Time Series Analysis of Daily Bitcoin Closing Prices

Jash Shah (202201016)*
Dhirubhai Ambani Institute of Information & Communication Technology,
Gandhinagar, Gujarat 382007, India

***Abstract*—This report presents a comprehensive time series analysis of daily Bitcoin closing prices from September 2014 to March 2025. Through techniques such as logarithmic transformation, moving averages, first-order differencing, and the Augmented Dickey-Fuller test, we address non-stationarity and volatility in the data. Residual analysis reveals the presence of heavy tails and leptokurtic behavior, which is further supported by fitted probability distributions. A positive correlation of 0.60 between trading volume and closing price is observed, indicating potential co-movement but not necessarily implying causality. This study highlights the complex yet structured behavior of Bitcoin prices and sets the stage for future forecasting or volatility modeling.**

I. Introduction

Bitcoin, the most popular cryptocurrency, experiences significant price fluctuations due to various economic and market factors. Accurate analysis of Bitcoin prices can help investors make informed decisions. This study applies time series techniques to analyze Bitcoin's closing prices.

The objectives of this research are:

- To understand the trend and seasonality of Bitcoin prices.
- To apply statistical techniques such as moving averages and first-order differencing.
- To analyze the residual after de-trending and de-seasonalizing the Bitcoin closing prices.

II. Data Overview

The dataset consists of historical Bitcoin closing prices recorded daily. The dataset is preprocessed to handle any missing values and ensure consistency.

The Bitcoin price data contains data from September 2014 to March 2025 and has been collected directly from Yahoo Finance, which is a reliable source. It includes:

- Date: Daily data as an index
- Opening Price: The price of Bitcoin at the beginning of the trading day.
- High Price: The highest price reached during the day.
- Low Price: The lowest price recorded during the day.
- Closing Price: The final price at the end of the trading day.
- Volume: The number of units traded on that day.

*Electronic address: 202201016@daiict.ac.in

We have focused more on analyzing the closing price and, to some extent, the volume of trades occurring over the last many years.

III. Exploratory Data Analysis (EDA)

EDA helps to understand the underlying patterns in the historical price data of Bitcoin, enabling the extraction of significant insights and inferences.

A. Time Series Visualization

A time series plot is created to visually identify possible trends and seasonal patterns in Bitcoin's closing prices.

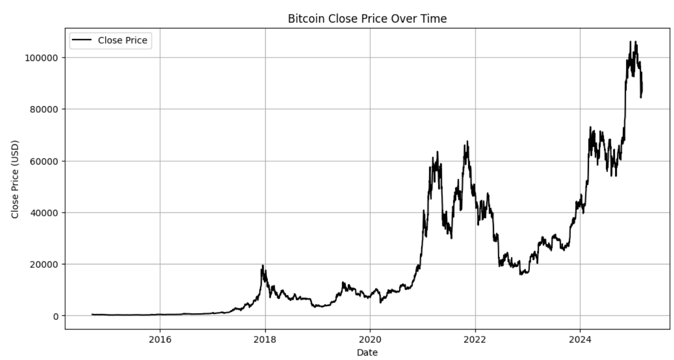

Fig. 1. Bitcoin Closing Price Over Time

B. Logarithmic Transformation

Logarithmic transformation is applied on Bitcoin closing price because of the following reasons:

- To better analyze the highly volatile data, as logarithmic transformation allows us to compress large values to smaller ones.
- To stabilize the visually increasing variance of the prices.

The Logarithmic transformation is applied as follows:

$$Y_t = \log(X_t) \tag{1}$$

where X_t represents the original Bitcoin price at time t, and Y_t is the transformed value.

C. Stationarity Check

1) Augmented Dickey-Fuller Test: To perform time series analysis effectively, important to assess whether the data is stationary. We applied the Augmented Dickey-Fuller (ADF) test to the Bitcoin closing prices to check for stationarity. The results were as follows:

- **ADF Statistic:** -2.0620
- **p-value:** 0.2601
- **Critical Values:**
  - 1%: -3.432
  - 5%: -2.862
  - 10%: -2.567

Since the p-value is greater than 0.05 and the ADF statistic is not less than any of the critical values, we fail to reject the null hypothesis. Therefore, the Bitcoin closing price time series is likely not stationary in its original form.

2) Rolling Statistics: Rolling means for different windows using the moving average method is plotted to observe any potential trends which were previously unobservable.

