Forecasting Bitcoin Prices

A Comprehensive Analysis Using Regression & Time Series Techniques

Team Members

Upam Chowdhury
241001161

A N M Zahid Hossain
241002061

Noor A Tanjum Saba Amin
241000161

Mohammad Shafiur Rahman
241000661

Kazi Nabila Tasnim
241001961

Brief Overview of the Project & Bitcoin

Bitcoin: Decentralized digital currency created in 2009 by Satoshi Nakamoto.

Trading: Traded on various cryptocurrency exchanges like Coinbase, Binance, and Kraken.

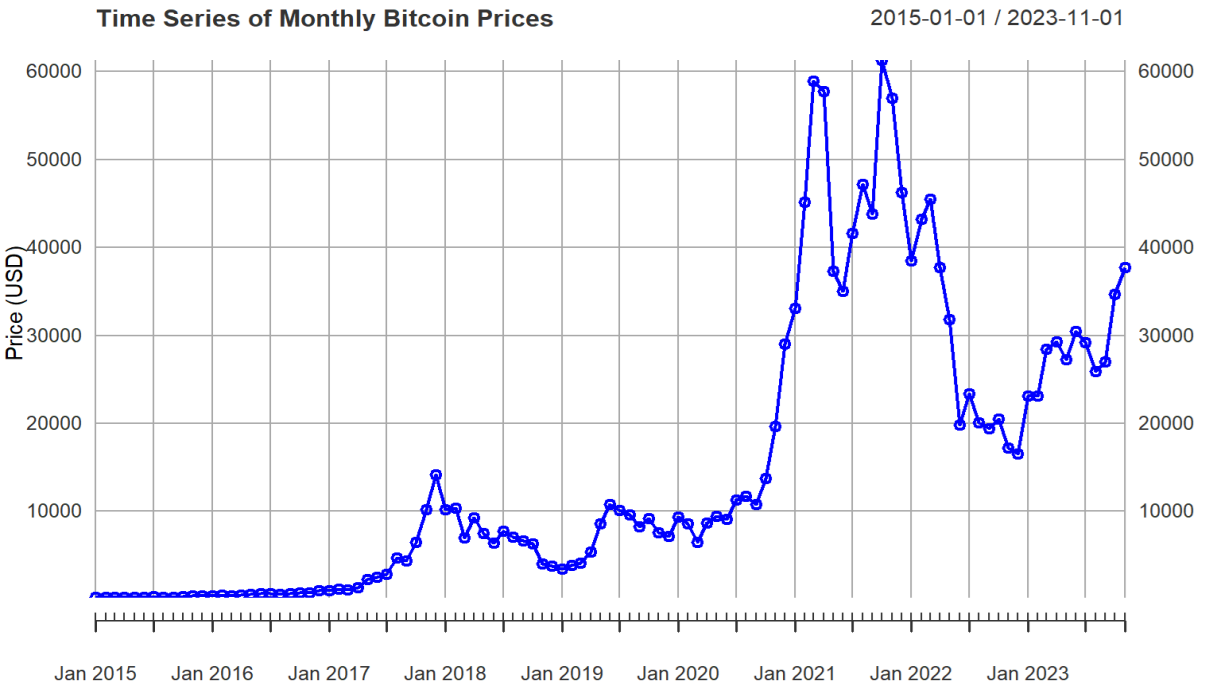
Price Volatility: Known for high volatility with significant price swings.

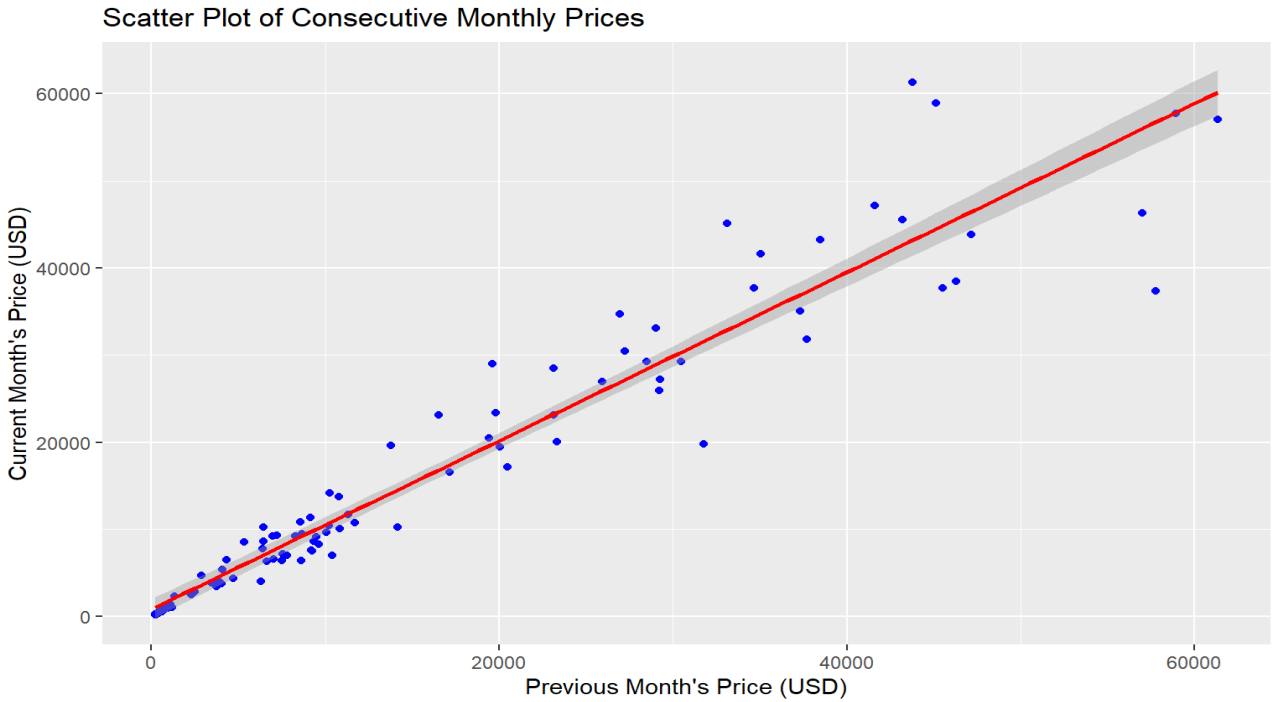
Importance: Understanding Bitcoin’s dynamics is crucial for investors, policymakers, and researchers.

