

Hayden Fargo

Tyler Lewis

Devyn Miller

Ponthea Zahraii

Cryptocurrency Price Analysis

Introduction and Background

The cryptocurrency market has experienced marked growth over the past decade, becoming a significant player in the global financial ecosystem. Bitcoin (BTC), the first and most prominent cryptocurrency, often leads the broader market, influencing the price dynamics of altcoins such as Ethereum (ETH), Solana (SOL), Binance Coin (BNB), and Ripple (XRP). Understanding how Bitcoin's historical prices affect the predictive accuracy of altcoin prices can benefit stakeholders involved in trading, investment, and risk management.

The dataset utilized in this analysis comprises historical daily cryptocurrency data from the CryptoCompare API, which included approximately 2,000 days of trading data. The dataset consists of daily metrics such as opening and closing prices, high and low prices, trading volume, and volatility indicators for each cryptocurrency- Bitcoin and the selected altcoins. The data was cleaned to address missing values and standardize timestamps across different cryptocurrencies. Daily average prices were calculated by averaging the high and low values to mitigate intraday volatility and provide a more stable metric for analysis.

Research Questions

Q1: How do macro factors, such as M2 money supply, the Fear & Greed Index, and alternative risk assets, influence short-term Bitcoin price movements, and can these factors predict Bitcoin's price trend?

Q2: How well can an ARIMA model predict the short-term price movements of altcoins (e.g., Ethereum, Litecoin, Sui, and Solana) based on historical Bitcoin price data, and which altcoin follows Bitcoin the most?

Q3: Does including past Bitcoin prices improve the predictions of current altcoin prices?

Q4: How does Bitcoin's volatility today affect altcoin volatility?

Question 1

Question 1 Analysis

Our initial research question aims to assess how macro factors such as M2 money supply, the Fear & Greed Index, and alternative risk assets correlate with short-term Bitcoin price movements and whether these factors may be valuable in predicting Bitcoin's price trend.

We identified seven potential indicators that we believed might correlate with Bitcoin's price. These indicators were chosen for their inherent alignment with cryptocurrency markets or their known correlation with macro trends.

We begin by conducting a detailed correlation analysis between the identified indicators, seeking trends between lagged snapshots of the data and the price of Bitcoin, which will be crucial in building a predictive model.

The sample data was accumulated via the following sources for Feb 2018 - Dec 2024:

Dependant variable:

- Bitcoin spot price —————(Binance)

Independent variables:

1. Bitcoin market volumes —————(Binance)
2. Bitcoin perpetual funding rates ———(Binance)
3. M2 Supply —————(FRED)
4. USD Strength Index (DXY)———(FRED)
5. Gold spot price —————(Yahoo Finance)
6. Gold futures price —————(Yahoo Finance)
7. Cryptocurrency Fear/Greed index——(Kaggle)

The data was mostly reduced to interval variances to efficiently summarize temporal movement and included offsets of the data by 1, 2, and 4 weeks. We aimed to determine if the past performance of our identified variables could act as an indicator of BTC price. Each correlation pair was computed with the following function and was used to produce the subsequent heat map figure:

$$r = \frac{\Sigma(x_i - x_\mu)(y_i - y_\mu)}{\sqrt{\Sigma(x_i - x_\mu)^2 \Sigma(y_i - y_\mu)^2}}$$

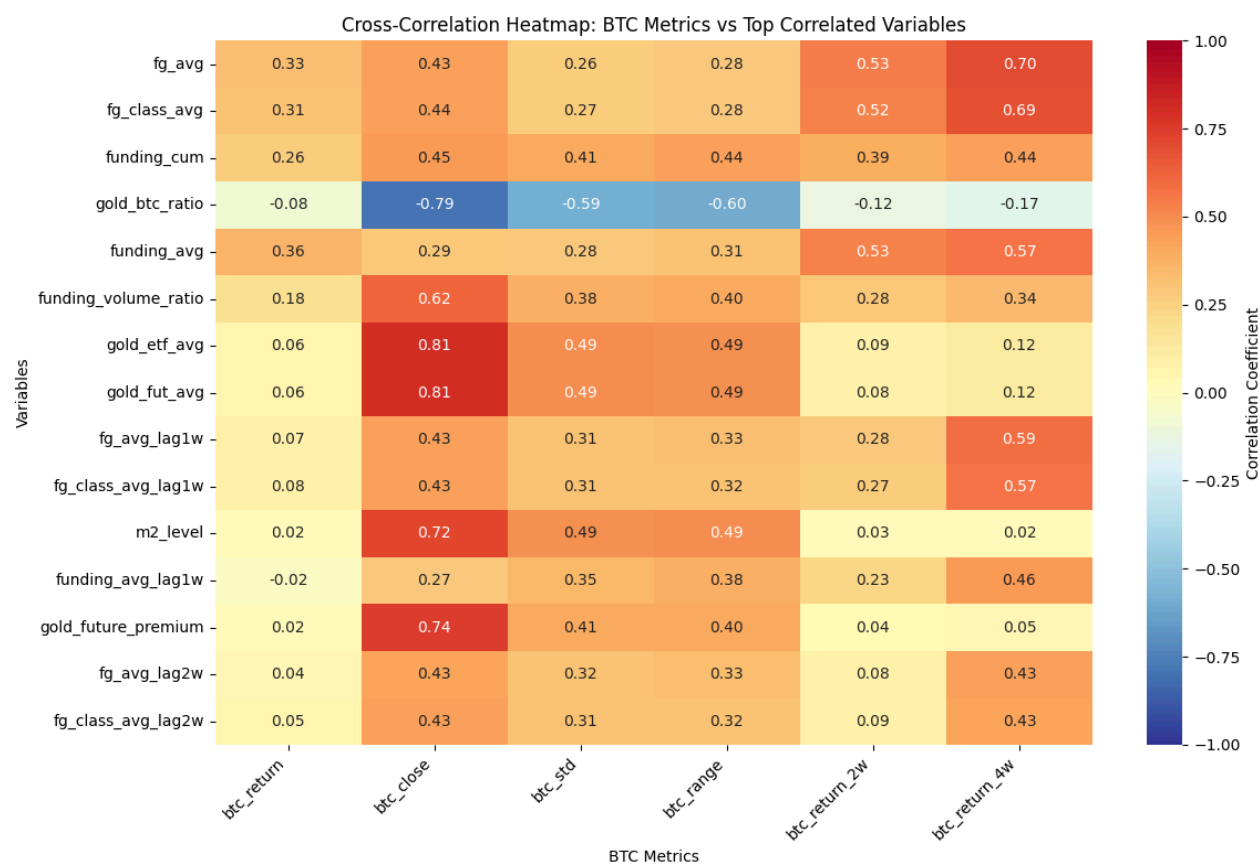


Figure 1.1: Pearson correlation heatmap between BTC price and identified indicators.

Our correlation analysis reveals relationships between Bitcoin's price dynamics and various market indicators across different timeframes. The analysis of weekly returns demonstrates moderate positive correlations with funding rates and Fear & Greed metrics and gold prices, suggesting a potential alignment in Bitcoin price movements and these two indicators. This observation is further supported by a strong positive relationship with the M2 money supply, indicating Bitcoin's responsiveness to monetary expansion.

Volatility measures, including standard deviation and price range, show moderate correlations (0.4-0.5) with both gold prices and M2 levels while maintaining consistent relationships with funding rates. This suggests that derivatives market activity may play a significant role in price volatility. When examining longer timeframes, the correlations become more pronounced.

Two-week returns appear to demonstrate stronger relationships with sentiment metrics compared to weekly returns, with Fear & Greed and funding rates showing correlations exceeding 0.5. This pattern becomes even more evident in the four-week returns, which exhibit very strong correlations (>0.69) with Fear & Greed metrics, including a notable memory effect where lagged Fear & Greed values maintain correlations above 0.5.

Question 1 Results

Our data analysis revealed correlation between Bitcoin price and fear/greed sentiment and gold markets in particular. To delve deeper into these correlations, we aim to produce results of the potential of these variables acting as short-term indicators of Bitcoin price movements, performing further analysis on higher resolution time frames.

1. Fear & Greed Index (F&G):

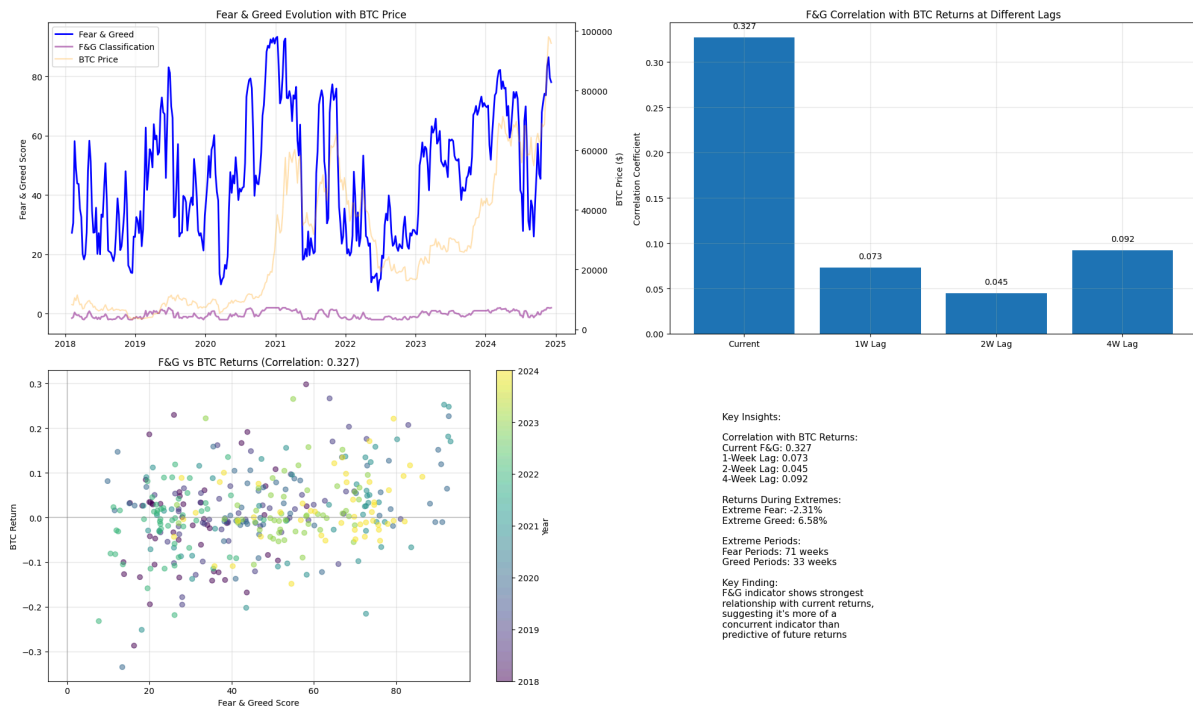


Figure 1.2: The Fear & Greed Index (F&G) shows its strongest correlation with BTC returns at the current value, with diminishing correlations over 1, 2, and 4-week lags.

