Time series analysis of change in Bitcoin prices during 2014 - 2022 to forecast future prices

• Table of contents :

• Introduction :

In recent years, Bitcoin has emerged in the world of Finance and Currency, captivating the attention of investors, technology enthusiasts, business men, and economists. The decentralized nature and blockchain technology, with its potential for substantial result has caused an exponential growth in its usage. However, due to its extreme uncertain nature, making accurate predictions of Bitcoin prices has become a very complex and challenging task.

The aim of this project is to analyse past data from previous 8 years spanning from 2014 to 2022 to forecast future Bitcoin prices. Using advanced statistical techniques and analysis the changes in data for 8 years, we seek to develop a reliable forecasting model which can be used by investors, traders and researchers to to take decisions and manage risk effectively.

This project seeks to leverage time series analysis to contribute to the growing field of knowledge surrounding Bitcoin and also will provide valuable insights into the future of digital currencies as a whole. As the world moves towards greater adoption of cryptocurrencies, the ability to forecast Bitcoin prices becomes increasingly vital to make rapid decisions in this fast pacing world of digital finance.

• What is Bitcoin :

Bitcoin is a decentralized digital currency that was created in 2009 by an anonymous existent or group of individualities. Unlike traditional edict currencies issued by governments, Bitcoin isn't controlled or regulated by any central authority or fiscal institution.

At its core, Bitcoin is a peer-to-peer electronic cash system that allows users to send and receive payments directly without the need for intermediaries such as banks. Transcations are verified by network nodes through cryptography anfd recorded on the blockchain, ensuring transparency and security. The bitcoin blockchain is a public tally that records bitcoin deals. It's enforced as a chain of blocks, each block containing a cryptographic hash of the former block up to the birth block in the chain. A network of communicating bumps running bitcoin software maintains the blockchain.

Bitcoin has gained significant attention and fashionability due to its eventuality as a store of value, a medium of exchange, and an investment asset. One of the crucial features of Bitcoin is its limited force. There will only ever be 21 million Bitcoins in actuality, which makes it a scarce asset. This failure, along with the growing demand and relinquishment, has contributed to its value appreciation over time.

• Abstract :

- ## Relevancy of the topic :

Time series analysis on Bitcoin prices is important because it will help us to realize that trend in the price change, also predict its future purposes. People who are interested to invest will be benefitted from this analysis as it will help them to realize when to invest to get maximum profit by forecasting the prices.

- ## Objective of the overall analysis :

Bitcoin is a cryptocurrency, a virtual currency designed to act as money and a form of payment outside the control of any one person, group or entity, thus removing the need for third party involvement in financial transactions. Time series analysis would be a good way to analyse or study the trend in change of prices of bitcoin in the period 2014-2022, also to detect any cyclical or seasonal pattern. This project will contain the visualization of data, identification of trend and fitting of mathematical functions. We will also fit an autoregressive model which is used for forecasting. We are going to use different softwares like R, excel and minitab to help analyse the data.

- ## Dataset description:

This dataset is collected from kaggle which contains a list of variables like date, open, high, low, close, adj close, volume etc.

https://www.kaggle.com/datasets/yasserh/bitcoin-prices-dataset

- ## Conclusion:

From the analysis, we might see that there is an increasing trend in the prices of the bitcoin and we also might be able to fit an autoregressive model for future prediction

• Data :

The time series data comprising Bitcoin prices from 2014 to 2022 provides a comprehensive and valuable resource for understanding the historical behaviour and trends of this cryptocurrency. This dataset includes columns such as High, Low, Open, Close, adj. close, volume which are sorted according to dates, which collectively offer insights into various aspects of Bitcoin's price movements.

The "High" column represents the highest recorded price of Bitcoin during a specific time period, while the "low" column indicates the lowest recorded prices. These data points help identify the price range within which Bitcoin fluctuated over time. The "Open" column represents the price at the beginning of the day and the "Close" column signifies the price of Bitcoin at the end of the day.
The column "adj. Close" represents the adjusted closing prices of Bitcoin which accounts for any corporate actions, such as stock splits and dividend payments, that may have affected the price. The "volume" column indicates the total number of total number of Bitcoin units traded during a specific time period. It represents the market activity and liquidity for Bitcoin, reflecting the level of interest and participation from buyers and sellers.

This dataset has been taken from the website "kaggle" and the link is given below:

https://www.kaggle.com/datasets/yasserh/bitcoin-prices-dataset

The first 30 rows of the dataset :

Date	Open	High	Low	Close	Adj Close	Volume
17-09-2014	465.864	468.174	452.422	457.334	457.334	21056800
18-09-2014	456.86	456.86	413.104	424.44	424.44	34483200
19-09-2014	424.103	427.835	384.532	394.796	394.796	37919700
20-09-2014	394.673	423.296	389.883	408.904	408.904	36863600
21-09-2014	408.085	412.426	393.181	398.821	398.821	26580100
22-09-2014	399.1	406.916	397.13	402.152	402.152	24127600
23-09-2014	402.092	441.557	396.197	435.791	435.791	45099500
24-09-2014	435.751	436.112	421.132	423.205	423.205	30627700
25-09-2014	423.156	423.52	409.468	411.574	411.574	26814400
26-09-2014	411.429	414.938	400.009	404.425	404.425	21460800
27-09-2014	403.556	406.623	397.372	399.52	399.52	15029300
28-09-2014	399.471	401.017	374.332	377.181	377.181	23613300
29-09-2014	376.928	385.211	372.24	375.467	375.467	32497700
30-09-2014	376.088	390.977	373.443	386.944	386.944	34707300
01-10-2014	387.427	391.379	380.78	383.615	383.615	26229400
02-10-2014	383.988	385.497	372.946	375.072	375.072	21777700
03-10-2014	375.181	377.695	357.859	359.512	359.512	30901200
04-10-2014	359.892	364.487	325.886	328.866	328.866	47236500
05-10-2014	328.916	341.801	289.296	320.51	320.51	83308096
06-10-2014	320.389	345.134	302.56	330.079	330.079	79011800
07-10-2014	330.584	339.247	320.482	336.187	336.187	49199900
08-10-2014	336.116	354.364	327.188	352.94	352.94	54736300
09-10-2014	352.748	382.726	347.687	365.026	365.026	83641104
10-10-2014	364.687	375.067	352.963	361.562	361.562	43665700
11-10-2014	361.362	367.191	355.951	362.299	362.299	13345200
12-10-2014	362.606	379.433	356.144	378.549	378.549	17552800
13-10-2014	377.921	397.226	368.897	390.414	390.414	35221400
14-10-2014	391.692	411.698	391.324	400.87	400.87	38491500
15-10-2014	400.955	402.227	388.766	394.773	394.773	25267100
16-10-2014	394.518	398.807	373.07	382.556	382.556	26990000