Forecasting Models: Models used include ARIMA, GARCH, machine learning algorithms, and regression analysis.

Project Focus: Predicting Bitcoin prices using Regression and Time Series analysis.

Findings from Descriptive Analysis

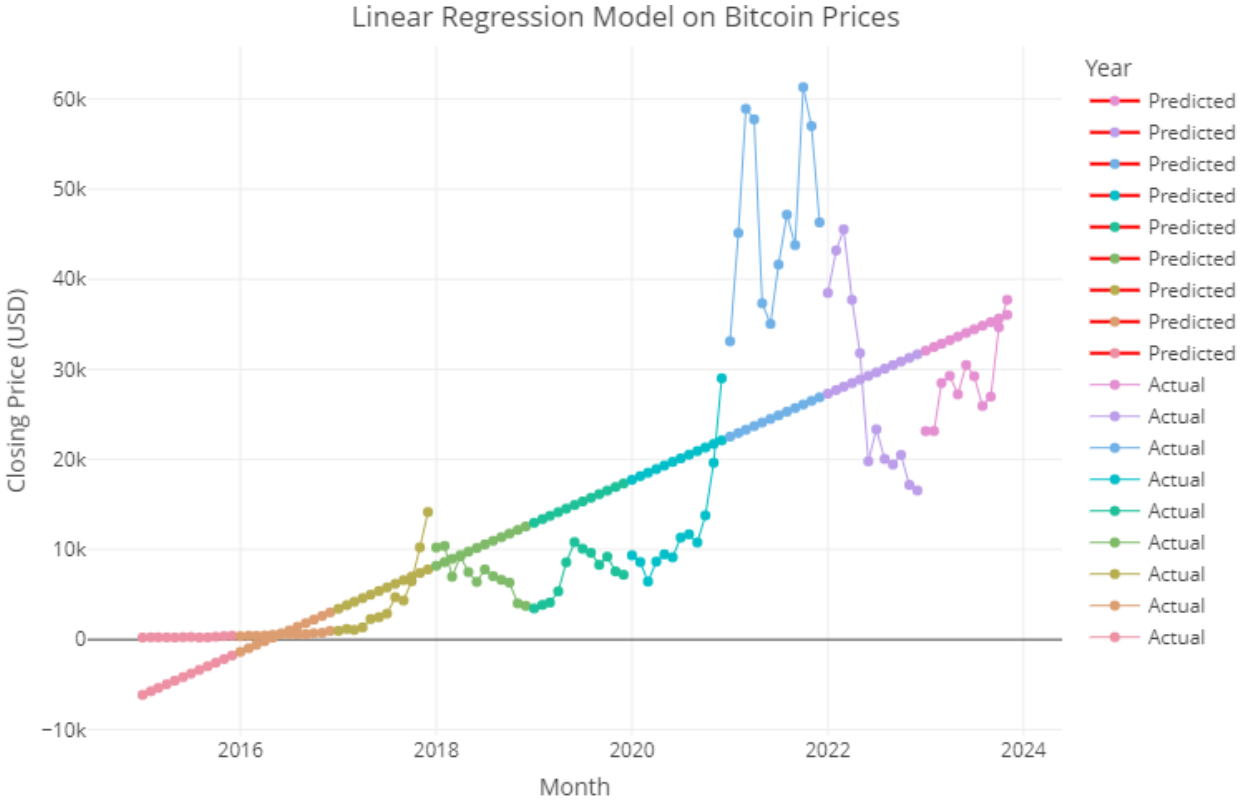


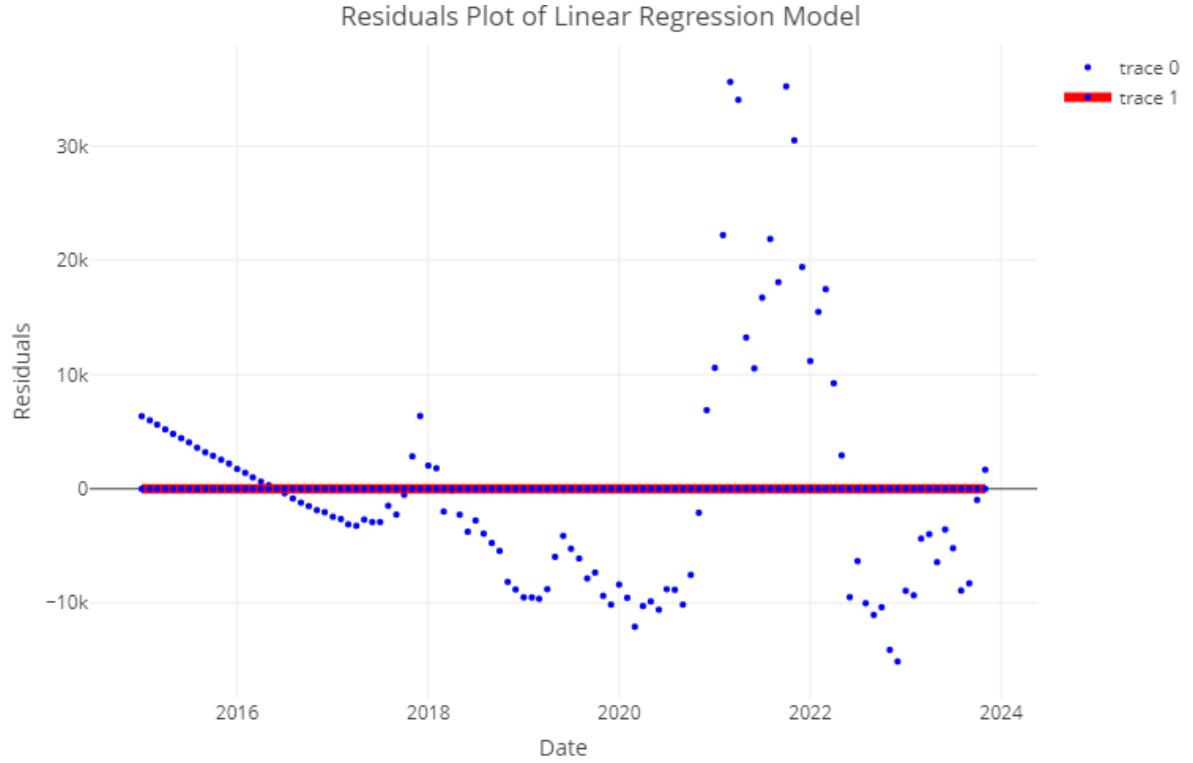


- Early years (2015-2017)
- Surge and volatility (2017-2021)
- Later years (2022-2023)