The Fear & Greed Index (F&G) demonstrates a moderate correlation with Bitcoin returns, strongest at the current value (0.327) and quickly diminishing at 1, 2, and 4-week lags (0.073, 0.045, and 0.092, respectively). While extreme sentiment shows clear trends—extreme fear aligns with a -2.31% decline and extreme greed with a 6.58% gain—these results emphasize F&G as a concurrent sentiment indicator rather than a forward-looking tool. The prolonged duration of fear periods (71 weeks) compared to greed periods (33 weeks) highlights its ability to reflect long-run market indications but does little to predict price movements. For our goal of identifying predictive indicators, the F&G Index offers limited utility, as it primarily captures current market psychology without providing actionable signals for anticipating price trends.

2. Gold markets

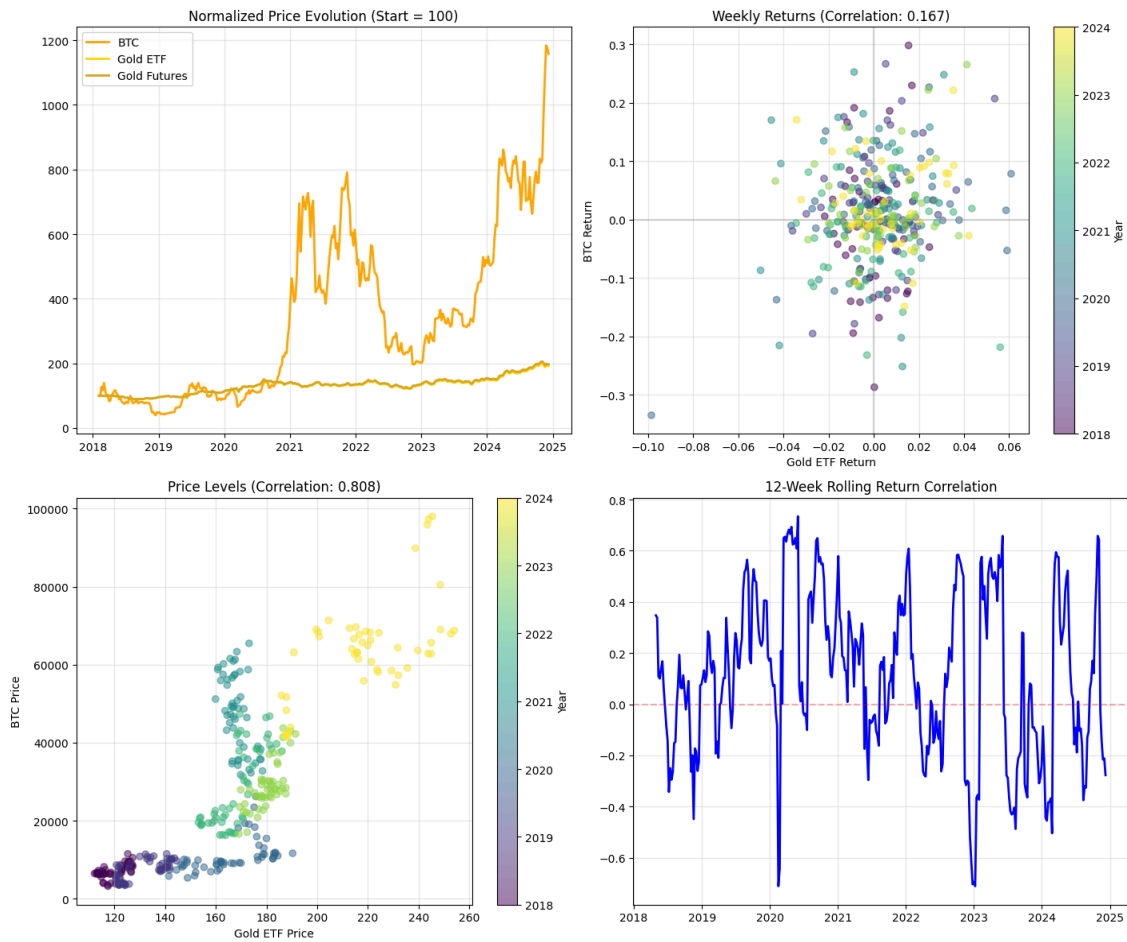


Figure 1.3: Relationship between Bitcoin (BTC) price and gold, including price level correlation, weekly return correlation, and the 12-week rolling return correlation.

The analysis of the relationship between Bitcoin and gold reveals a strong correlation in price levels, particularly post-2020, with a correlation coefficient of 0.808. This suggests that Bitcoin and gold move similarly over longer timeframes, implying that Bitcoin may act as a store of value similar to gold. However, when examining weekly returns and the 12-week rolling return correlation, the relationship weakens significantly, with a low correlation of 0.167 and fluctuating returns. These findings suggest that while gold's price can provide some insight into Bitcoin's long-term price trends, its short-term movements are not reliable for predicting Bitcoin price fluctuations. Therefore, while gold can serve as a general indicator of Bitcoin's price over extended periods, it may not be a useful tool for short-term price forecasting.

3. Seasonality of Returns

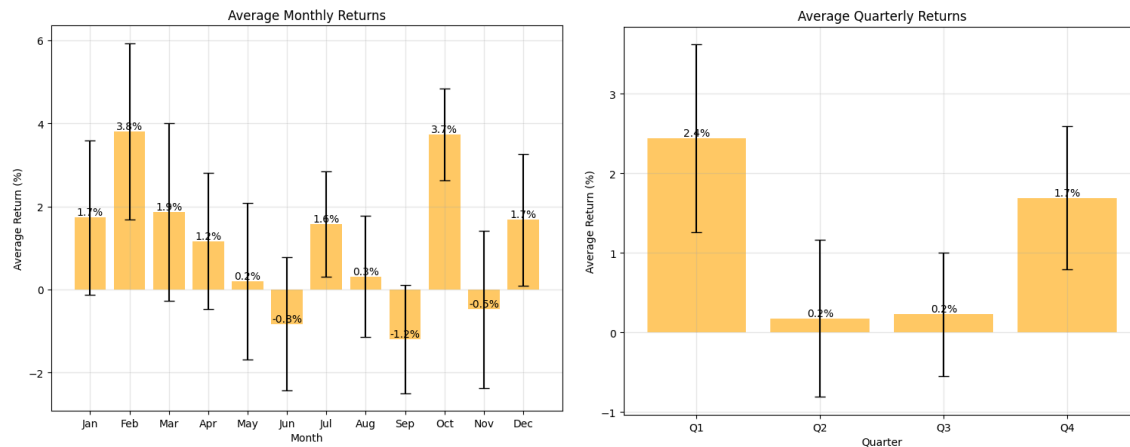


Figure 1.4: Average monthly and quarterly returns for Bitcoin

The analysis of Bitcoin's monthly and quarterly returns reveals notable seasonal patterns. February and October stand out with significantly higher returns, averaging 3.8% and 3.2%, respectively, suggesting these months tend to perform well. In contrast, June and September show weaker performance, with returns averaging -0.8% and -1.2%. This indicates that Bitcoin's price movements may follow certain seasonal trends, with some months exhibiting stronger returns and others experiencing more frequent declines. On a quarterly basis, Q1 shows the highest average return at 2.4%, while Q2 and Q3 are nearly flat, with a modest 0.2% return. Q4 performs on average at a 1.7% return. These findings suggest that the first and final quarters tend to be the most favorable for Bitcoin, while the second and third quarters tend to be the weakest and subdued. While these seasonal patterns could inform market strategies, the volatility of Bitcoin should be considered when applying them.

Question 1 Discussion

In this study, we set out to explore how macroeconomic factors, such as M2 money supply, the Fear & Greed Index, and alternative risk assets like gold, correlate with short-term Bitcoin price movements and whether these factors could be predictive of Bitcoin's price trends. Our analysis revealed moderate correlations between Bitcoin's price and some of these macro indicators, notably with funding rates, sentiment metrics, and M2 money supply. While these factors show strong correlations over longer time frames, especially in the case of M2 money supply and gold, their ability to predict Bitcoin's price in the short term is more limited.

For instance, the Fear & Greed Index showed a moderate correlation with Bitcoin returns but was only predictive of current market sentiment, with diminishing correlations as we moved to longer lags. This suggests that the index is more of a concurrent indicator, capturing market psychology rather than providing actionable signals for short-term price trends. Similarly, while

gold and Bitcoin exhibited a strong correlation in price levels, the short-term correlation, especially on a weekly basis, was weak, making gold a poor short-term predictor of Bitcoin's price. We also observed seasonal trends in Bitcoin's returns, with certain months and quarters displaying consistently higher returns, particularly in February and Q1. However, despite these trends, Bitcoin's inherent volatility means that these seasonal patterns are not necessarily reliable indicators for short-term price movements.

