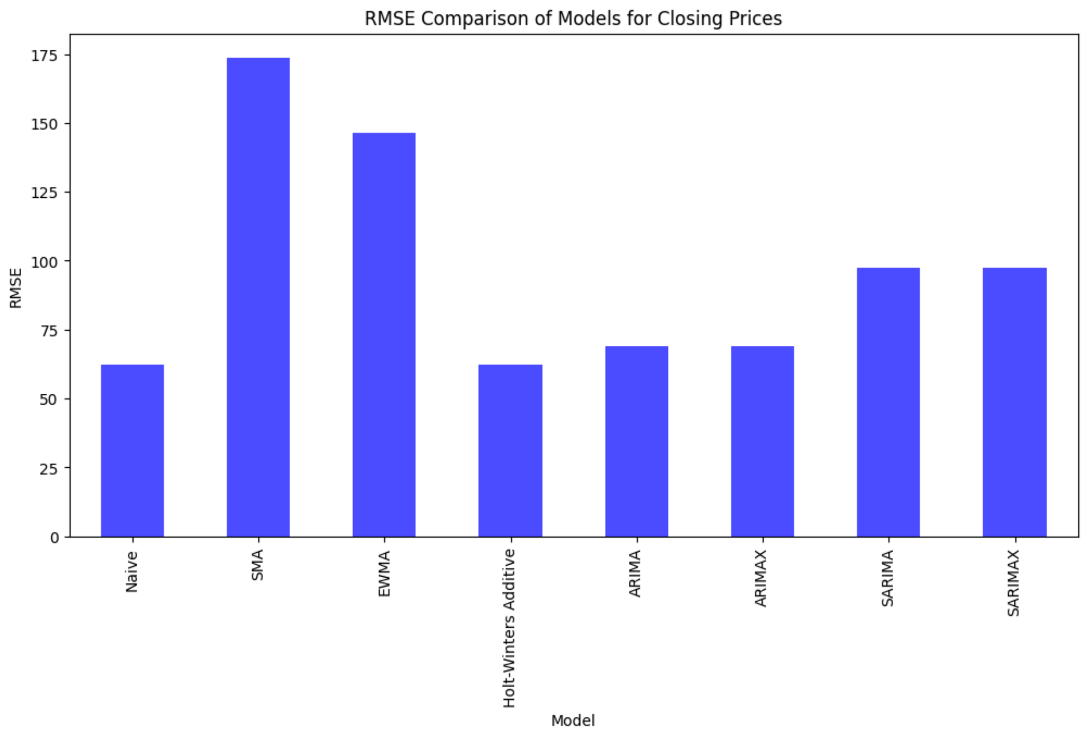

Figure 4.9.2 : *RMSE Comparison of Models for Closing Prices*

This bar chart compares the RMSE of different models for predicting closing prices. The comparison reveals that the Naive model performs reasonably well, while models like SMA and EWMA show higher RMSE due to their lagging nature. The ARIMA and SARIMA models demonstrate competitive performance in capturing trends and seasonality in the data.

:

5. Justification and Application in Financial Engineering

Financial engineering relies heavily on the application of advanced quantitative methods and models to solve complex problems in finance. Time series forecasting plays a crucial role in areas like risk management, derivative pricing, and portfolio optimization. In this section, we will justify the selection of the various forecasting models used in this study and explore their practical applications in financial engineering, particularly in forecasting returns and closing prices of financial assets like the FTSE 100 index.

5.1) Model Justification

Each model used in this analysis brings unique strengths, and their application in financial engineering can vary based on their forecasting capability, computational efficiency, and the nature of the financial data.

- **Naive Model**: While simplistic, the Naive model serves as a baseline in financial forecasting. It assumes that the most recent value is the best predictor of the next. This model is often used in financial markets where no clear trend or seasonality is evident, and investors may use recent prices as proxies for future prices, especially in short-term trading strategies. However, the limited accuracy demonstrated by this model, particularly in volatile markets, justifies its minimal use for more sophisticated financial applications.
- **Simple Moving Average (SMA)**: The SMA model smooths out short-term fluctuations and highlights longer-term trends, making it useful in identifying general market movements. In technical analysis, moving averages are widely applied to determine price direction, identify support and resistance levels, and create trading signals such as "golden crosses" and "death crosses." However, as observed in the model comparison, its lagged nature reduces its effectiveness for predictive accuracy, which limits its usage in real-time financial decision-making.
- **Exponentially Weighted Moving Average (EWMA)**: The EWMA model gives more weight to recent observations, which is critical in financial markets where recent data often have greater relevance. Its ability to quickly adapt to changing market conditions makes it applicable in volatility forecasting models, such as those used in options pricing (e.g., GARCH models) or risk management strategies like Value at Risk (VaR). The performance of the EWMA model in this study highlights its advantage over simpler moving averages, particularly in capturing rapid market shifts.
- **Holt-Winters Additive Model**: The Holt-Winters model's ability to capture both trend and seasonality makes it valuable in predicting assets with known cyclical behaviors, such as certain commodities or economic indices. However, the significant overestimation of volatility observed in the returns forecast using this model (Figure 7, Holt-Winters Additive Model vs Actual Returns) suggests that its applicability is limited in highly volatile financial markets, where seasonality might not be as pronounced.
- **ARIMA Model**: ARIMA is one of the most widely used models in financial forecasting due to its ability to handle non-stationary data by incorporating differencing and lagged observations. Its application is particularly strong in predicting financial time series with no clear seasonality, such

as stock prices or returns. In this study, the ARIMA model demonstrated good accuracy in forecasting both closing prices and returns (Figure 9, Arima Forecast vs Actual Closing Prices and Figure 10, ARIMA Forecast vs Actual ARIMA Forecast Returns). This model is frequently used in pricing models for bonds and derivatives, where understanding price movements over time is critical.