- Positive correlation
- Volatility

Findings from Linear Regression Model

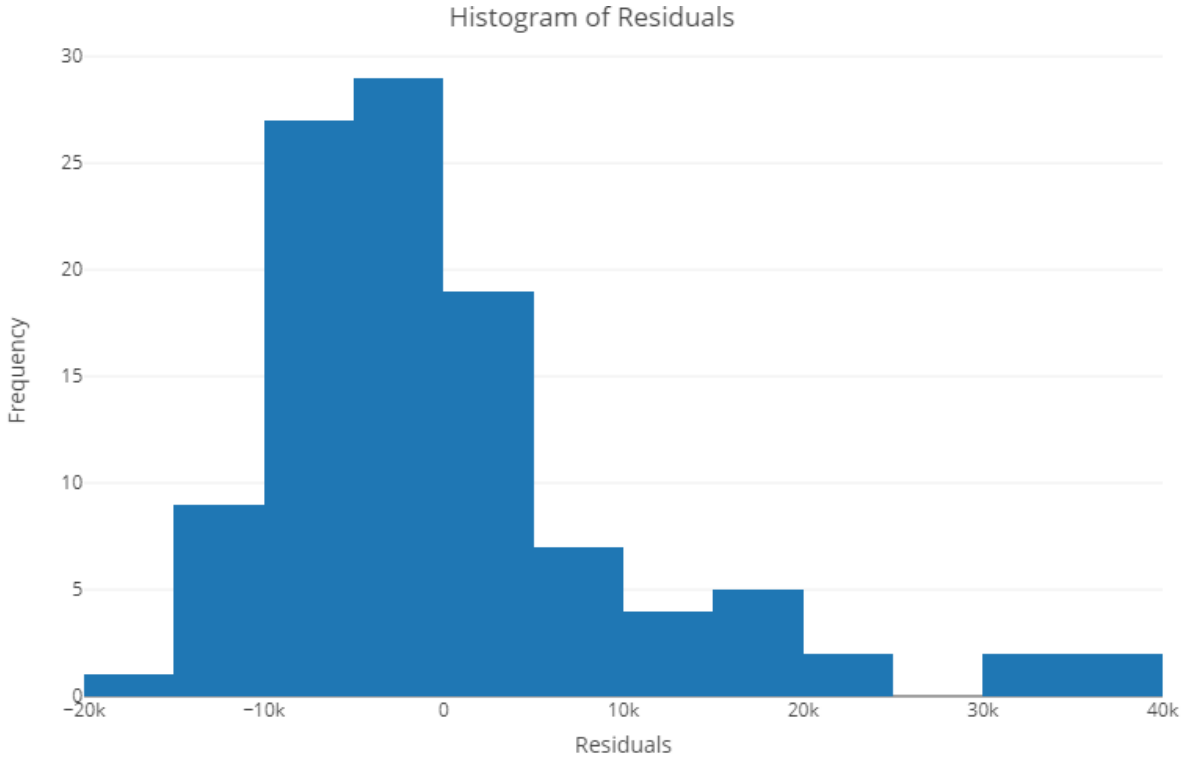




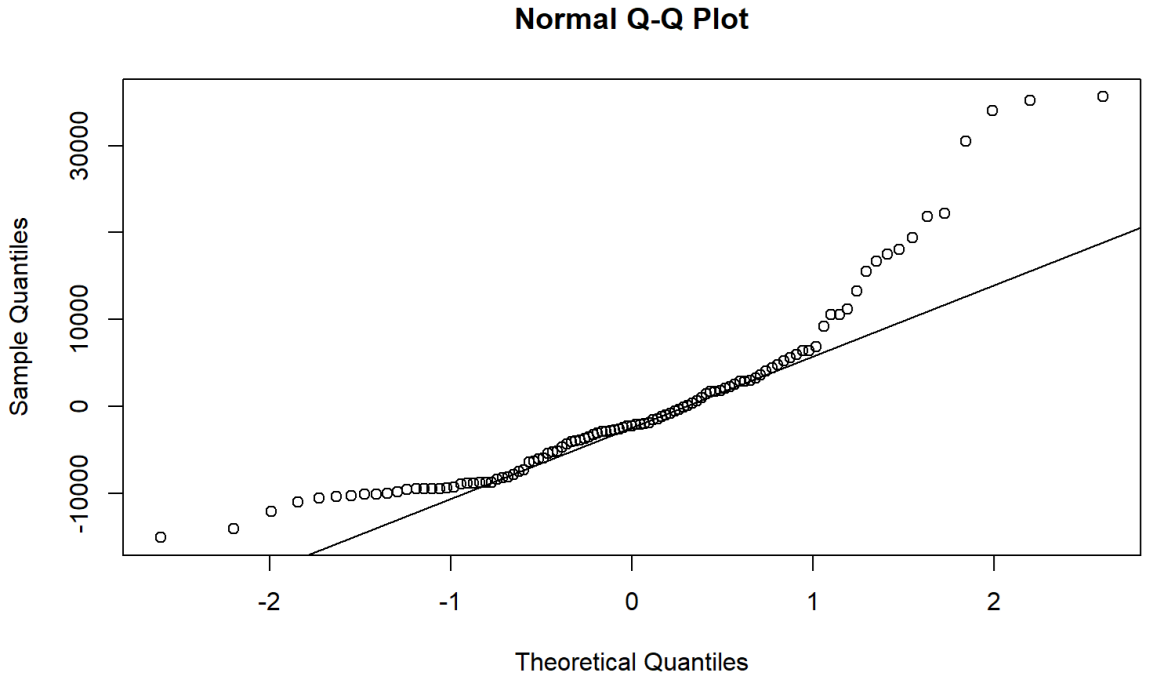
- Upward trajectory
- Price fluctuation
- Model limitation

- Homoscedasticity
- Faltered fit