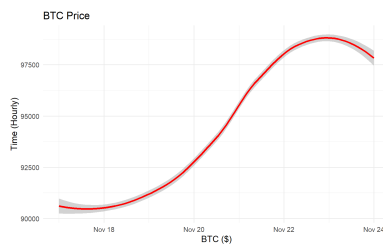
Ultimately, while macroeconomic factors like M2 money supply, gold, and sentiment metrics offer valuable insights into long-term trends, they are less effective at providing short-term price predictions for Bitcoin. The analysis indicates that while there are correlations to explore, pinpointing direct short-term indicators for Bitcoin price movements remains a challenging task, and more sophisticated models may be needed to capture the dynamic and volatile nature of Bitcoin's market behavior.

Question 2

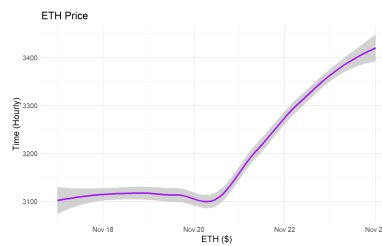
Question 2 Analysis

Cryptocurrency close data for a 7-day period was selected, as seen in Fig. 2.1, dating from November 17, 2024, to November 24, 2024. A 7-day period was selected because of the short-term focus of this problem. Data was looked at on the hourly level; this resulted in $7 \cdot 24$ data points. This specific week was chosen because Bitcoin hit an all-time high of ~\$99k. Bitcoin (BTC) data and data for four altcoins were selected: Ethereum (ETH), Solana (SOL), Sui (SUI), and Litecoin (LTC). These specific altcoins were chosen due to their popularity.

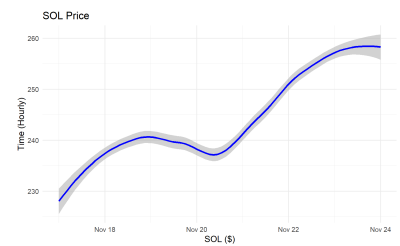
2.1.1



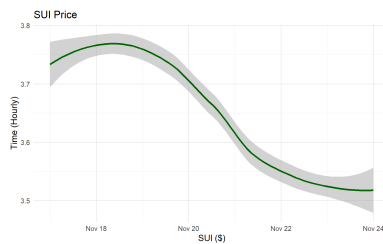
2.1.2



2.1.3



2.1.4



2.1.5

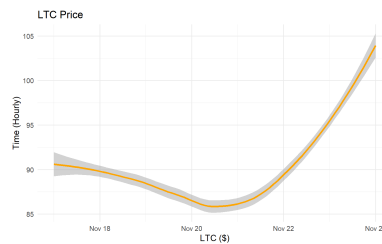
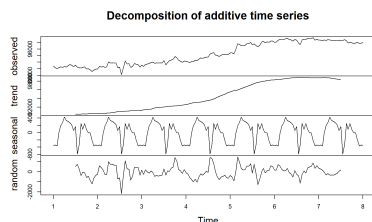


Figure 2.1 shows BTC and altcoin close prices (e.g., ETH, SOL, SUI, and LTC) for a 7-day period at an hourly level. Prices range from November 17, 2024, to November 24, 2024.

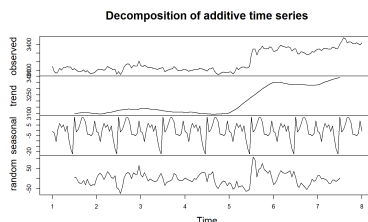
As seen in Fig. 2.2, the seasonal decomposition of each coin was taken to further understand the nature of the data. A SARIMAX model, selected by the autoarima function, was ultimately used due to the high daily seasonality and the desire to see what influence BTC has on altcoin prices. To further analyze the seasonal decomposition, there is a 24-hour seasonality in the market, which aligns with the times of day people are available to trade. Although cryptocurrency has a 24-hour market, there are downtimes when the majority of traders are not trading due to time-zone constraints. This was a determining factor in selecting a SARIMAX model instead of an ARIMAX model.

In addition to seasonal decomposition, we conducted time series analysis tests to understand the nature of the data at hand. An ADF test was taken to determine whether the raw time series data was stationary. As seen in Table 2.1, it was determined that all of the time series were non-stationary. Autocorrelation and partial autocorrelation plots, shown in Fig. 2.3, were used to analyze the data further to determine what parameters should be used. The data shows high autocorrelation but no partial autocorrelation across hourly lag variables. The high autocorrelation can be credited to the volatility of cryptocurrencies. The lack of partial autocorrelation can be credited to the complex predictability of the cryptocurrency market.

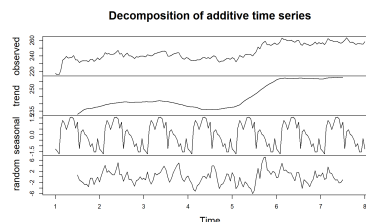
2.2.1



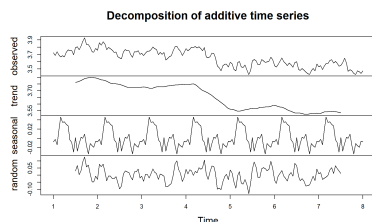
2.2.2



2.2.3



2.2.4



2.2.5

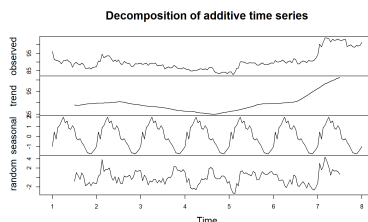
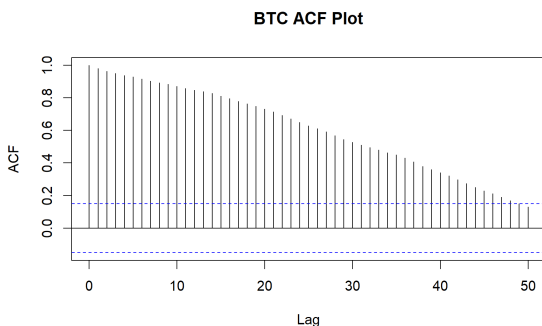
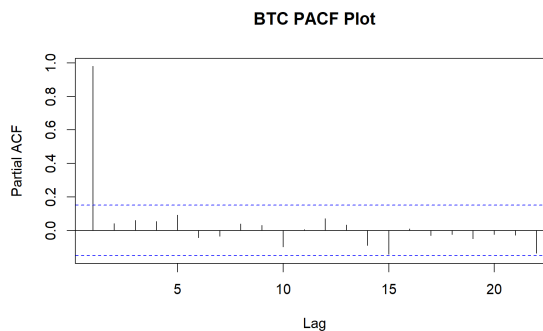


Figure 2.2 shows seasonal decomposition for cryptocurrency close prices (e.g., BTC, ETH, SOL, SUI, and LTC) for a 7-day period at an hourly level.

2.3.1

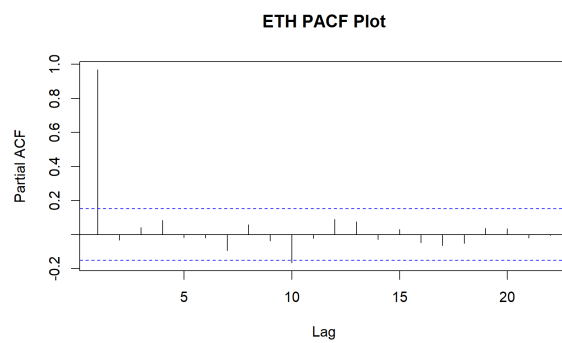
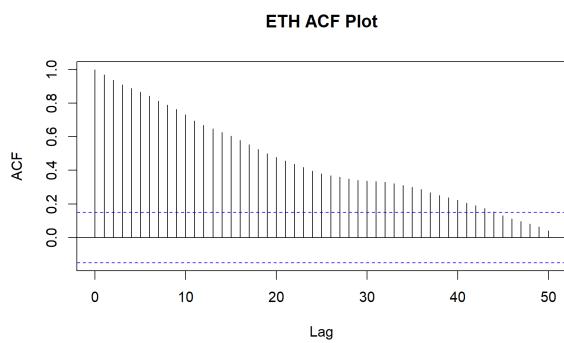