The last 30 rows of the dataset :

Date	Open	High	Low	Close	Adj Close	Volume
22-12-2021	48937.1	49544.8	48450.94	48628.51	48628.51	2.44E+10
23-12-2021	48626.34	51332.34	48065.84	50784.54	50784.54	2.82E+10
24-12-2021	50806.05	51814.03	50514.5	50822.2	50822.2	2.44E+10
25-12-2021	50854.92	51176.6	50236.71	50429.86	50429.86	1.9E+10
26-12-2021	50428.69	51196.38	49623.11	50809.52	50809.52	2.1E+10
27-12-2021	50802.61	51956.33	50499.47	50640.42	50640.42	2.43E+10
28-12-2021	50679.86	50679.86	47414.21	47588.86	47588.86	3.34E+10
29-12-2021	47623.87	48119.74	46201.5	46444.71	46444.71	3E+10
30-12-2021	46490.61	47879.96	46060.31	47178.13	47178.13	2.67E+10
31-12-2021	47169.37	48472.53	45819.95	46306.45	46306.45	3.7E+10
01-01-2022	46311.75	47827.31	46288.48	47686.81	47686.81	2.46E+10
02-01-2022	47680.93	47881.41	46856.94	47345.22	47345.22	2.8E+10
03-01-2022	47343.54	47510.73	45835.96	46458.12	46458.12	3.31E+10
04-01-2022	46458.85	47406.55	45752.46	45897.57	45897.57	4.25E+10
05-01-2022	45899.36	46929.05	42798.22	43569	43569	3.69E+10
06-01-2022	43565.51	43748.72	42645.54	43160.93	43160.93	3.02E+10
07-01-2022	43153.57	43153.57	41077.45	41557.9	41557.9	8.42E+10
08-01-2022	41561.46	42228.94	40672.28	41733.94	41733.94	2.81E+10
09-01-2022	41734.73	42663.95	41338.16	41911.6	41911.6	2.13E+10
10-01-2022	41910.23	42199.48	39796.57	41821.26	41821.26	3.21E+10
11-01-2022	41819.51	43001.16	41407.75	42735.86	42735.86	2.63E+10
12-01-2022	42742.18	44135.37	42528.99	43949.1	43949.1	3.35E+10
13-01-2022	43946.74	44278.42	42447.04	42591.57	42591.57	4.77E+10
14-01-2022	42598.87	43346.69	41982.62	43099.7	43099.7	2.36E+10
15-01-2022	43101.9	43724.67	42669.04	43177.4	43177.4	1.84E+10
16-01-2022	43172.04	43436.81	42691.02	43113.88	43113.88	1.79E+10
17-01-2022	43118.12	43179.39	41680.32	42250.55	42250.55	2.17E+10
18-01-2022	42250.07	42534.4	41392.21	42375.63	42375.63	2.24E+10
19-01-2022	42374.04	42478.3	41242.91	41744.33	41744.33	2.31E+10
20-01-2022	41736.53	42034.73	41724.46	41933.55	41933.55	2.35E+10

- # Methodology:

We will first visualize the data which will help us understand the increasing or decreasing trend of the data as well as give an overall view of the data.

Using trend analysis to fit mathematical equations to the data which results in detecting underlying pattern or trend in the data. However, the fitted equations might contain seasonality and trend which can cause inaccurate forecasting. Hence, we have to analyse the data more accurately using different autoregressive or moving average modes.

Time series forecasting of this Bitcoin price data is quite different from other forecastings because

1. It is time dependent, so the basic assumption of a linear regression model that the observations are independent does not hold in this case.

2. Along with an increasing or decreasing trend, most time series have some form of seasonality trends, i.e. variations specific to a particular time frame.

When working with Bitcoin data, it is common to encounter non-stationarity due to its inherit volatile nature and changing trends. Stationarity is a crucial concept when analysing time series data, as it ensures that the statistical properties of the series remains constant over time. To check stationarity, we will use Augmented Dickey Fuller test.

If the data is stationary, we will proceed to further analysis, and if not, we have to make it stationary. One commonly used technique is log transforming the series, which stabilizes the variance and mitigates the impact of extreme values. Additionally, using differencing can help to remove any trend or seasonality by calculating the difference between consecutive observations.

After making the data stationary, different order models can be fitted to forecast future prices. AR(autoregressive), MA(Moving Average) or ARIMA(Autoregressive Integrated Moving Average) models can be used. AR models use the previous values of the series to predict future values, while MA models incorporate past forecast errors. ARIMA models combine both autoregressive and moving average components while also accounting for differencing to achieve stationarity. We will then compare the errors of different values to understand which model is best suited for forecasting future prices.