- **ARIMAX Model**: The ARIMAX model extends ARIMA by incorporating exogenous variables, which can improve forecast accuracy when additional factors, such as macroeconomic indicators (e.g., inflation rates), influence the asset prices. The inclusion of inflation data in this study, for instance, helped capture broader market dynamics in the returns forecast (Figure 11, file: ARIMAX Forecast vs Actual Returns (with Inflation)). ARIMAX models are particularly useful in scenarios where external economic factors significantly affect asset performance, such as in bond pricing or interest rate forecasting.
- **SARIMA and SARIMAX Models**: Seasonal ARIMA (SARIMA) and SARIMA with exogenous variables (SARIMAX) are designed to handle data with both non-stationarity and seasonality, making them highly relevant for assets with periodic behaviors, such as energy commodities or agricultural products. In financial engineering, SARIMA models are often applied in predicting seasonal variations in stock indices or commodity prices. The results from SARIMA and SARIMAX models in this study showed effective performance in capturing both long-term trends and short-term fluctuations (Figure 14, file: SARIMA Forecast vs Actual Closing Prices (s=4) and Figure 16, file: SARIMAX Forecast vs Actual Closing Prices). This makes them suitable for applications such as managing seasonal risk in portfolios or pricing seasonal derivatives.

5.2) Application in Risk Management and Forecasting

- **Risk Management**: Forecasting models like ARIMA and SARIMAX are integral to risk management in finance. They help predict future price movements, which are used to assess the risk of financial assets. For example, Value at Risk (VaR) models, which are crucial in portfolio risk management, often rely on ARIMA-type models to forecast the potential losses in a portfolio. By incorporating external variables like interest rates or inflation, ARIMAX and SARIMAX models can offer a more comprehensive risk assessment, particularly in macro-sensitive assets such as bonds or currencies.
- **Derivative Pricing**: Accurate forecasting models are essential in derivative pricing, where the value of options, futures, or swaps depends on the predicted price of the underlying asset. Models like ARIMA and SARIMAX are used in pricing equity and commodity derivatives, where time series data of asset prices are modeled to forecast future price trajectories. Additionally, these models can be employed in dynamic hedging strategies, where the hedge position is adjusted as new price data is forecasted.
- **Portfolio Optimization**: Portfolio managers use time series forecasting models to optimize asset allocation by predicting future returns and risks. By incorporating forecasting models like SARIMAX, which include exogenous factors such as inflation or interest rates, portfolio managers can adjust their portfolios based on anticipated macroeconomic conditions. This ensures that portfolios are hedged against adverse economic shifts, improving long-term performance and stability.

- **Interest Rate and Bond Yield Forecasting**: ARIMA and ARIMAX models are widely applied in forecasting interest rates and bond yields, which are critical in fixed-income markets. These models help central banks, institutional investors, and corporate treasuries predict future rate movements, assisting in decisions regarding bond issuance, refinancing, and investment.
- **Volatility Forecasting**: The ability of models like EWMA and GARCH (generalized autoregressive conditional heteroskedasticity) to forecast volatility is fundamental to derivative pricing, risk management, and market-making. These models are particularly effective when recent data points are more relevant to future volatility, as demonstrated in the EWMA model's performance in capturing market fluctuations (Figure 6, file: EWMA vs Actual Returns).

6. Conclusion

This report has explored and evaluated various forecasting models applied to the FTSE 100 index, focusing on both closing prices and returns. The analysis spanned a range of models, from baseline techniques such as the Naive and Simple Moving Average (SMA) models to more advanced methods like ARIMA, SARIMA, and SARIMAX, which incorporate both time series dynamics and exogenous variables.

Each model demonstrated unique strengths and limitations, with the more complex models generally providing better predictive performance. The ARIMA and SARIMAX models, in particular, showed strong accuracy for both closing price and return forecasting. The addition of exogenous variables, such as inflation in the ARIMAX and SARIMAX models, allowed for improved predictive accuracy by considering external economic factors that influence market behavior.

- **Naive and SMA Models**: While simple and computationally efficient, these models were outperformed by more advanced methods, particularly in volatile or non-stationary markets.
- **EWMA and Holt-Winters Models**: These models captured short-term fluctuations more effectively than the simpler techniques, but their performance was limited in scenarios involving long-term trends or significant market volatility.
- **ARIMA and ARIMAX Models**: These models demonstrated strong forecasting accuracy, particularly for non-seasonal time series, and were highly applicable for assets influenced by external economic factors.
- **SARIMA and SARIMAX Models**: The seasonal extensions of ARIMA performed well in capturing cyclical behavior in the data, making them suitable for financial assets with seasonal variations. The inclusion of exogenous variables further enhanced the SARIMAX model's performance, making it the most versatile and accurate model in the analysis.

From a financial engineering perspective, the models evaluated in this report have wide-ranging applications in risk management, derivative pricing, and portfolio optimization. The predictive power of the ARIMAX and SARIMAX models, particularly in capturing the effects of exogenous variables, highlights their practical utility in financial forecasting, especially when anticipating market movements influenced by macroeconomic factors such as inflation.

In conclusion, the results of this study underscore the importance of selecting the appropriate forecasting model based on the financial asset and the specific market conditions. The integration of external variables into time series models provides enhanced accuracy, which is critical in high-stakes financial decision-making processes. The SARIMAX model, in particular, stands out as the most comprehensive and reliable tool for time series forecasting in the financial domain, offering valuable insights for both theoretical research and practical applications in financial engineering.