2.3.2

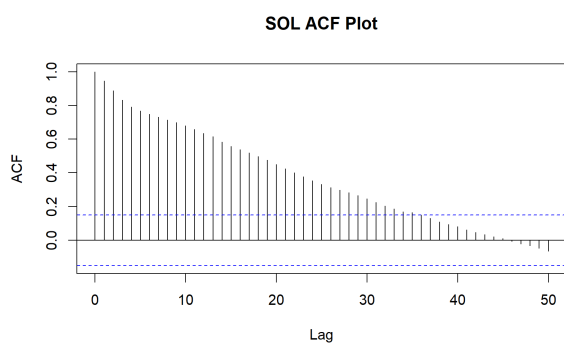


2.3.3

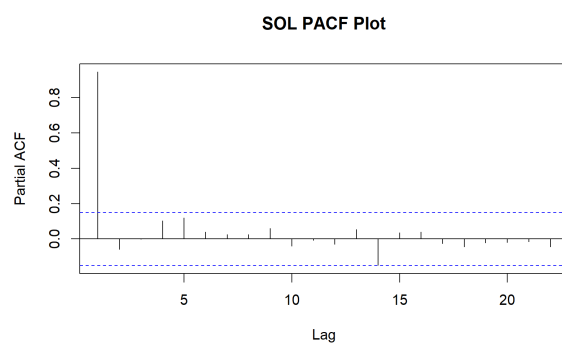
2.3.4



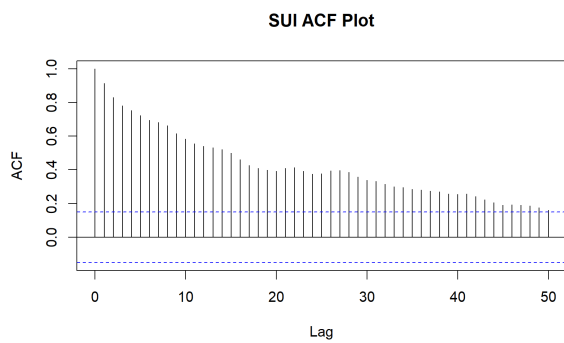
2.3.5



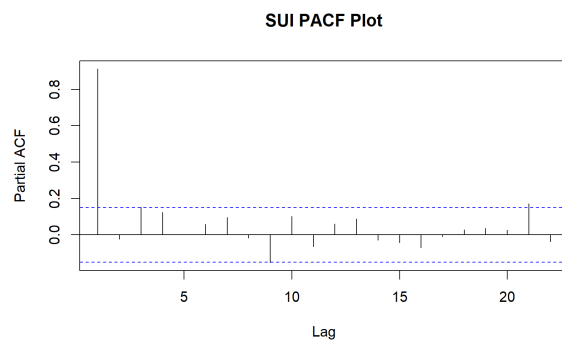
2.3.6



2.3.7



2.3.8



2.3.9

2.3.10

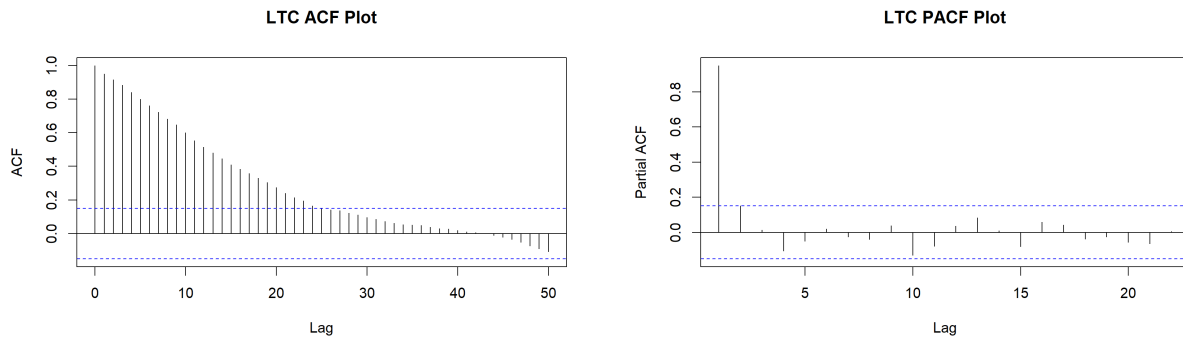


Figure 2.3 shows autocorrelation and partial autocorrelation plots for cryptocurrency close prices (BTC, ETH, SOL, SUI, and LTC).

Cryptocurrency	ADF Statistic	p-value	Stationarity
BTC	-1.684	0.708	Non-Stationary
ETH	-1.875	0.629	Non-Stationary
SOL	-2.161	0.510	Non-Stationary
SUI	-3.421	0.053	Non-Stationary
LTC	-1.653	0.721	Non-Stationary

Table 2.1 shows the results of the ADF tests done on cryptocurrency close prices (BTC, ETH, SOL, SUI, and LTC).

Question 2 Results

The autoarima with seasonality and exogenous variables (BTC close prices) was run for each altcoin, and parameters were algorithmically determined. ETH and SUI ran with an ARIMA(1, 0, 0) regression model, SOL ran with an ARIMA(2,0,2) regression model, and LTC ran with an ARIMA(0,1,0) regression model. A seasonal component was not determined by the autoarima function. This can be explained by the small size of the dataset. Although seasonality was specified as True, the function could not detect seasonality as significant. These results differ from the results of the EDA tests that were done and are likely due to the small size of the dataset. Figure 2.4 shows forecasting data for the selected altcoins, with SOL's forecasted close prices being the closest to the actual close prices.

The results of the SARIMAX, as seen in Fig. 2.4, for the selected altcoins, show varying levels of accuracy and error across both the training and testing datasets. ETH had a training MAE of 13.917 with an MPAE of 0.439% and a testing MAE of 78.323 with an MPAE of 2.307%, indicating a drop in forecasting performance. This can be attributed to the unpredictability of the market and the fact that there are no actual tools to accurately predict the market.

For SOL, the training set had great accuracy with an MAE of 1.566 and a MAPE of 0.655%. The model maintained its effectiveness for the testing set with an MAE of 4.541 and MAPE of 1.757%. This indicates that, although there is a slight reduction in performance, the results are still reliable.

SUI had an excellent training accuracy with an MAE of 0.033 and an MAPE of 0.893%. The testing set also has great capabilities, although there is higher error, with an MAE of 0.153 and 4.366%. This reflects a high accuracy for both the training and testing set.

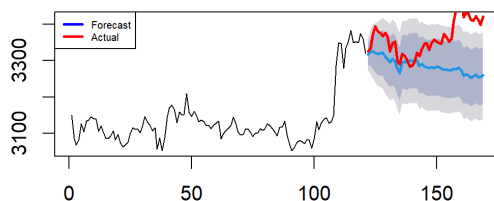
LTC, shows an MAE of 0.655 and an MAPE of 0.736%, suggesting that the model was able to fit well to the training data. The MAE of the testing set was 6.171 and a MAPE of 6.123%. This shows that there is higher error, suggesting that the model struggles to see the unseen data.

Correlations were calculated between BTC and altcoins, with ETH having a correlation of 0.495, SOL having a correlation of 0.508, SUI having a correlation of 0.481, and LTC having a correlation of 0.416.

Overall, while the SARIMAX models for SUI and SOL demonstrated strong performance in both accuracy and consistency, the models for ETH and LTC showed greater variability, particularly in the test results for ETH, which suggests room for improvement in capturing the underlying patterns for forecasting. The correlation values suggest there is the strongest correlation between BTC and SOL, which is reflected in the results of the SARIMAX model. LTC exhibits the weakest correlation, which suggests why the model is unable to predict LTC very well.

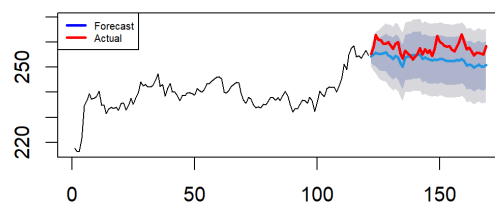
2.4.1

ETH Close Forecast (SARIMAX with BTC)



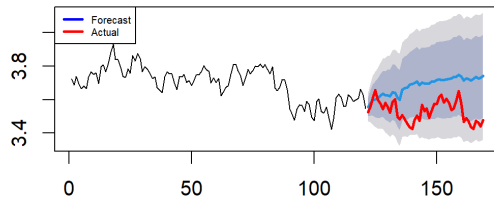
2.4.2

SOL Close Forecast (SARIMAX with BTC)



2.4.3

SUI Close Forecast (SARIMAX with BTC)



2.4.4

LTC Close Forecast (SARIMAX with BTC)

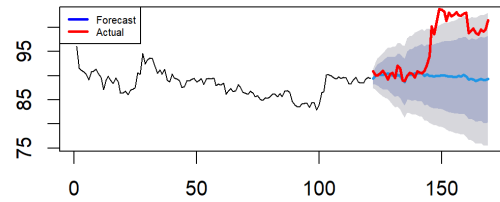


Figure 2.4 shows forecasting results for SARIMAX models of altcoin cryptocurrencies with BTC as an exogenous variable.

Question 2 Discussion

The results of the SARIMAX model suggest that a simple model can predict some altcoin prices in the short term with variability. This shows that, with improvements to the model, such as lagged BTC variables, more training data, or picking lag variables based on what coin the altcoin is built on. These minor improvements could help make more accurate predictions and even allow the autoarima to detect a seasonal component. There are significant implications to answering this research question as this could be a method for making smart trading decisions. Furthermore, the greatest limitation of this research question was that the data was on an hourly basis. If the data had been more granular, microtrends in the data could have been detected and potentially resulted in a better SARIMAX model.

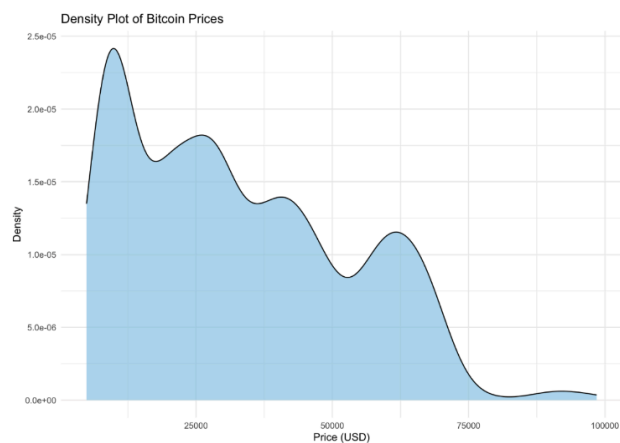
Question 3

Question 3 Analysis

This analysis examines whether integrating historical Bitcoin prices improves our ability to forecast current altcoin prices. To explore this relationship, we used time series data for Bitcoin and selected altcoins to dissect patterns in trends, seasonality, and volatility. By comparing two predictive models—one using solely current Bitcoin prices (Model A) and another incorporating lagged Bitcoin prices along with control variables such as altcoin trading volume and volatility (Model B)—we assessed whether including historical Bitcoin performance improves predictive precision. We seek to quantify this potential improvement to give us better insights into market dynamics and discuss practical implications for traders, investors, and portfolio managers within the cryptocurrency sector.

We used several figures to visualize the patterns, relationships, and potential anomalies in our data. Density plots (Figures 1 and 2) of Bitcoin and altcoin prices had right-skewed distributions, a characteristic feature of financial assets. Bitcoin prices displayed clustering around notable price levels, while altcoins exhibited similar trends with varying degrees of skewness. The decomposition of the time series data (Figures 3 and 4) highlighted clear trend components and seasonal patterns. Bitcoin showed a gradual upward trend with cyclical fluctuations, whereas altcoins showed more pronounced volatility and shorter cycles. Figure 5 shows price trends for bitcoin and altcoins on one graph. Through residual analysis of the decomposed time series, we pinpointed potential areas for improving model accuracy by finding irregular components.