Findings from Linear Regression Model



- Centered
- Price fluctuation
- Outliers
- Non-normal distribution



- Deviation from diagonal
- Tail deviation
- Non-normality

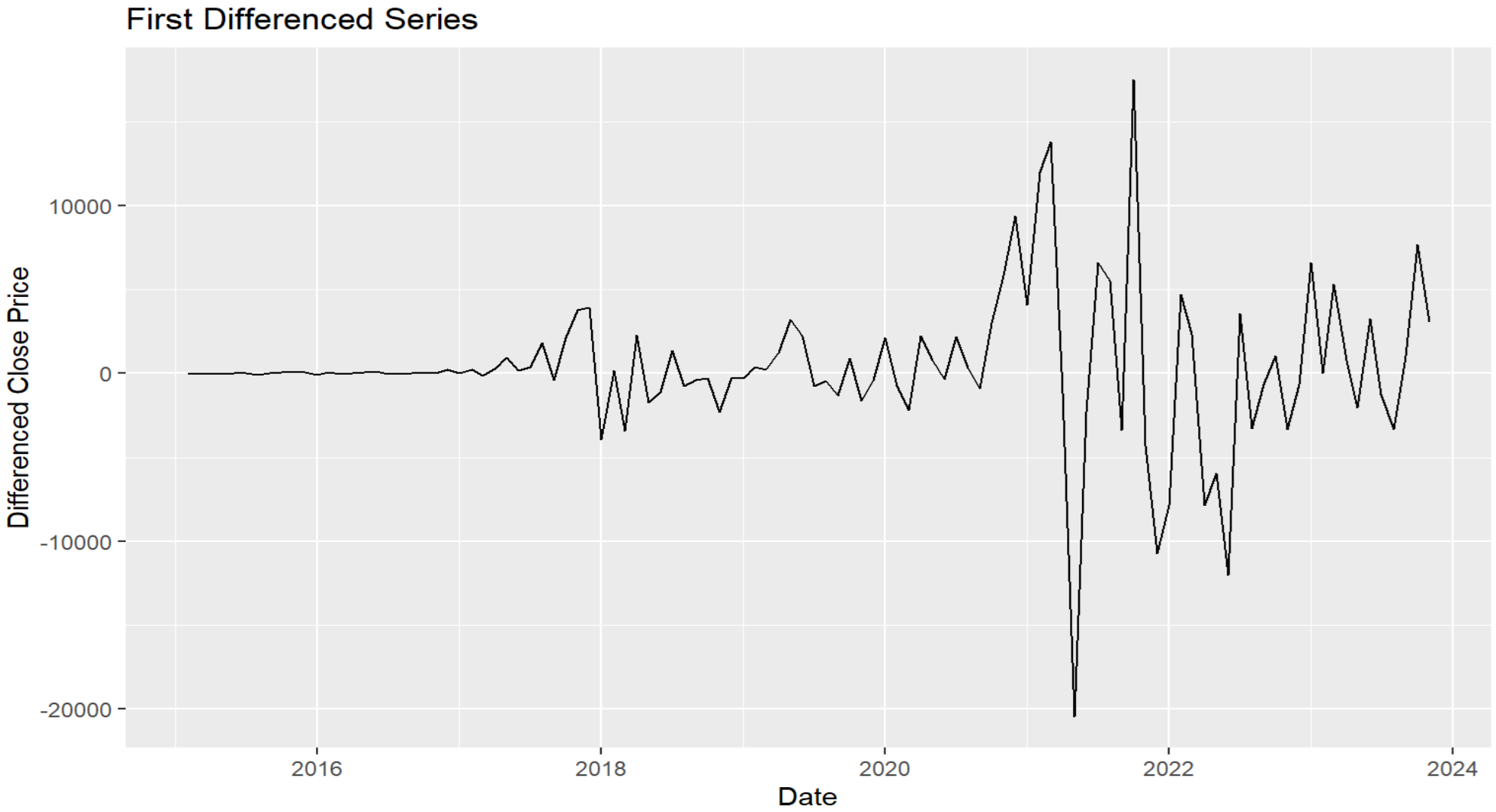
Whether Linear Regression Model is Appropriate or not