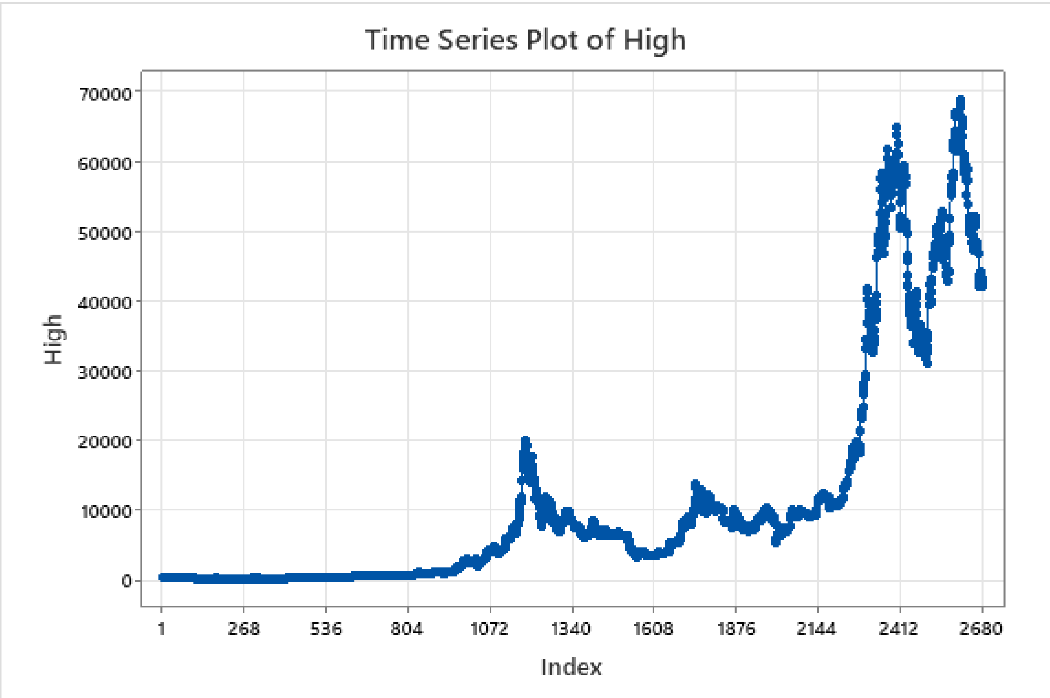
Softwares like Minitab, excel are going to be the ones used in making this analysis.

• Data visualization :

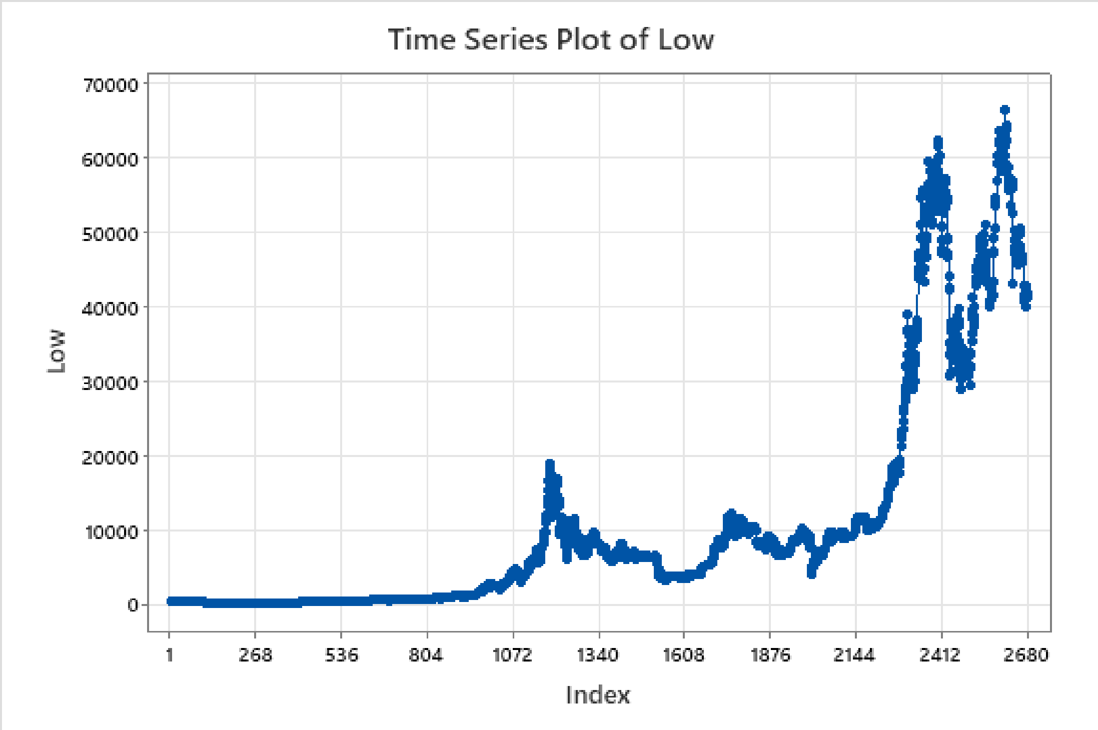
Data visualization of Bitcoin's prices over time provides a concise and insightful summary of the changes during 2014 – 2022. By plotting these values on a graph against their corresponding dates, we can gain a visual understanding of Bitcoin point fluctuations and potential patterns.

Visualization can help investors and traders to make mindful decisions. It helps identify optimal entry or exit points, understand price ranges for setting stop-loss or take-profit levels and also gauge the overall market sentiment.