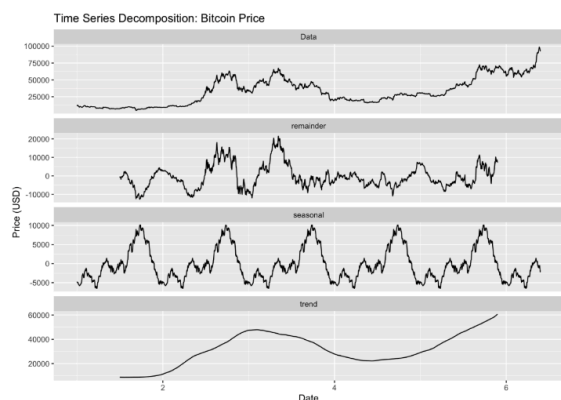
Question 3, Figure 1



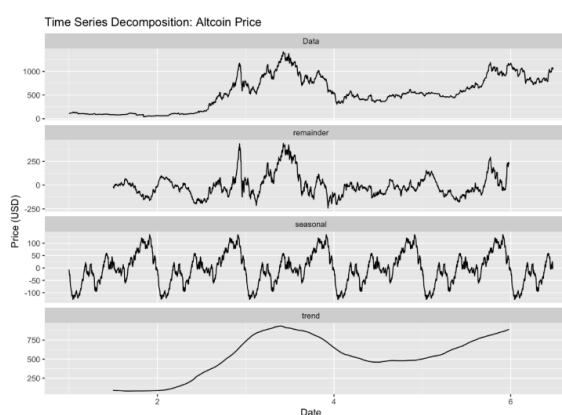
Question 3, Figure 2



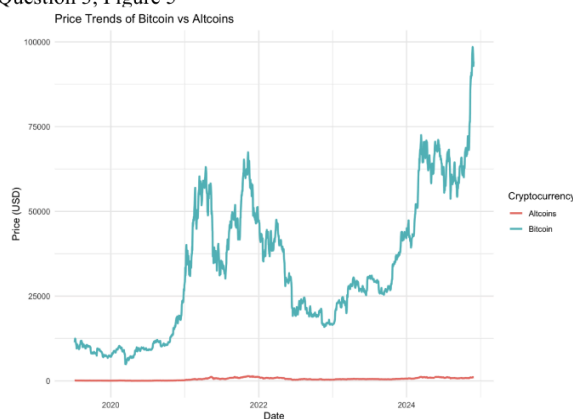
Question 3, Figure 3



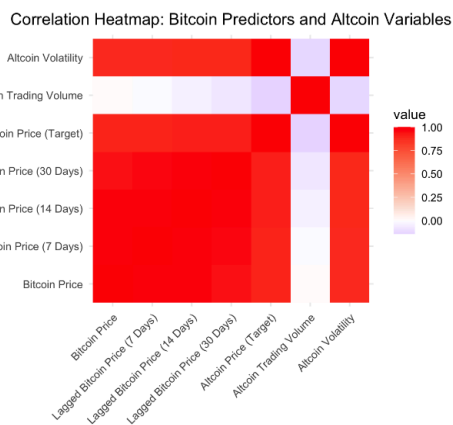
Question 3, Figure 4



Question 3, Figure 5



Question 3, Figure 6



We used both Frequentist and Bayesian approaches to ensure that our results were both robust and interpretable. We conducted an ordinary least squares regression for an initial model fitting, and we used Bayesian techniques to give us insights into parameter uncertainty and generated predictive distributions. Time-series cross-validation corroborated the models on unseen data to ensure generalizability. We evaluated our models with the following performance metrics: root mean square error (RMSE), mean absolute error (MAE), adjusted R^2 , and Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC).

A correlation heatmap (Figure 6) revealed significant relationships between Bitcoin's current and lagged prices (e.g., 7-day, 14-day, and 30-day lags) and altcoin prices. Strong positive correlations suggested that past Bitcoin performance impacted subsequent altcoin movements. Trading volume and volatility metrics for altcoins were moderately correlated to their respective prices, which emphasizes their importance as control variables in our predictive models.

To answer our question, we created and compared two regression models. Model A served as the baseline and used only the current Bitcoin price to predict the target altcoin's daily price. Model B extended this by including lagged Bitcoin prices (7-day, 14-day, and 30-day lags) and control variables (altcoin trading volume and volatility). These additional predictors were used to

capture delayed effects and broader market dynamics. We then compared these models using time-series cross-validation, Bayesian posterior predictive checks, and metrics such as AIC and BIC.

Below are the formulas used for each model:

Model A (baseline):

$$Y_{altcoin_t} = \beta_0 + \beta_1 X_{bitcoin_t} + \epsilon$$

Model B (lagged model):

$$Y_{altcoin_t} = \beta_0 + \beta_1 X_{bitcoin_t} + \beta_2 X_{bitcoin_{t-1}} + \beta_3 X_{bitcoin_{t-2}} + \dots + \beta_k Z_k + \epsilon$$

Where Z_k represents control variables.

The additional predictive power of lagged Bitcoin prices was assessed using an ANOVA test, allowing us to quantify how much including lagged variables improved the model fit. Furthermore, we used a Diebold-Mariano test to provide a formal statistical comparison of the forecasting errors between the two models to validate predictive accuracy independently.

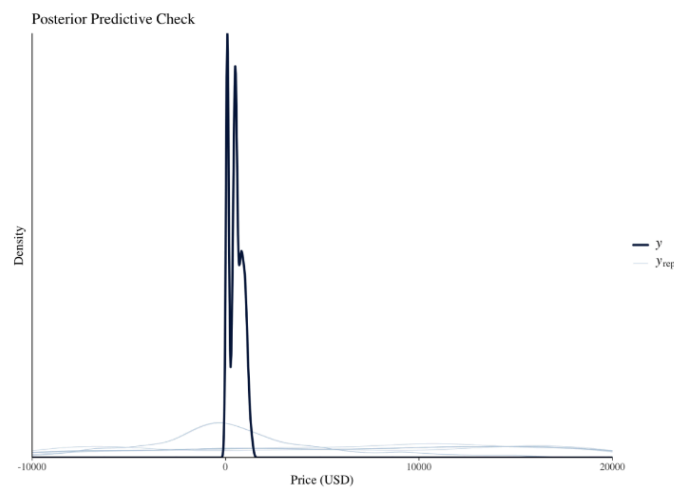
Question 3 Results

Model B demonstrated superior predictive accuracy compared to Model A. The enhanced model achieved an 84% reduction in Root Mean Square Error (RMSE) compared to Model A, alongside a marked decrease in Mean Absolute Error (MAE). The adjusted R² values indicated that Model B explains 99.5% of the variance in altcoin prices, a clear improvement over Model A's 80.5%. It is clear that Model B had greater explanatory power because it used the additional predictors. Additionally, Model B had lower AIC and BIC values, indicating a better balance between model fit and complexity.

Model	RMSE	MAE	Adjusted R ²	df	AIC	BIC
Model A (Current Bitcoin Price)	154.86113	120.38075	0.80481	3	25477.10	25493.86
Model B (Lagged Bitcoin Prices)	25.37588	21.05761	0.99475	8	18357.09	18401.78

Our posterior predictive check in Figure 7 showed a strong alignment between the observed density (y) and the central predictive densities (y_rep) of Model B, indicating that it adeptly captured the overall patterns in the data. While deviations in the tails indicated that the model encountered challenges with extreme price fluctuations (likely attributed to the volatile nature of the cryptocurrency market), it nevertheless offered a robust fit for the central tendencies.

Question 3, Figure 7



Our ANOVA results confirmed that including lagged Bitcoin prices and control variables leads to a significant reduction in residual variance, once again providing evidence in favor of Model B's predictive abilities.

The hypotheses for this ANOVA test were as follows:

- Null Hypothesis (H_0): Including lagged Bitcoin prices does not improve predictive accuracy for altcoin prices.
- Alternative Hypothesis (H_A): Including lagged Bitcoin prices improves predictive accuracy for altcoin prices.

The following formula was used to find the F-statistic:

$$F = \frac{\text{explained variance}}{\text{residual variance}}$$

The high F-statistic (14236) and low p-value ($< 2.2e-16$) from the ANOVA test provided strong evidence against the null hypothesis, showing that lagged Bitcoin prices significantly improved predictive accuracy. A Diebold-Mariano was used with the following hypotheses:

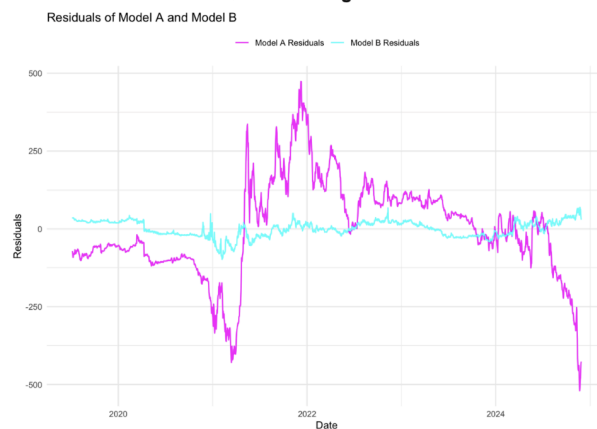
- Null Hypothesis (H_0): Model B's predictive accuracy is not better than Model A's.
- Alternative Hypothesis (H_A): Model B's predictive accuracy is better than Model A's.

The Diebold-Mariano statistic of -4.02 (with a p-value of 0.00003) confirmed Model B's superiority in forecasting accuracy.

Residual analysis (Figure 8) showed that Model B (cyan) produced lower-magnitude and more stable residuals compared to Model A (magenta), indicating a better fit to the observed altcoin prices, especially during periods of market volatility (such as in 2022 and 2024). On the other hand, Model A had large residual spikes during these periods, an indication of its inability to capture the complex dynamics between Bitcoin and altcoin prices.

Figure 9 further supports the improved performance of Model B by comparing the predictions for Models A and B to the actual altcoin prices (black). Model B closely tracked actual price trends, particularly during volatile periods. Model A would often underestimate or overshoot price movements because of its limited ability to account for temporal dependencies and market influences.