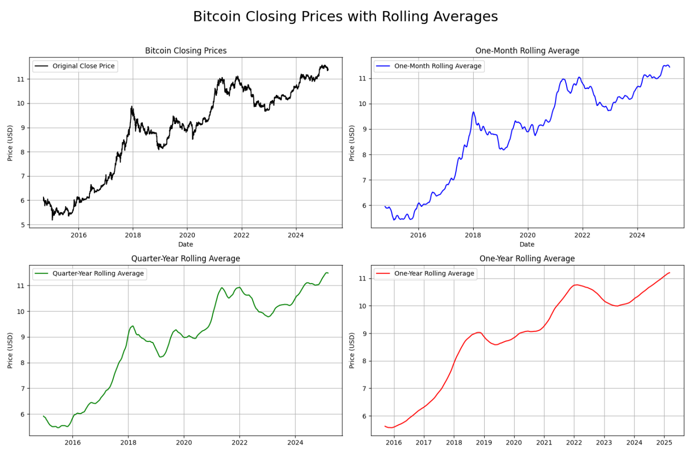

Fig. 2. Rolling Mean and Standard Deviation

D. Moving Average

A moving average with a window of k periods was applied to smoothen short-term fluctuations and observe possible trend in the data:

$$\mathrm{MA}_t = \frac{1}{k}\sum_{i=0}^{k-1} Y'_{t-i} \tag{2}$$

where MA_t represents the moving average at time t.

E. First Order Differencing

After performing the Rolling Mean analysis, it is visually observed that the data possess a somewhat linear trend. To detrend the data, we perform First Order Differencing as follows:

$$Y'_t = Y_t - Y_{t-1} \tag{3}$$

where Y'_t represents the differenced series. After performing the First Order Differencing, we plot the data.

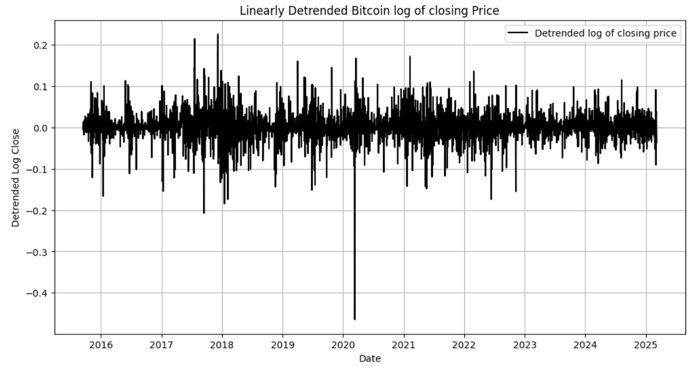

Fig. 3. First Differenced Bitcoin Price Series

The plot above visually appears to be centered around zero, indicating a constant mean, and the variance also remains stable over time (with the exception of a few outliers), suggesting weak stationarity. We performed one more ADF test on our detrended data, which yielded a negligible p-value close to 0.00, confirming our visual observation.

F. Outlier Analysis

We use the Interquantile Range (IQR) method to identify outliers, which essentially treats the first and final 25 percentile of the data as outliers.

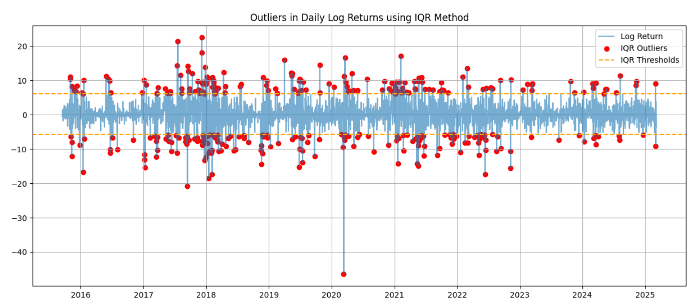

Fig. 4. Identification of Outliers Using IQR Method

The above plot suggests that, while logarithmic returns tend to remain within a narrow band, there are frequent and sometimes extreme outliers. A lot of outliers are clustered around the time of financial turmoil.

G. Analysis of the Stationary Residuals

To analyze the stationary residual, we plot the daily and monthly density distributions. This will allow us to analyze the nature of fluctuations, the frequency of outliers, and skewness of the fluctuations. In financial time series, it becomes especially important to study these fluctuations to perform a better risk analysis.