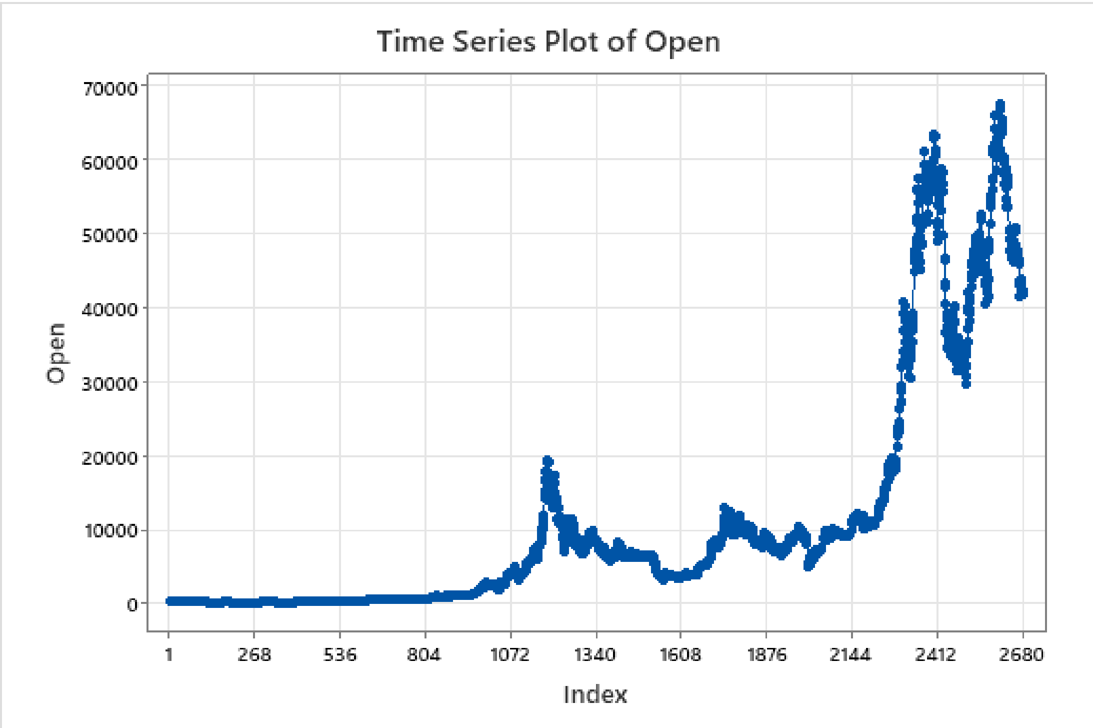
Time series plot for "High" :



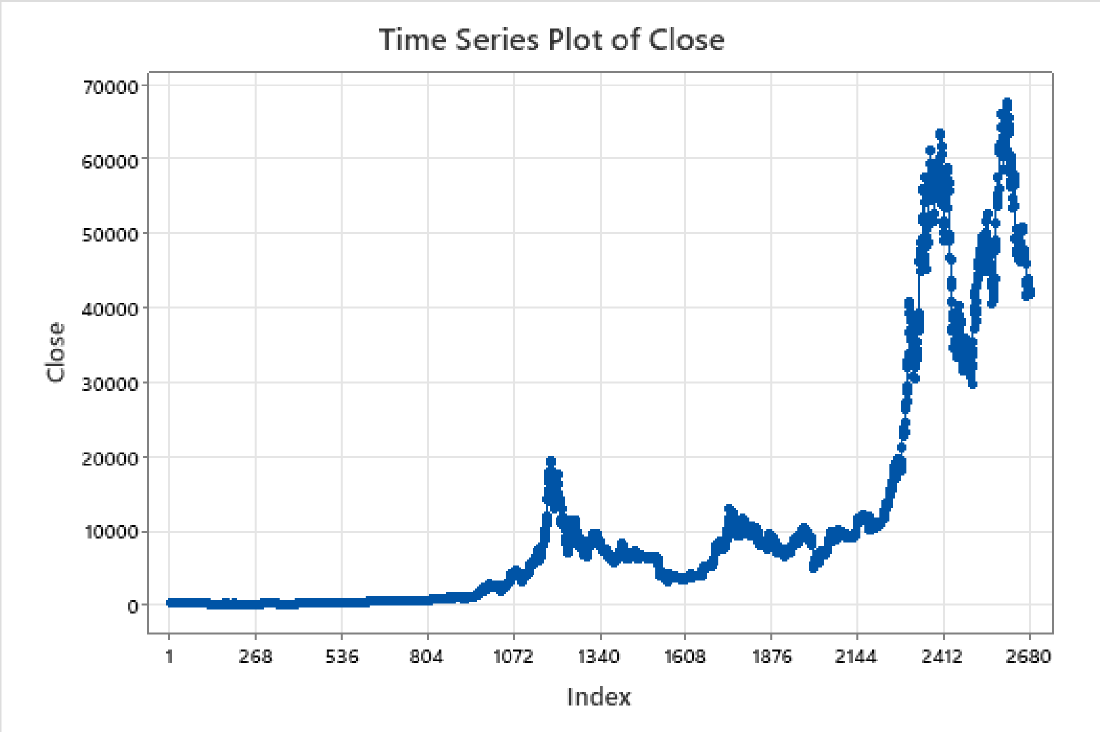
Time series plot for "Low" :



Time Series plot for "Open" :



Time Series plot for “Close” :



In these graphs, we can clearly see significant and consistent upward trend which provides the evidence of Bitcoin’s long term value appreciation and the potential for profit upon investment over time. This clear increasing trend in price shows the relativity of Bitcoin in the finance and cryptocurrency world highlights the potential for investors to gain huge profits by investing in Bitcoin.

However, it is necessary to understand that past do not guarantee future performance as Bitcoin is a highly volatile asset and is susceptible to market forces. This is why, it is necessary for us to further analyse the data and try to fit a close to accurate forecasting model to predict future prices.

• Trend analysis:

Fitting a mathematical trend equation to this time series data will help us to find a functional form that best represents the underlying pattern or trend within the data. This will help us to capture the relationship between Bitcoin prices and time over a given period. We are going to try to fit linear, polynomial or exponential curve to estimate the parameters that optimize the fit into the data.