Question 3, Figure 8



Question 3, Figure 9



Overall, Model B outperformed Model A across all metrics, including RMSE, MAE, adjusted R^2 , AIC, BIC, ANOVA, and Diebold-Mariano tests. Model B, which included lagged Bitcoin prices and control variables, achieved an 84% reduction in RMSE and explained 99.5% of the variance in altcoin prices. We learned that the 7-day lag was the most predictive and that including lagged Bitcoin prices significantly improved predictive accuracy. While minor deviations occurred during extreme market conditions, Model B demonstrated greater robustness and reliability than Model A.

Question 3 Discussion

Our analysis provides several practical takeaways for stakeholders. Individuals could develop more sophisticated trading algorithms that improve upon the timing of buying and selling decisions based on the findings from Model B. Traders and investors could make more informed risk assessments, and portfolio managers might adjust asset allocations to mitigate the risk of losses during market volatility.

The evidence shows that the market often does not immediately implement the available information. Traders could leverage these inefficiencies for short-term gains. Model B could help investors better understand how Bitcoin's historical performance affects other altcoin markets and balance their portfolios accordingly.

However, our analysis did have some limitations. For instance, we only analyzed four major altcoin currencies and, as a result, likely ignored the dynamics of smaller cryptocurrencies. We were also limited to historical data, which means we were likely not able to fully capture future market behaviors, especially given the cryptocurrency market's rapidly growing, unpredictable nature. Both models make linear assumptions, such as linear relationships, stationarity of time series data, and normally distributed residuals, which could be problematic in high-stakes scenarios involving the volatile and non-linear nature of cryptocurrency markets.

The models also did not take into account regulatory changes, technological advancements, market sentiment, macroeconomic indicators, or significant geopolitical events, all of which can significantly influence cryptocurrency prices.

While we used regularization techniques like Ridge Regression, there is still a risk of models overfitting, especially given Model B's high Adjusted R-squared value. This challenge points to a need for ongoing model validation and exploration of other modeling approaches.

To address the identified limitations and further enhance predictive capabilities, in the future, we could try some more strategies such as:

1. Incorporating a greater range of altcoins to capture more diverse market behaviors and correlations.
2. Employing non-linear modeling approaches like neural networks or ensemble methods to better capture the data's complex relationships.
3. Integrate external market indicators such as regulatory news, technological developments, macroeconomic variables, and social media sentiment to improve the predictive models.
4. Use more advanced regularization and feature selection techniques to mitigate overfitting and improve the generalizability of our model.
5. Explore real-time data feeds to dynamically adapt models to account for rapidly changing market conditions.

In conclusion, Model B outperforms Model A in all areas: RMSE, MAE, Adjusted R-squared, AIC, BIC, ANOVA, and the Diebold-Mariano test.

To better understand altcoin market dynamics, it is important to consider both current and past Bitcoin prices, as well as altcoin trading volume and volatility. Because Bitcoin's historical

performance significantly impacts the altcoin market, Model B can be useful to stakeholders looking to improve trading strategies, risk management practices, and investment decision-making within the field of cryptocurrency. Despite our limitations, this analysis lends itself to future research that improves upon existing predictive models and explores the multifaceted relationships governing cryptocurrency markets.

Question 4

Question 4 Analysis

Fundamental Analysis

For this analysis, I utilized data obtained from the Cryptocompare.com API, narrowing the dataset to the most recent 1500 days. This selection provides a four-year window, aligning with the creation of Solana (SOL) in March 2020. The initial intent was to examine the relationship between Bitcoin volatility and altcoin price changes. However, due to the inherent nature of volatility ranging from 0 to infinity and price change encompassing all real numbers, the findings were nothing of interest. Consequently, the focus shifted to a comparative analysis of volatilities between BTC and selected altcoins.

To calculate volatility, I employed a standard economic convention of a 30-day window for computing the standard deviation. This method ensured consistency in assessing short-term volatility trends across the dataset, which begins in December 2020 due to the 30-day calculation window. Notably, there were no missing data points apart from the initial 30 days used to compute volatility. The altcoins chosen for comparison were Ethereum (ETH), Solana (SOL), Binance Coin (BNB), and Ripple (XRP), given their popularity and significant market presence. All volatility values in the dataset were greater than zero (obviously), with no values exceeding 0.2, reflecting the constrained nature of volatility during the observed period. This upper bound (0.2) suggests relatively low volatility compared to typical cryptocurrency market behavior, which can often experience much higher fluctuations in 30-day periods.

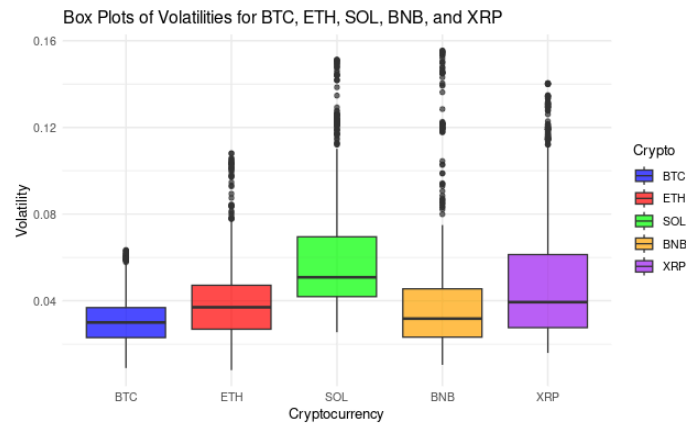
To further analyze the data, I used scatter plots to visualize the relationships between BTC volatility and the volatilities of each altcoin, highlighting potential correlations. Boxplots were employed to compare the distributions of volatility across assets, providing insight into variability and relative stability. A temporal analysis of volatility trends over time allowed for the identification of significant patterns or anomalies in market behavior. Additionally, density functions were utilized to examine the overall distribution of volatility for each cryptocurrency, aiding in understanding their risk profiles. Finally, correlation coefficients were calculated to quantify the strength and direction of relationships between BTC volatility and that of the altcoins, offering a clearer picture of interdependencies within the market.

Scatter Plots



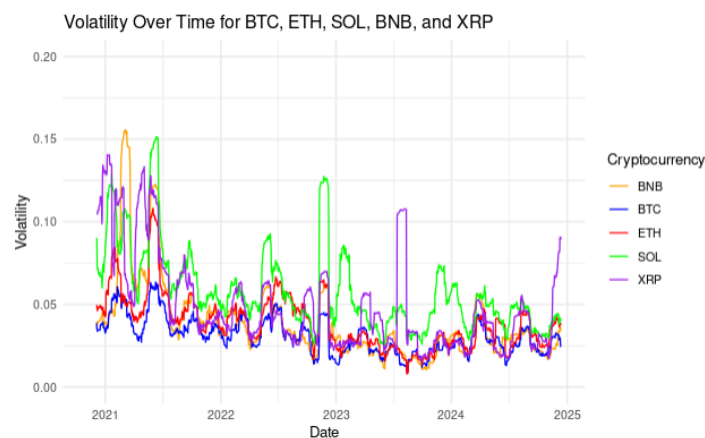
The scatter plots illustrate a positive relationship between Bitcoin's volatility and the volatility of the selected altcoins. Across all plots, as BTC volatility increases, altcoin volatility increases as well. However, the strength and spread of this relationship vary slightly between altcoins. ETH, SOL, and BNB exhibit a tighter clustering of points, indicating a potentially stronger correlation with BTC volatility. In contrast, XRP shows a slightly broader spread, suggesting relatively more variability in its response to changes in BTC volatility. Overall, the scatter plots reveal a clear, albeit differing, pattern of alignment between BTC and altcoin volatilities, reinforcing the idea of Bitcoin's dominant influence on broader cryptocurrency market dynamics.

Box Plots



The boxplots reveal distinct patterns in the 30-day volatilities of cryptos over the last 4 years. Bitcoin exhibits the lowest median volatility (~0.02), with a relatively narrow interquartile range, indicating greater stability compared to the altcoins. Ethereum follows with a slightly higher median volatility (0.035) and a somewhat broader IQR, suggesting moderate variability. Solana stands out with the highest median volatility (~0.05) and a significantly wider IQR compared to BTC and ETH, indicating it experiences greater fluctuations. Binance Coin and Ripple both exhibit median volatilities near 0.04 and 0.03, but XRP has a slightly wider IQR than BNB, suggesting more variability in its volatility. Additionally, SOL, BNB, and XRP show numerous outliers extending beyond 0.12, highlighting periods of extreme volatility for these altcoins. Overall, the boxplots demonstrate that BTC is the most stable cryptocurrency in terms of volatility, while SOL has higher average volatility and XRP has more variation.

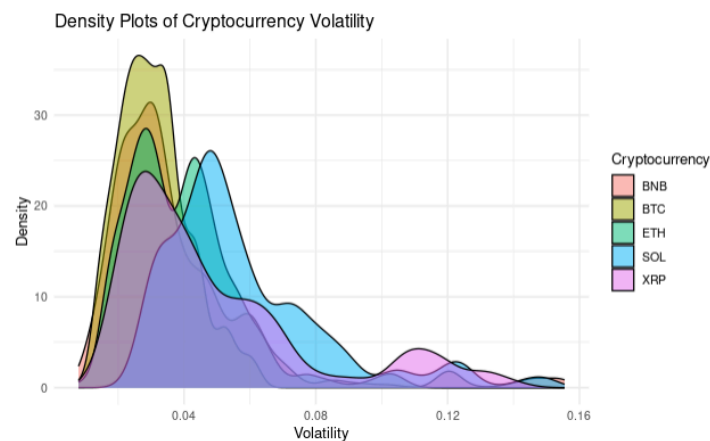
Scatter Plots



The volatility over time plot reveals the dynamic behavior of BTC, ETH, SOL, BNB, and XRP from December 2020 to today. Bitcoin exhibits consistent stability, with volatility rarely exceeding 0.05 and remaining tightly clustered. Ethereum also shows relative stability, with