The daily density distribution plot has a significantly higher peak than a theoretical normal distribution, implying that it is leptokurtic. It also has a relatively fat tail suggesting a

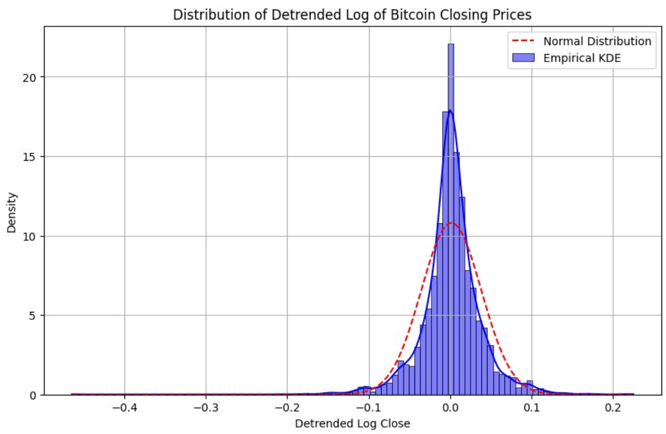

Fig. 5. Daily Density Distrubution

higher probability of extreme values which is an expected characteristic given the highly volatile nature of bitcoin price.

These observations are in line with the four statistical moments of the distribution which were calculated numerically:

- **Mean:** 0.001601
- **Variance:** 0.001366
- **Skewness:** -0.671608
- **Kurtosis:** 11.666637

We also corroborated these observations with the help of a Q-Q plot, which shows that normality may be followed for the central values such as mean and median, but has deviation in the tails, which shows that the distribution may have heavier tails than a theoretical normal distribution. Overall this indicates slight deviation from normality, particularly in the extremes, which is common in financial time series data.

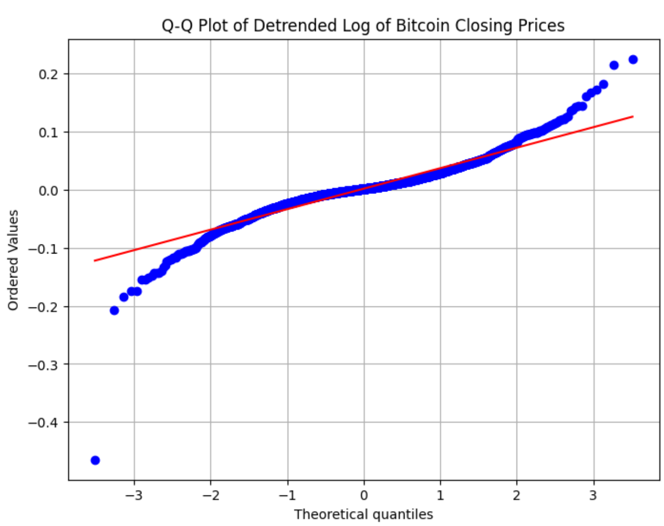

Fig. 6. Q-Q Plot to visualize distribution of the residuals

We then proceeded to check the distribution of the residuals in a monthly moving average. We observed that the distribution was more central than daily residuals due to the aggregation smoothing out volatility and extremes. Daily residuals retain more raw volatility and noise, resulting in a heavier-tailed, less ”central” shape.

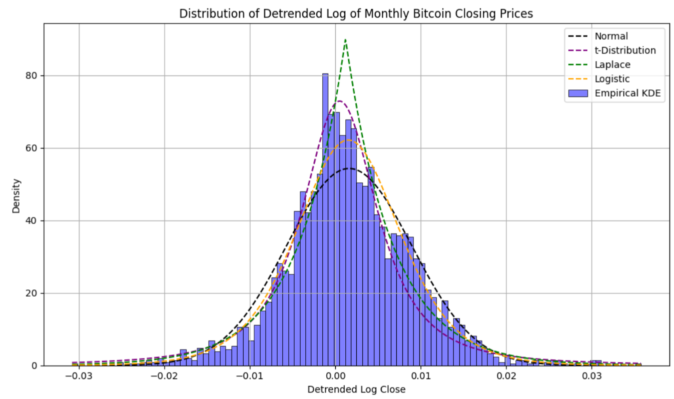

Fig. 7. Fitted Distributions of Monthly Residuals

We fitted multiple distributions to check for the best fit and we observed that the t-distribution fits best to the monthly distribution.

1) Auto-correlation in Detrended Data: To analyze the residuals of our data, we examined the autocorrelation of the detrended log-transformed series. The ACF plot indicates that, apart from the lag at zero, all autocorrelations lie within the 95% confidence interval, suggesting that the series exhibits no significant autocorrelation.