1. Fitting trend equation to "High" :

We will try to fit a linear curve to the data.

Method

Model type	Linear Trend Model
Data	High
Length	2683
NMissing	0

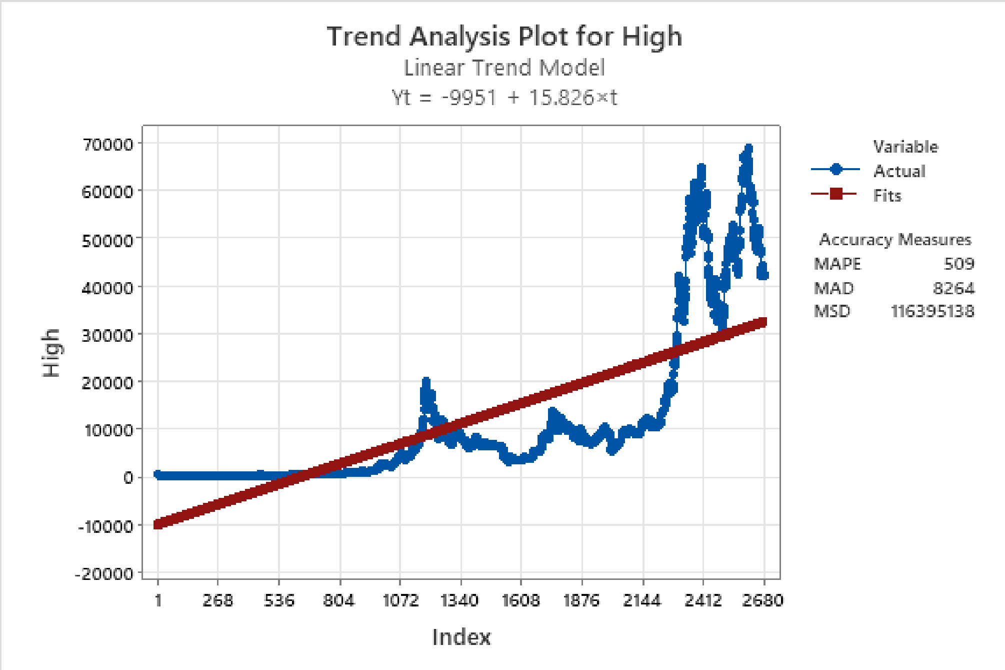
Accuracy Measures

MAPE	509
MAD	8264
MSD	116395138

And the fitted trend equation is:

Fitted Trend Equation

$$Yt = -9951 + 15.826 \times t$$



As we can see, the MAPE value is 509 which is way too high, hence this is not a good fit.

Now we will try to fit an exponential curve.

Method

Model type	Growth Curve Model
Data	High
Length	2683
NMissing	0

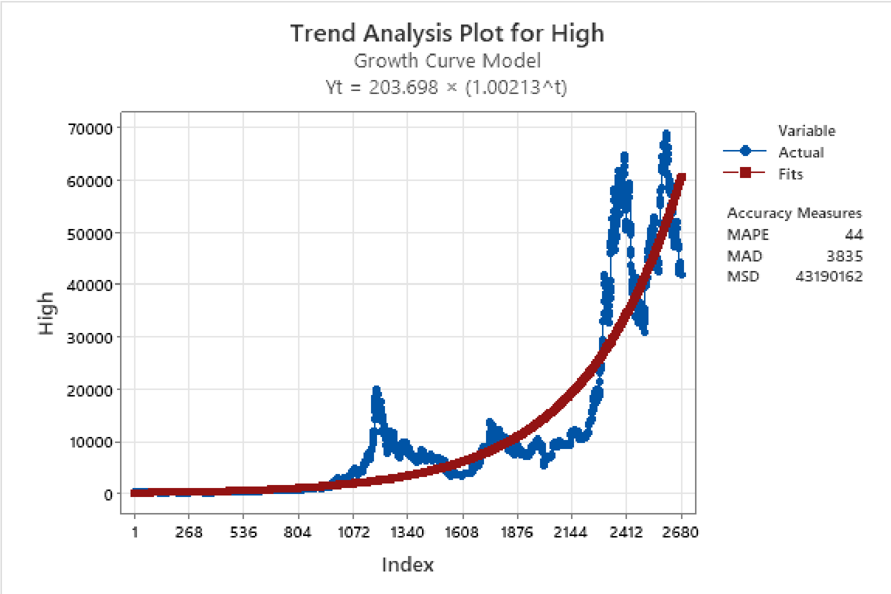
Accuracy Measures

MAPE	44
MAD	3835
MSD	43190162

And the fitted trend equation is:

Fitted Trend Equation

Yt = 203.698 × (1.00213^t)



Here we can see that the MAPE value is 44% which is moderate and the closest fit among others. Now we will fir exponential trend equation to all of the rest columns as they are similar type of data.