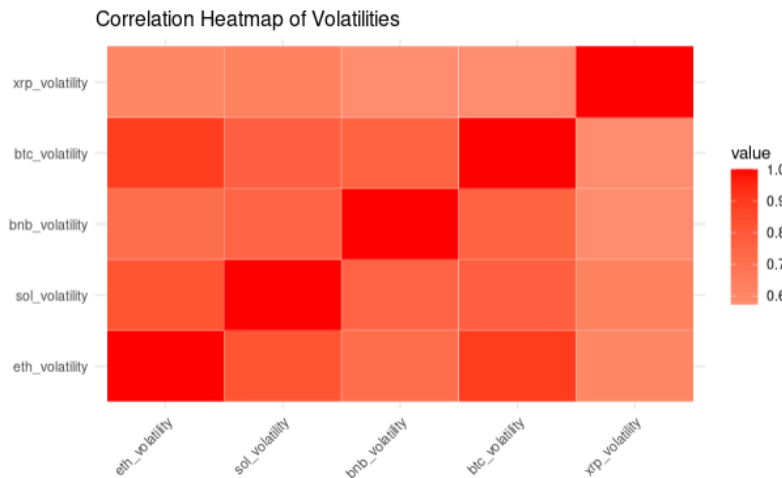
volatility generally fluctuating between 0.03 and 0.08, as well as occasional spikes that reach approximately 0.10. In contrast, Solana, Binance Coin, and Ripple display pronounced spikes in volatility, often exceeding 0.10, with all three reaching ~0.15. These extreme spikes highlight periods of significant market turbulence for these altcoins. The timing of these spikes also varies. For SOL, large volatility events are most prominent in early 2021 and late 2022. XRP shows similar periods of heightened activity but exhibits more frequent smaller spikes throughout the time frame. BNB demonstrates spikes that coincide with market-wide turbulence but less frequently than SOL or XRP. These dramatic volatility fluctuations are largely absent in BTC and less common in ETH, emphasizing the comparative stability of these assets. The plot suggests that while altcoins like SOL, BNB, and XRP are prone to sudden and extreme market movements, BTC and ETH provide more consistent risk profiles over the analyzed period. This distinction underscores Bitcoin's role as the most stable cryptocurrency for the broader market.

Density Plots



The density plots provide a visual representation of the distribution of volatilities for BTC, ETH, SOL, BNB, and XRP over the last four years. Bitcoin exhibits a highly concentrated density of around 0.04, reflecting its stability and minimal variability compared to the altcoins. Ethereum, Binance Coin, and Solana have distributions that are more similar to BTC, with their peaks occurring slightly above 0.04 and tapering off gradually. These similarities suggest that these assets experience comparable levels of volatility, though they exhibit slightly more variation than BTC. Ripple, on the other hand, displays a significantly wider variation in its volatility distribution (with a peak at approximately 0.03), but its variation looks a lot different from that of the others. It also has a second bump around 0.12, indicating a substantial number of periods with extreme volatility that are absent in BTC. This wider distribution and higher volatility highlight XRP's relatively higher risk profile compared to BTC and even the other altcoins. The density plots thus emphasize that while ETH, BNB, and SOL exhibit volatility profiles closer to BTC, XRP stands out as an outlier with significantly higher variability and a notable concentration of extreme volatility.

Correlation



The correlation matrix provides numerical insights into the relationships between the volatilities of BTC and the selected altcoins (ETH, SOL, BNB, and XRP) over the last four years. The strongest correlation is observed between BTC and ETH volatility, with a coefficient of 0.9047, indicating a very close and almost linear relationship. This aligns with the density and scatter plots, where ETH volatility mirrored BTC's stability and showed minimal deviation, suggesting that ETH's market movements are closely tied to BTC. Solana (SOL) has a moderately strong correlation with BTC, with a coefficient of 0.7704. This supports the scatter plot observation that SOL generally follows BTC's volatility trends but experiences occasional significant spikes, reflecting periods of decoupling during heightened market activity. Binance Coin (BNB) exhibits a correlation of 0.7618 with BTC, slightly lower than SOL's but still substantial. Like SOL, BNB volatility follows BTC's general trend with some divergence during periods of extreme activity, as seen in both the scatter and density plots. Ripple (XRP), however, shows the weakest correlation with BTC, at 0.5752, indicating a more independent volatility pattern. This aligns with XRP's broader and more varied density distribution, as well as the scatter plots where XRP demonstrated more significant and frequent spikes compared to BTC. These findings reinforce the visual patterns observed earlier. Cryptocurrencies like ETH, SOL, and BNB have volatility behaviors that are closely linked to BTC, with ETH being the most correlated. XRP, on the other hand, displays much more independent behavior. This divergence underscores XRP's higher risk and less predictable relationship with the overall market.

Question 4 Results

P-Scores and Standard Error

The results of the GLMs and GAMs for each altcoin (ETH, SOL, BNB, and XRP) that were used to assess the effect of BTC volatility as a confounding variable on altcoin volatility are displayed in the table below.

Model Type and Metric	ETH	SOL	BNB	XRP
GLM P-Score	<2e-16	<2e-16	<2e-16	<2e-16
GAM P-Score	<2e-16	<2e-16	<2e-16	<2e-16
GLM Significance	***	***	***	***
GAM Significance	***	***	***	***
GLM Standard Error	0.0170478	0.03864	0.038735	0.0583498
GAM Standard Error	0.000437	0.001014	0.001046	0.001505

The results of the Generalized Linear Models and Generalized Additive Models for each altcoin (ETH, SOL, BNB, and XRP) demonstrate a statistically significant relationship between Bitcoin volatility and the volatilities of these altcoins. The P-scores for both models are extremely low (<2e-16), which is far below the conventional significance threshold of 0.05. This strong significance is further supported by the "***" notation in the table, indicating high confidence in the effect of BTC volatility on these altcoins. The extremely small P-scores suggest that BTC's volatility serves as a meaningful explanatory variable for the volatility of the altcoins included in the analysis.

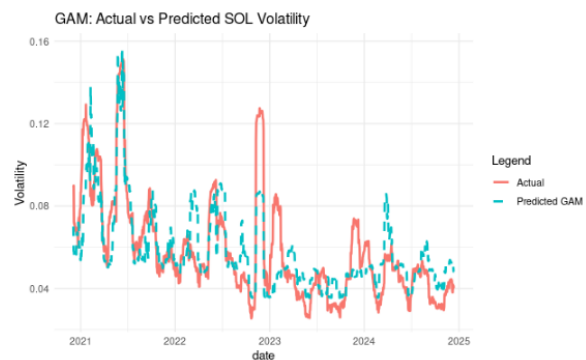
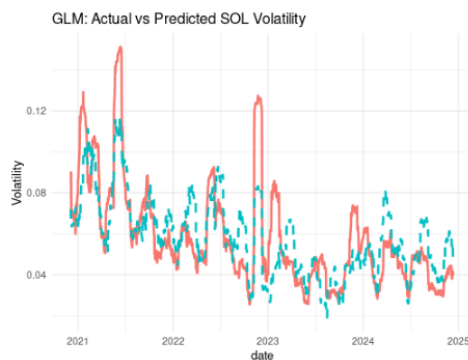
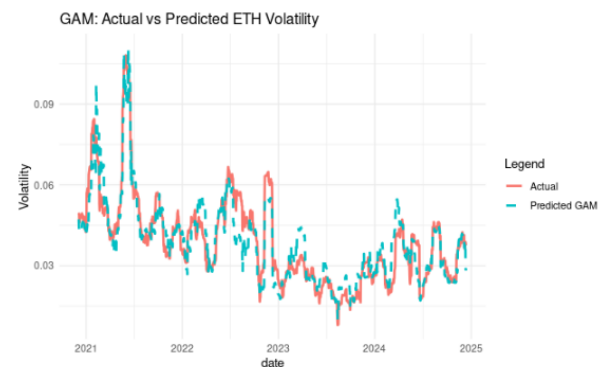
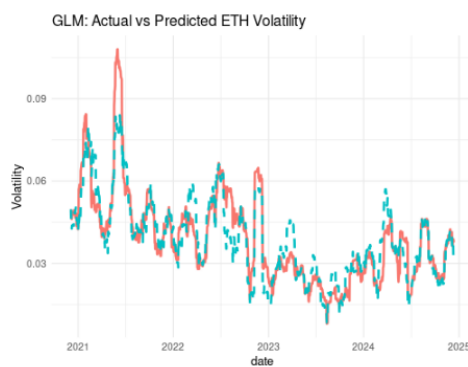
The standard errors for the GLM and GAM models vary slightly across the altcoins. For example, ETH has a GLM standard error of 0.0170478 and a GAM standard error of 0.000437, both of which are relatively low. This indicates a strong and consistent relationship between BTC and ETH volatilities. The other altcoins, such as SOL and BNB, also show low standard errors in their GAM models (0.001014 and 0.001046, respectively), which suggests that the GAM models fit the data more precisely than the GLM models for these cryptocurrencies. However, XRP has the highest GLM standard error (0.0583498) and GAM standard error (0.001505) among the altcoins, reflecting greater variability in its relationship with BTC volatility, consistent with the lower correlation observed in the correlation matrix.