- **Wide Residual Range**: Residuals range from -15,114 to 35,626, indicating large prediction errors.
- **Model Fit**: Explains only 58.6% of price variability (R-squared = 0.586), with a residual standard error of 10,430 showing substantial unexplained variability.
- **Significant Coefficients**: While coefficients are significant, the overall model fit is poor.
- **Residual Patterns**: Residual plots reveal patterns and increasing variability, suggesting issues like heteroscedasticity.
- **Non-Normal Residuals**: The histogram shows a right-skewed distribution with outliers, indicating non-normal residuals and violating key linear regression assumptions.
- **Model Simplicity**: The model is too simple to capture the complexities of Bitcoin's price

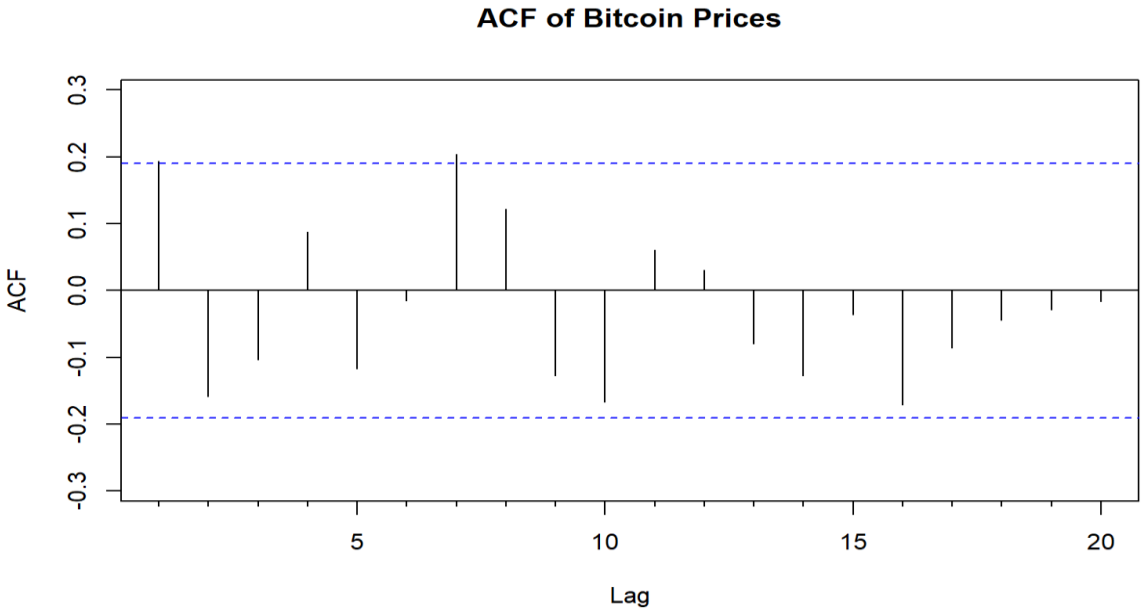
Whether Quadratic Regression Model is Appropriate or Not

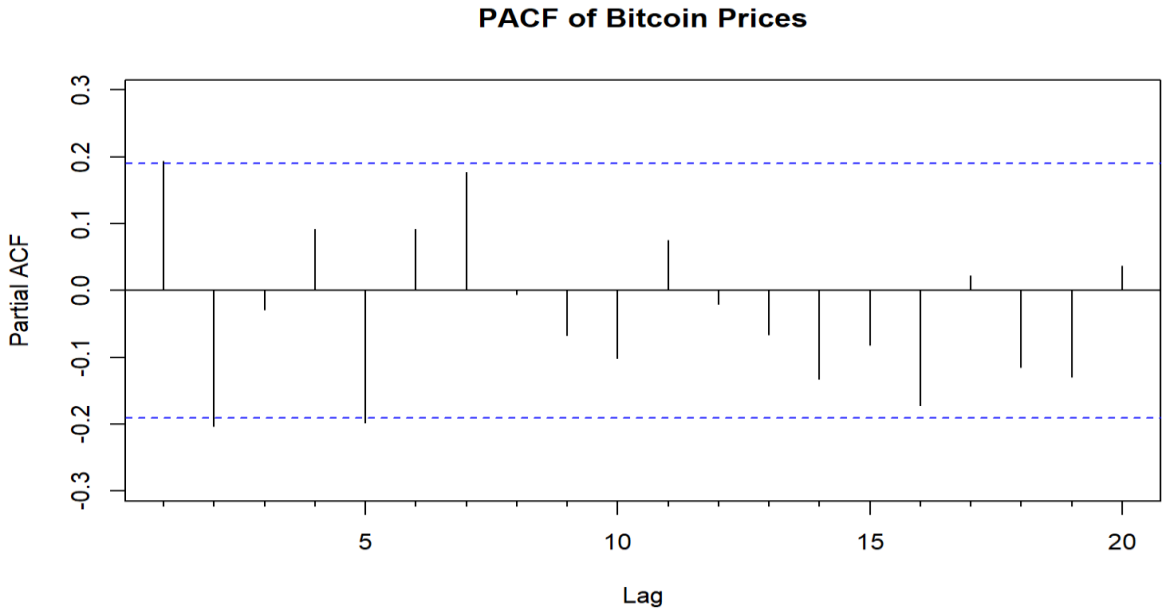
- **Model Variables:** Predicts Bitcoin’s closing prices using a second-degree polynomial of the numeric date.
- **Coefficient Significance:** Coefficients for intercept, first-degree term, and second-degree term are not statistically significant (p-values: 0.798, 0.614, 0.432).
- **Variance Explanation:** The model explains about 58% of the variability in Bitcoin’s closing prices (R-squared = 0.5885, adjusted R-squared = 0.5806).
- **Residual Standard Error:** The residual standard error is 10,450, indicating moderate dispersion of the residuals.
- **Overall Model Fit:** The model is highly significant, with an F-statistic of 74.36 and a p-value less than 2.2e-16, indicating a significant fit to the data overall. currency market.