2. Fitting trend equation to “Low” :

Method

Model type	Growth Curve Model
Data	Low
Length	2683
NMissing	0

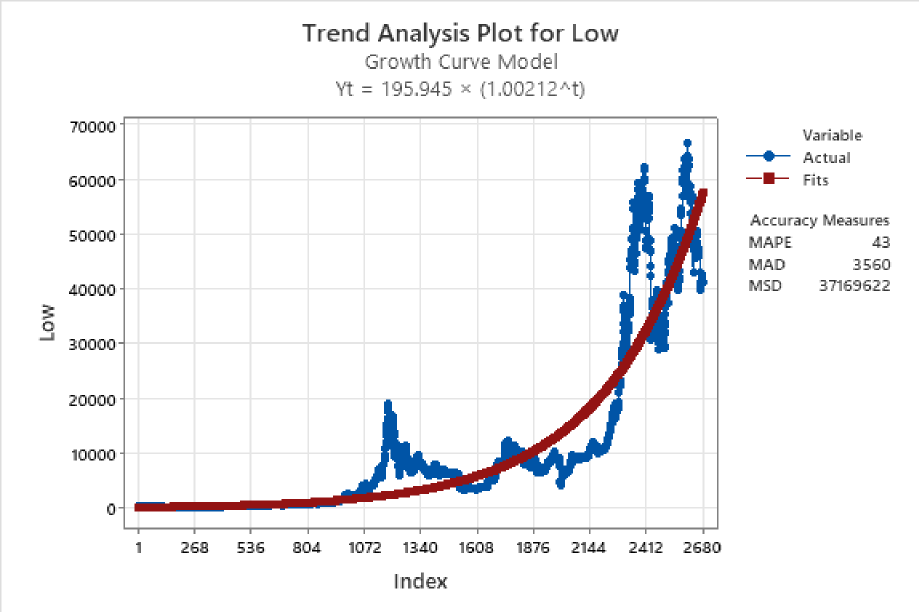
Accuracy Measures

MAPE	43
MAD	3560
MSD	37169622

And the fitted trend equation is:

Fitted Trend Equation

$$Yt = 195.945 \times (1.00212^{t})$$



3. Fitting trend equation to “Open” :

Method

Model type	Growth Curve Model
Data	Open
Length	2683
NMissing	0

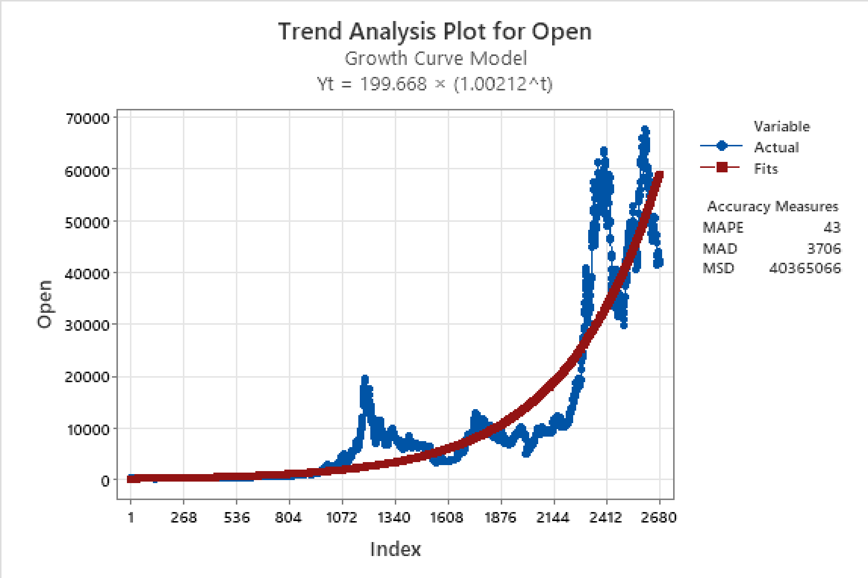
Accuracy Measures

MAPE	43
MAD	3706
MSD	40365066

And the fitted trend equation is:

Fitted Trend Equation

Yt = 199.668 × (1.00212^t)



4. Fitting trend equation to "Close" :

Method

Model type	Growth Curve Model
Data	Close
Length	2683
NMissing	0

Accuracy Measures

MAPE	43
MAD	3712
MSD	40476253

And the trend equation is:

Fitted Trend Equation

Yt = 199.879 × (1.00212^t)

From the above trend equations, we can observe that in the form a*(b^t), the values of 'b' are greater than 1, meaning the curve shows exponential growth resulting increase in Bitcoin's prices as 't' increases. But the values of MAPE are around the range

40-45 which is moderate and not accurate enough for forecasting. Also there are presence of trend and seasonality in the model which we have to remove in order to get more accurate results.

• Analysis :

1. Analysis for “High” :

To check the stationarity of the data stored in the column “High”, we are going to perform Augmented Dickey Fuller Test. In this test,

The null Hypothesis: The data is not stationary.

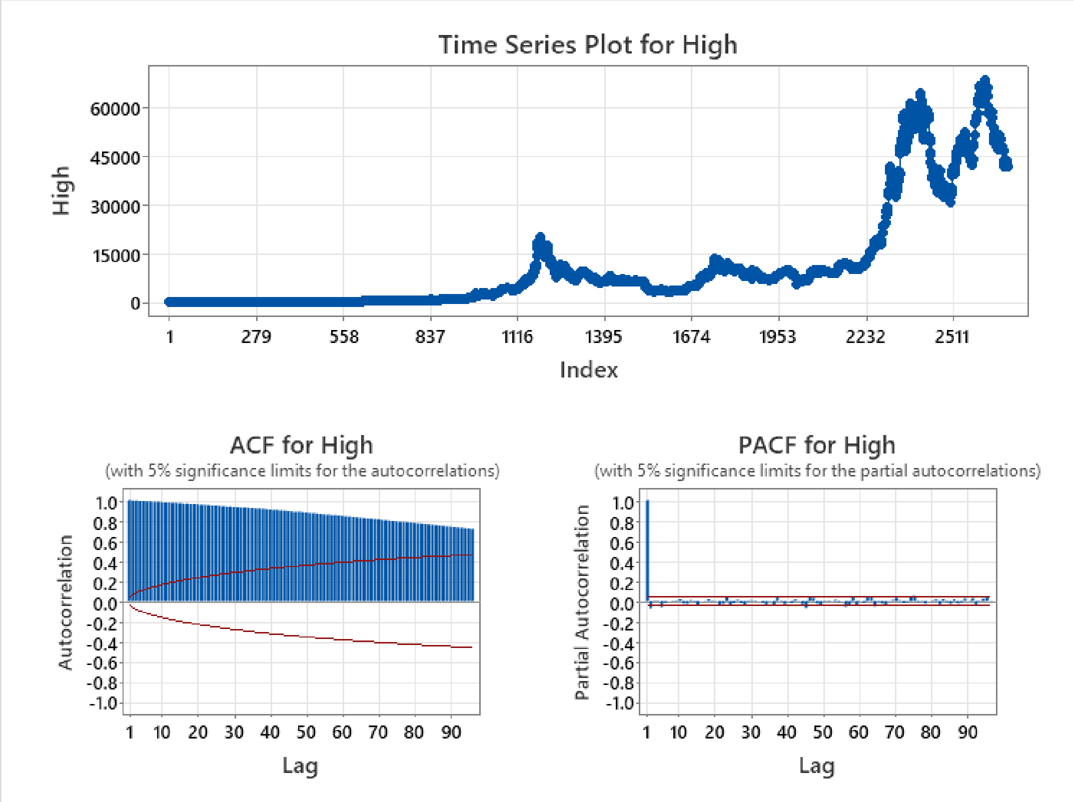
The alternative hypothesis: The data is stationary.