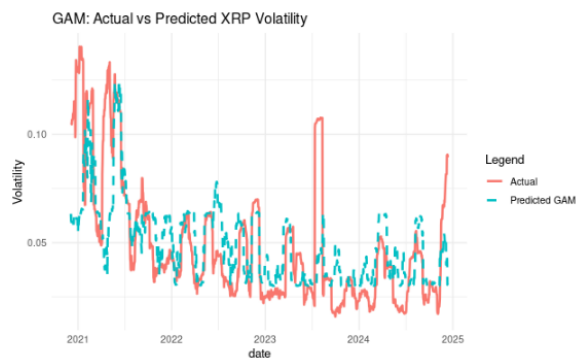
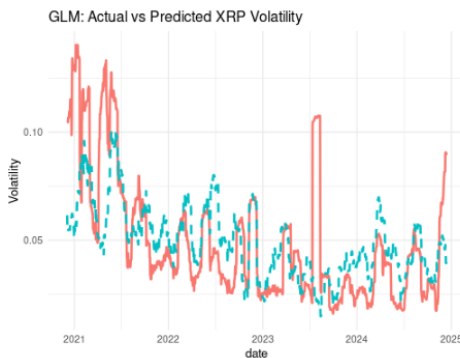
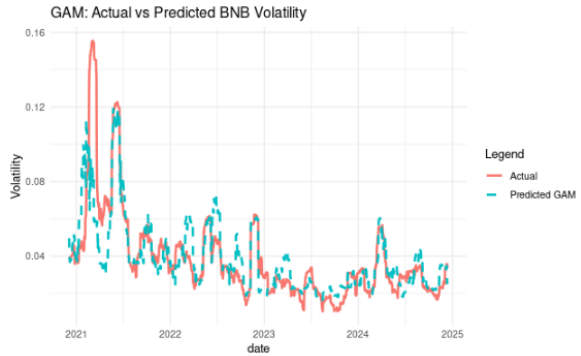
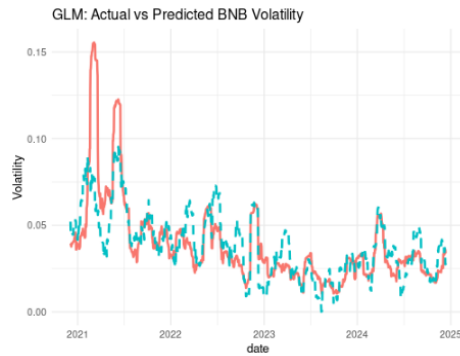
The potential overshooting of the P-scores could be due to several reasons. First, the size and structure of the dataset might contribute to this effect. With a large sample size (1500 days), even small effects can become statistically significant, potentially inflating the P-scores. Additionally, the relationship between BTC and altcoin volatilities may exhibit nonlinearity. While GAMs are designed to handle such relationships, the residual confounding or overfitting in the model could

amplify the statistical significance. Another possibility is multicollinearity or autocorrelation in the data, which could make the models overly sensitive to small variations, further exaggerating the P-scores.

Overall, the results strongly suggest that BTC volatility significantly impacts the volatilities of ETH, SOL, BNB, and XRP. However, the degree of this effect varies, with ETH showing the strongest and most stable relationship and XRP exhibiting the weakest and most variable connection. While the P-scores underline the importance of BTC as a confounder, their extremely low values warrant careful interpretation, as they might be influenced by the dataset's characteristics or model sensitivities.

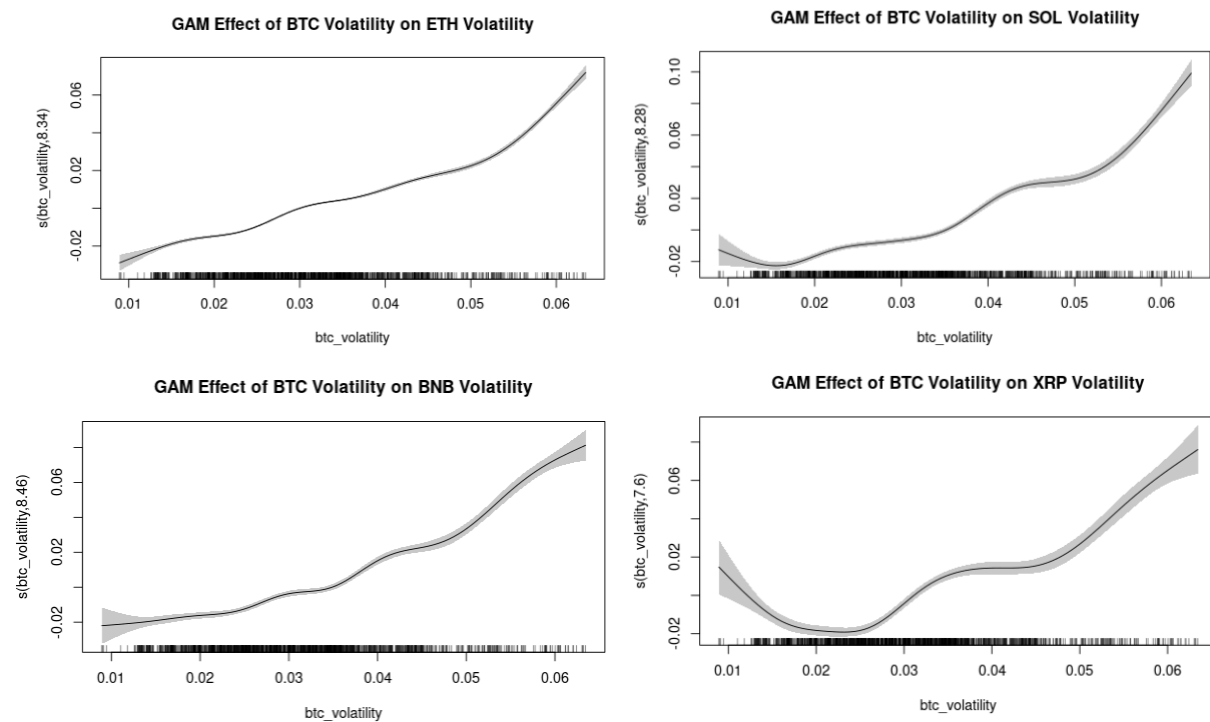
Actual vs. Predicted Probabilities





To evaluate how well BTC volatility explains altcoin volatility, I compared actual vs. predicted volatility using the GLMs and GAMs created in the previous section. These results show how well each model predicts the altcoin volatility patterns and highlight their strengths and limitations. The actual vs. predicted graphs reveal that GAMs consistently outperform GLMs when predicting periods of higher volatility. The GAM predictions align more closely with the actual values during these spikes, indicating that GAMs are better suited to capture the nonlinear relationships inherent in the data. However, both models struggle with extremely high volatility values, where predictions tend to deviate significantly from the actual values. This limitation is likely due to the rarity and unpredictability of extreme volatility events, which reduce the models' ability to generalize well in such scenarios. Additionally, during lower volatility periods, both GLM and GAM predictions closely match the actual values, demonstrating strong predictive power for more stable market conditions. Overall, while GAMs provide a more nuanced and accurate depiction of higher volatility trends, both models face challenges in capturing extreme fluctuations effectively.

Gam Effect of BTC Volatility Graph



These graphs were used to model and visualize the nonlinear effects of Bitcoin volatility on the volatilities of the altcoins using the same GAMs as before. The purpose of these visualizations is to better understand how BTC volatility influences these altcoins and whether the effects change depending on the level of BTC volatility. The smoothed curves and confidence intervals allow for a nuanced interpretation of the relationship, capturing nonlinear patterns and quantifying uncertainty in the estimates.

The graphs reveal that BTC volatility has a significant and dynamic effect on altcoin volatilities, with the impact increasing as BTC volatility rises. For XRP and SOL, the effect begins near zero or slightly negative at lower BTC volatility levels (below 0.02) and dips before rising sharply after 0.03, showing their heightened sensitivity to BTC volatility spikes. This makes sense due to bitcoin volatility being unchanging during the biggest spikes for XRP and SOL. BNB exhibits a smoother and more linear increase, suggesting a consistent response across the range of BTC volatility. ETH, while also linear, shows the smallest deviations and the most stable relationship, reflecting its closer alignment with BTC's volatility behavior. The confidence intervals indicate high reliability for most of the range, though some uncertainty exists at extreme BTC volatility levels due to fewer data points.

These findings demonstrate that BTC volatility is a key driver of altcoin volatility but affects each altcoin differently. XRP and SOL exhibit nonlinear and more pronounced reactions, particularly during periods of high BTC volatility, reflecting their susceptibility to market

shocks. In contrast, BNB and ETH have steadier and more predictable responses, with ETH showing the strongest alignment to BTC. This indicates that while all altcoins are influenced by BTC volatility, the degree and nature of this influence depend on the specific altcoin, highlighting BTC's role as a central force in the cryptocurrency market's volatility dynamics.

Question 4 Discussion

The analysis demonstrates that Bitcoin volatility significantly impacts the volatilities of major altcoins, as evidenced by the results from both Generalized Linear Models and Generalized Additive Models. The extremely low P-scores across all models indicate a robust statistical relationship. However, these results should be interpreted cautiously due to potential underlying biases inflating the P-scores. Factors such as the dataset size, multicollinearity, autocorrelation, or residual confounding effects may have contributed to the apparent overstatement of significance. Addressing these biases is a crucial step for improving the accuracy and reliability of such analyses in future research.

One of the key findings is that BTC is a strong predictor of low volatility changes across the selected altcoins. For periods where volatility remains below 0.06, BTC's volatility closely mirrors that of Ethereum, Solana, Binance Coin, and Ripple. This suggests that during stable or mildly fluctuating market conditions, BTC serves as a reliable anchor for the broader cryptocurrency market, with its volatility trends aligning well with those of other major altcoins.

However, the analysis also highlights a limitation: BTC volatility is less effective in predicting high volatility spikes in altcoins. Neither the GLM nor the GAM models performed well in capturing these extreme fluctuations, though GAMs offered slightly better predictions. These high-volatility periods often correspond to significant market events or shocks that disproportionately affect individual altcoins, such as protocol updates, regulatory news, or unique market behaviors. The inability to accurately model these spikes underscores the complexity and idiosyncratic nature of altcoin volatility, which is only partially explained by BTC's influence.

If I were to revisit this project, addressing the reasons behind the hyper-low P-scores would be a priority. Identifying and mitigating potential biases, such as multicollinearity or overfitting, would improve the robustness of the statistical models. Additionally, exploring alternative modeling approaches or incorporating external factors, such as macroeconomic indicators or sentiment analysis, could provide a more nuanced understanding of the relationships between BTC and altcoin volatilities.

For future research, it would be particularly intriguing to investigate the effects of BTC price changes on altcoin prices and volatilities. Testing whether lagged BTC volatility or price movements influence altcoins could offer valuable insights into temporal dependencies and lead-lag relationships within the cryptocurrency market. Such an analysis could help uncover

whether BTC serves as not only a leading indicator of market-wide volatility but also a driver of altcoin price behavior over time.

In conclusion, while BTC volatility is pivotal in shaping altcoin market dynamics, its predictive power is context-dependent. Future research that addresses existing model limitations and expands the scope to include price changes and temporal effects will be instrumental in advancing our understanding of cryptocurrency market behavior.