Findings from ARIMA Model



Findings from ARIMA Model





ACF Plot Analysis:

- Starts high and declines slowly
- Indicates a potential AR or ARMA model
- Significant spikes at specific lags.

PACF Plot Analysis:

- Significant spikes at lags 1 and 2.
- Spikes drop off after lag 2.
- Suggests an AR process.

Findings from ARIMA Model

Models selected for testing:

ARIMA(1,1,1):
The ARIMA(1,1,1) model is often a good starting point because it balances simplicity with the ability to capture both AR and MA components. It's chosen because the EACF shows significance in the first AR and MA terms without too much complexity.

ARIMA(2,1,1):
Adding an additional AR term can sometimes capture more subtle dependencies in the data. The EACF shows significance at AR(2) and MA(1), suggesting this model could better fit data with a slightly more complex autoregressive structure.

ARIMA(3,1,2):
This model includes both higher-order AR and MA terms. The presence of significant terms in the EACF table at these positions suggests that this model could better capture complex dynamics and interactions between past values and past forecast errors.

Findings from ARIMA Model

ARIMA(2,1,1) Model:

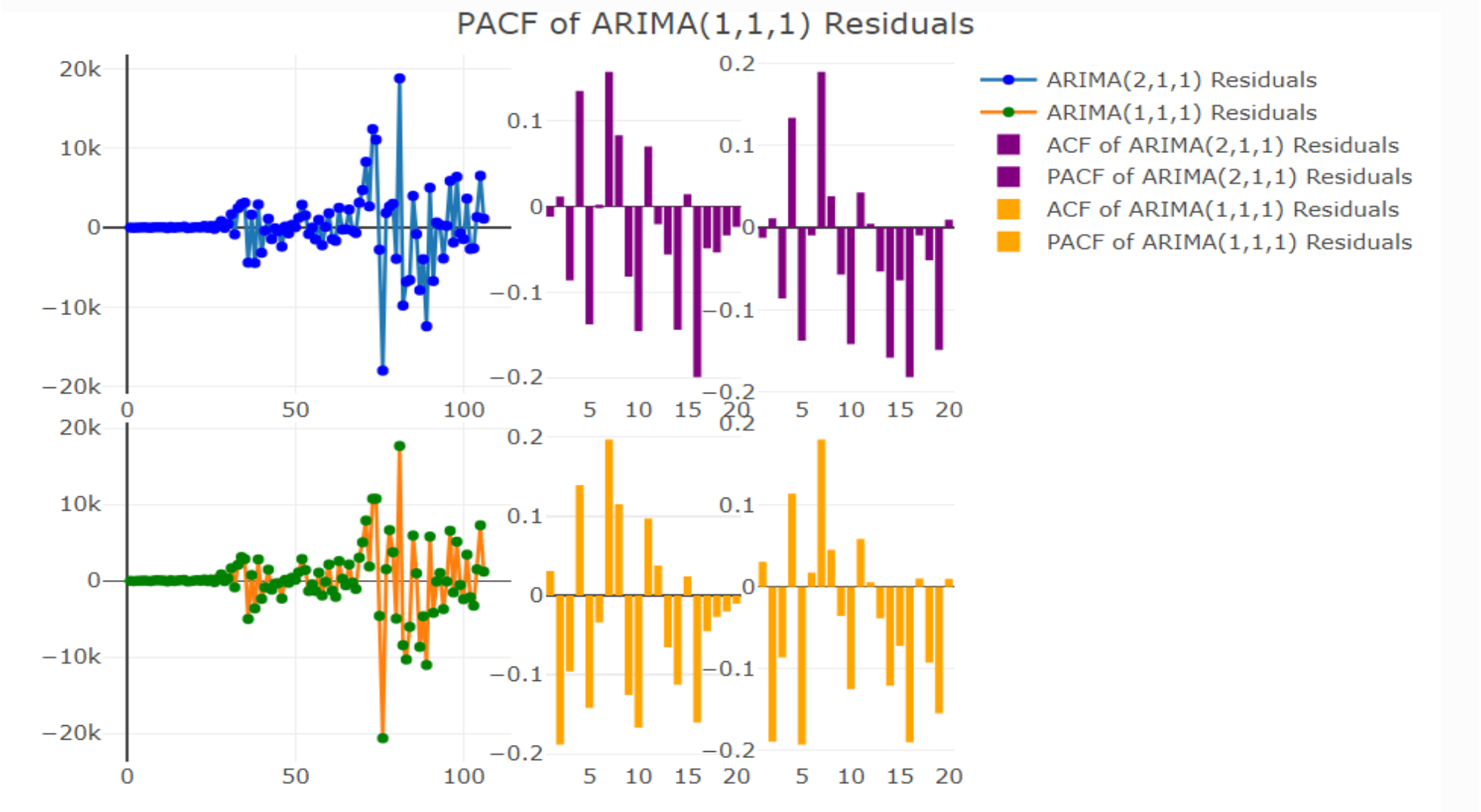
- Mean Absolute Error (MAE): 436.1328
- Mean Squared Error (MSE): 228802.7
- Root Mean Squared Error (RMSE): 478.3333
- Mean Absolute Percentage Error (MAPE): 2.756649%

.