If the p-value of the test is greater than 0.05, then we accept the hypothesis, otherwise we reject it.

Augmented Dickey-Fuller Test

Null hypothesis: Data are non-stationary
Alternative hypothesis: Data are stationary

Test Statistic	P-Value	Recommendation
-0.998499	0.754	Test statistic > critical value of -2.86263. Significance level = 0.05 Fail to reject null hypothesis. Consider differencing to make data stationary.



As the p-value is 0.754>0.05, we accept the null hypothesis. The data is stationary. Now we have to log transform the series to make it stationary.

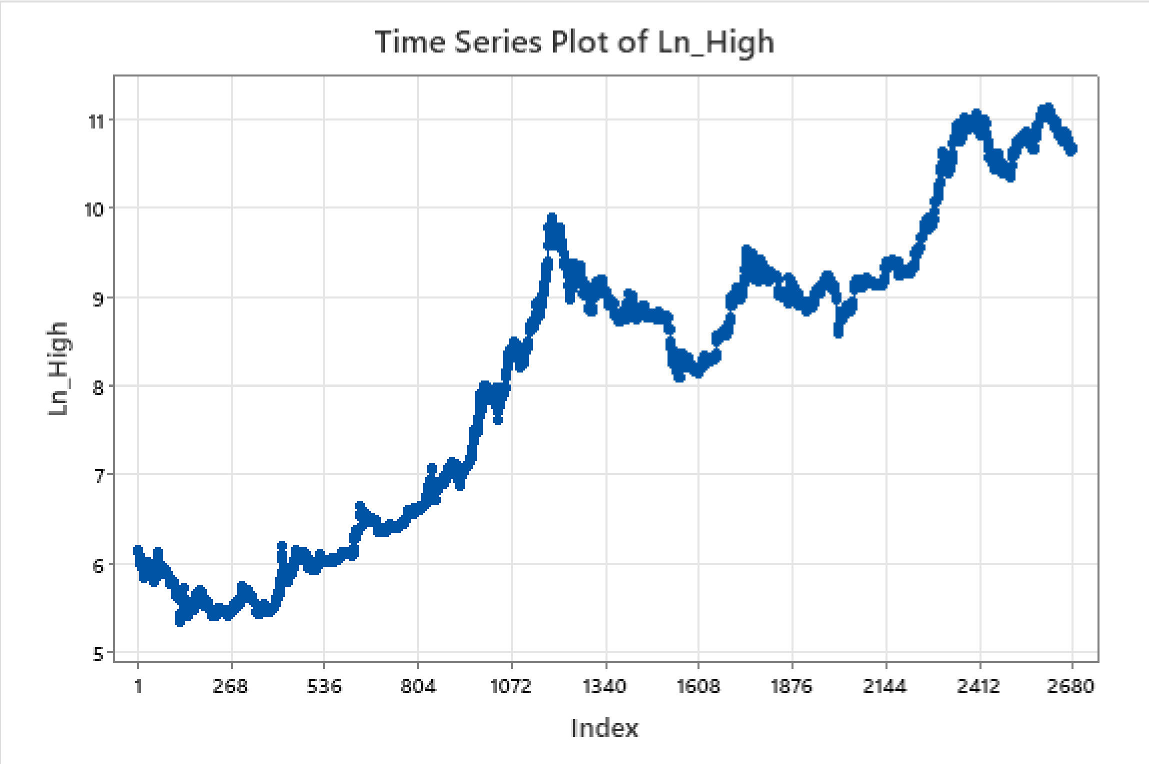
Log transforming the series :

Log transformation is used to unskew highly skewed data by stabilizing variance and eliminating the effect of extreme values.

The log transformation has several benefits in achieving stationarity. It can linearize exponential growth or decay pattern, making the data easier for modelling. It also reduces the magnitude of large fluctuations, making the series more consistent and suitable for analysing.