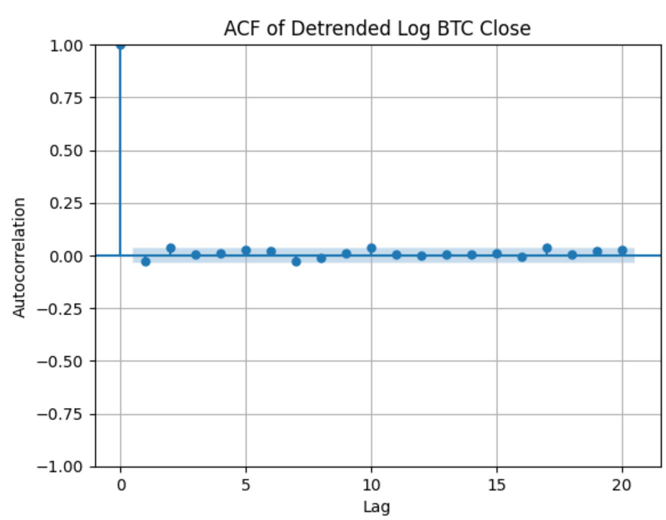

Fig. 8. ACF of Detrended Data

2) Partial Auto-correlation in Detrended Data: To further analyze the temporal structure of the detrended log-transformed Bitcoin closing prices, we examined the Partial Autocorrelation Function (PACF). The PACF helps in identifying the direct relationship between a time series and its lagged values, controlling for the influence of intervening lags.

As illustrated in Fig. 9, the partial autocorrelation at lag 1 is significantly different from zero, while all higher-order lags fall well within the 95% confidence interval. This indicates that only the first lag has a statistically significant effect on the current value, and subsequent lags do not contribute meaningful information.

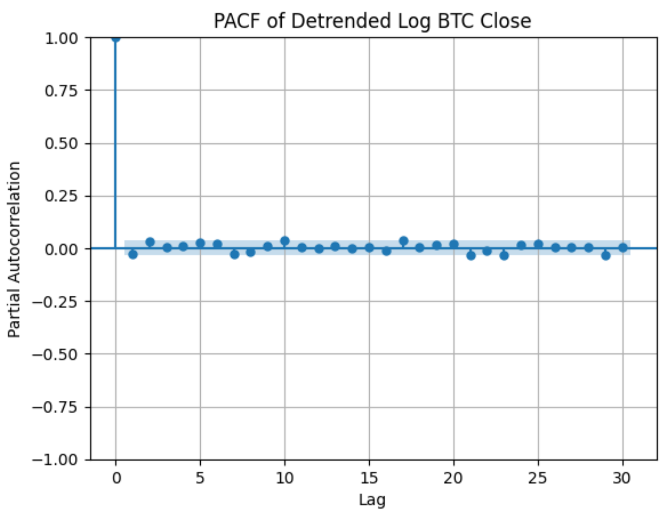

Fig. 9. PACF of Detrended Data

IV. Conclusion

This study explored the dynamic behavior of Bitcoin's daily closing prices using a variety of time series techniques. Through logarithmic transformation and differencing, we effectively addressed the non-stationarity inherent in the raw data. Visual tools such as rolling statistics and autocorrelation plots helped uncover structural characteristics, while statistical tests confirmed the presence of heavy tails and excess kurtosis in the residuals — a common trait in financial time series.

Our analysis revealed a moderate positive correlation between trading volume and price, suggesting a degree of co-movement, though without implying direct causality. Furthermore, fitting multiple probability distributions to the residuals showed that the data exhibits strong leptokurtic behavior, with the t-distribution providing the best fit.

Overall, the results tell us about the complex, volatile nature of Bitcoin prices, while showing that statistical techniques can still give us meaningful patterns and characteristics within the data.

V. Future Scope

While this study has explored the highly volatile nature of Bitcoin's daily closing prices, there remain several areas for future work. These maybe outlined as follows:

- **Forecasting**: Future work could deal with applying robust predictive models such as LSTM Regression which can account for highly volatile nature of the data.
- **Volatility Modeling**: Volatility models such as ARCH or GARCH can be used to better understand clustering and persistence of volatility given the leptokurtic and fat-tailed nature of residual density distribution.
- **Multivariate Analysis**: Including other aspects such as volume of trade or prices of other cryptocurrencies which would allow us to do a more in-depth analysis of trends in closing prices.

VI. Link to Google Colab Notebook

All analyses, visualizations, and statistical computations were performed using Python in a Google Colab notebook. The complete code and interactive results can be accessed at the following link: Colab Notebook

References

[1] D. A. Dickey and W. A. Fuller, "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, vol. 74, no. 366, pp. 427–431, 1979.
[2] Yahoo Finance. "Bitcoin Historical Price Data," 2025.