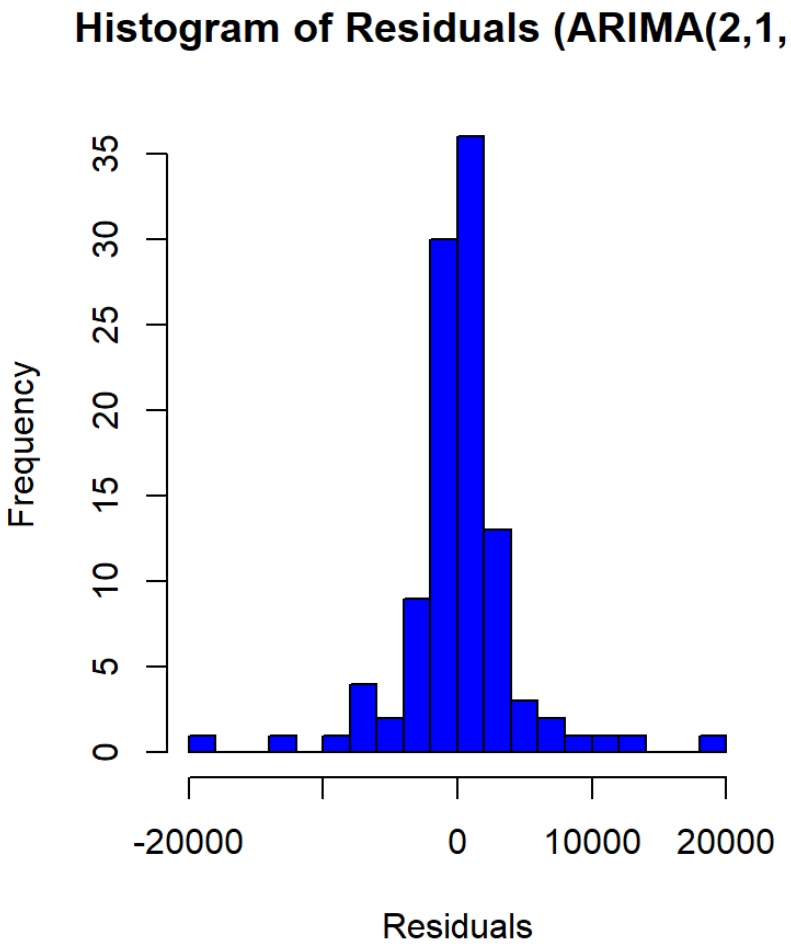
ARIMA(1,1,1) Model:

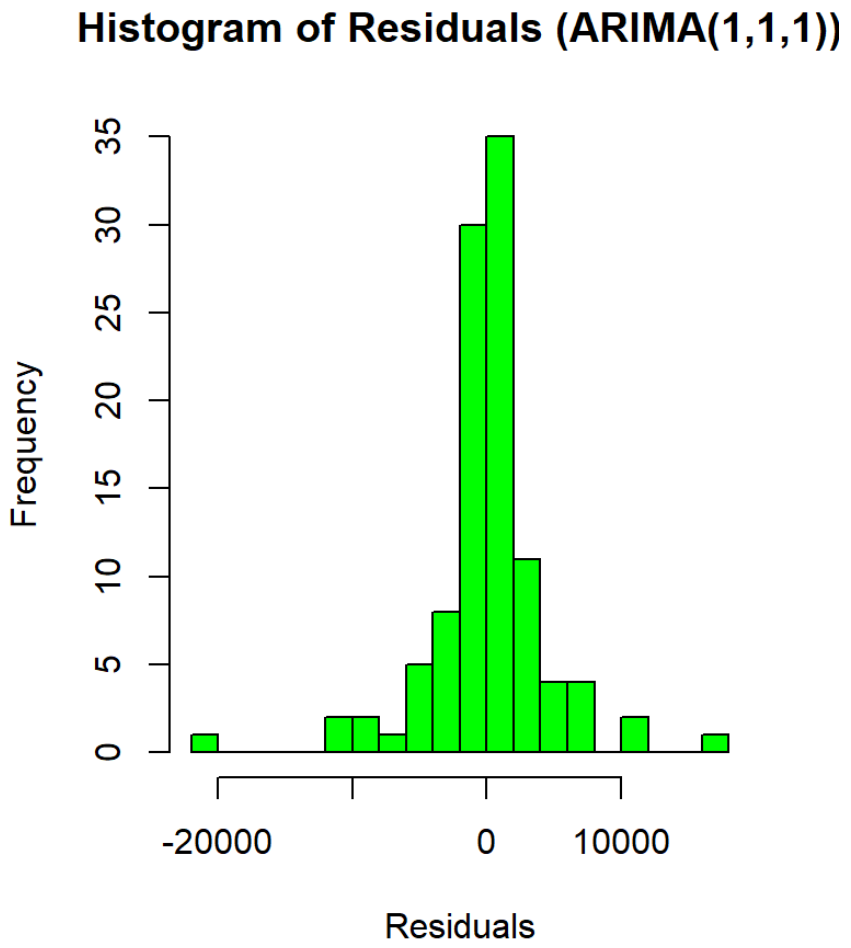
- Mean Absolute Error (MAE): 425.8315
- Mean Squared Error (MSE): 235393.7
- Root Mean Squared Error (RMSE): 485.1739
- Mean Absolute Percentage Error (MAPE): 2.698167%

Findings from ARIMA Model



Findings from ARIMA Model





ARIMA(2,1,1) Residuals:
Shape: Most residuals clustered around zero.
Symmetry: Slight right skew, indicating some higher-than-actual predictions.
Spread: Concentrated around zero with fewer extreme values, indicating a good fit.

ARIMA(1,1,1) Residuals:
Shape: Most residuals clustered around zero, similar to ARIMA(2,1,1).
Symmetry: Slight right skew, indicating some higher-than-actual predictions.
Spread: Slightly more spread out with a few more extreme values, suggesting some larger prediction errors compared to ARIMA(2,1,1).

Selected ARIMA Model

ARIMA(2,1,1) - AIC: 2066.482 - BIC: 2077.098 - MAE: 436.1328 - MSE: 228802.7 - RMSE: 478.3333 - MAPE: 2.756649%

ARIMA(1,1,1) - AIC: 2068.509 - BIC: 2076.47 - MAE: 425.8315 - MSE: 235393.7 - RMSE: 485.1739 - MAPE: 2.698167%

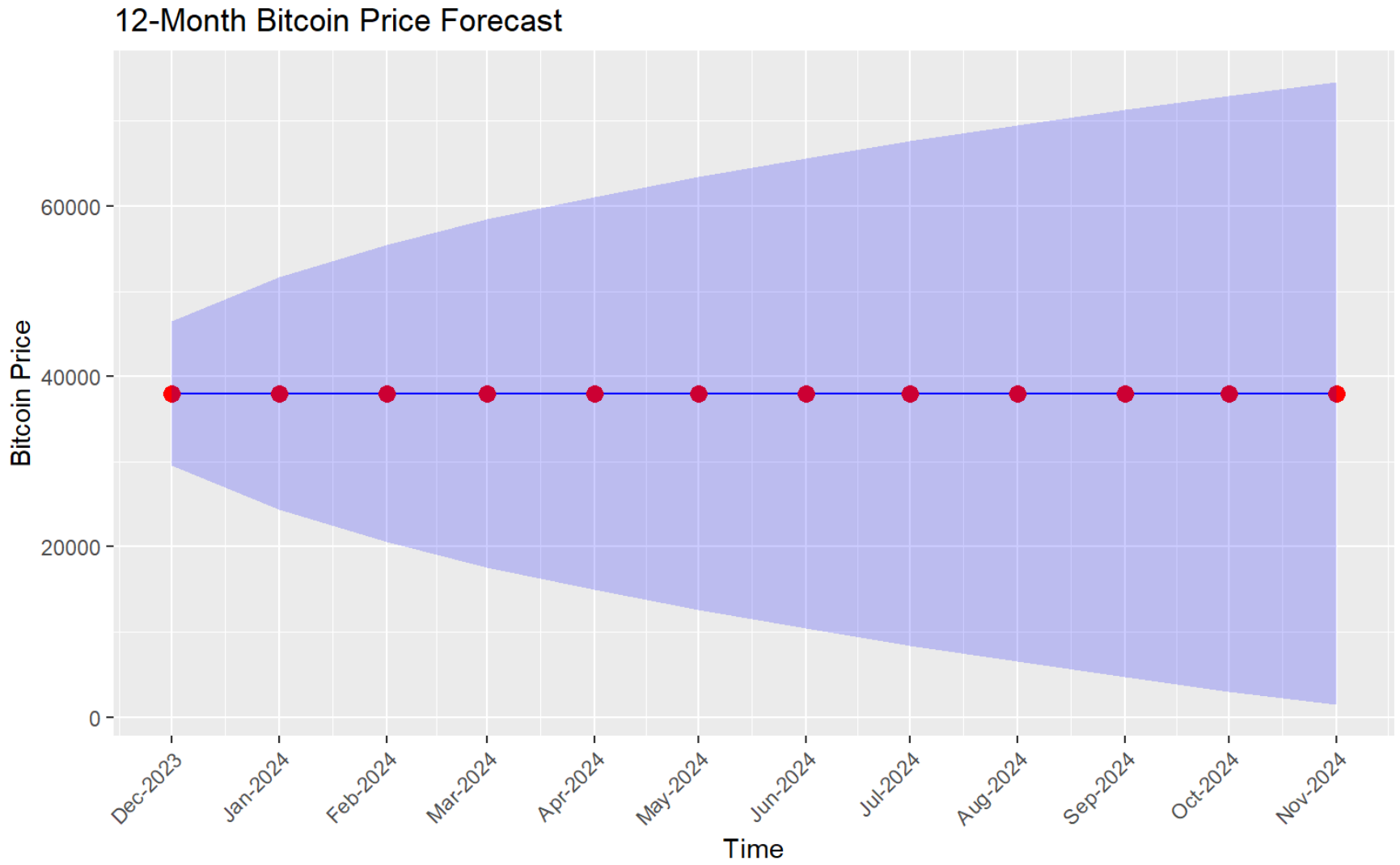
Overall Forecasting Accuracy:
Better MAE and MAPE, indicating more reliable practical forecasting.

Model Metrics:
Competitive AIC and BIC values.

Error Metrics:
Lower absolute error (MAE) and percentage error (MAPE).

Forecasting

).



.

Overall Findings from the Project

Evaluation Process:
Assessed linear, quadratic, and ARIMA models for forecasting Bitcoin prices.

Selection Rationales for ARIMA Model:
Performance: Superior performance compared to linear and quadratic models.
Adaptability: Accurately captures the complex and volatile nature of Bitcoin prices.
Components: Integrates autoregressive and moving average components for better accuracy.

Comparison with Simpler Models:
Linear and Quadratic Models:

- Struggled with Bitcoin's non-linear trends and high volatility.
- Less robust and accurate compared to the ARIMA model.

Conclusion:
ARIMA model is the optimal choice for reliable Bitcoin price forecasting in the unpredictable cryptocurrency market.


THANK YOU