High
468.174
456.86
427.835
423.296
412.426
406.916
441.557
436.112
423.52
414.938
406.623
401.017
385.211
390.977
391.379

ln_High
6.14884
6.124377
6.058738
6.048072
6.022057
6.008607
6.090307
6.077899
6.048601
6.028129
6.007886
5.994004
5.953791
5.968649
5.969676

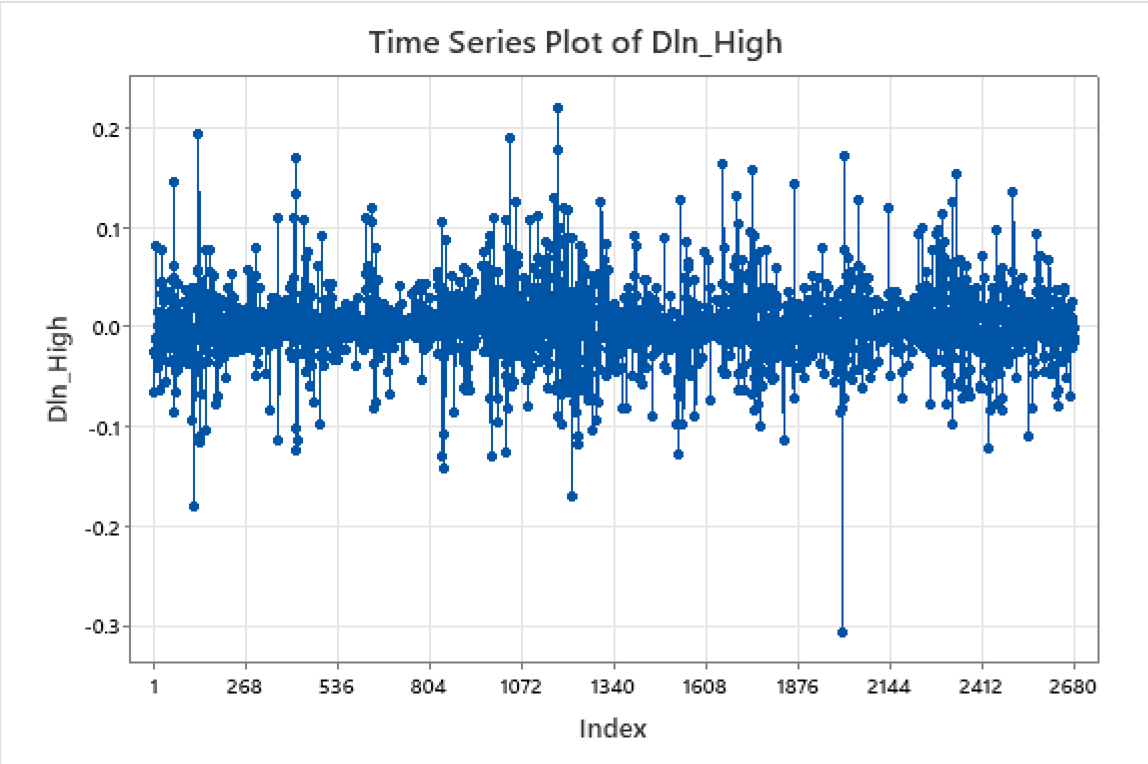


But there are still effects of trend and seasonality. To remove this effect, we will use differencing by calculating the difference between two consecutive values.

Differencing:

ln_High
6.14884
6.124377
6.058738
6.048072
6.022057
6.008607
6.090307
6.077899
6.048601
6.028129
6.007886
5.994004
5.953791
5.968649
5.969676

Dln_High
-0.02446
-0.06564
-0.01067
-0.02601
-0.01345
0.0817
-0.01241
-0.0293
-0.02047
-0.02024
-0.01388
-0.04021
0.014857
0.001028



Now we are going to do Augmented Dickey Fuller test to check the stationarity.

Augmented Dickey-Fuller Test

Null hypothesis: Data are non-stationary
Alternative hypothesis: Data are stationary

Test Statistic	P-Value	Recommendation
-24.3012	0.000	Test statistic <= critical value of -2.86262. Significance level = 0.05 Reject null hypothesis. Data appears to be stationary, not supporting differencing.

As we can see that the p-value of this test is 0.00<0.05, hence this data is now stationary and we can fit various models to forecast.

Model Fitting:

Now, we are going to fit different Autoregressive, Moving Average and ARIMA models to check which one has lowest error and more accuracy for prediction.

I. AR(1) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	0.1431	0.0191	7.49	0.000
Constant	0.001435	0.000657	2.18	0.029
Mean	0.001675	0.000767		

Residual Sums of Squares

DF	SS	MS
2680	3.10650	0.0011591

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	24.15	48.06	59.59	81.43
DF	10	22	34	46
P-Value	0.007	0.001	0.004	0.001

As we can see, the value of RSS in this model is 3.1065.

II. AR(2) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	0.1502	0.0193	7.78	0.000
AR 2	-0.0498	0.0193	-2.58	0.010
Constant	0.001508	0.000657	2.30	0.022
Mean	0.001676	0.000730		

Residual Sums of Squares

DF	SS	MS
2679	3.09881	0.0011567

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	17.82	42.94	54.42	76.32
DF	9	21	33	45
P-Value	0.037	0.003	0.011	0.002

For this model, the value of RSS is 3.09881.

III. MA(1) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
MA 1	-0.1540	0.0191	-8.07	0.000
Constant	0.001675	0.000758	2.21	0.027
Mean	0.001675	0.000758		

Residual Sums of Squares

DF	SS	MS
2680	3.10131	0.0011572

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	19.76	44.26	55.95	77.70
DF	10	22	34	46
P-Value	0.032	0.003	0.010	0.002

For this model, the value of RSS is 3.10131.

IV. MA(3):

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
MA 1	-0.1516	0.0193	-7.84	0.000
MA 2	0.0238	0.0195	1.22	0.223
MA 3	0.0195	0.0193	1.01	0.312
Constant	0.001677	0.000728	2.30	0.021
Mean	0.001677	0.000728		

Residual Sums of Squares

DF	SS	MS
2678	3.09851	0.0011570

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	17.55	42.70	54.07	76.08
DF	8	20	32	44
P-Value	0.025	0.002	0.009	0.002

In this Model, the value of RSS is 3.09851.

V. MA(5) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
MA 1	-0.1490	0.0193	-7.71	0.000
MA 2	0.0262	0.0195	1.34	0.180
MA 3	0.0154	0.0195	0.79	0.432
MA 4	-0.0444	0.0195	-2.28	0.023
MA 5	-0.0285	0.0193	-1.47	0.141
Constant	0.001672	0.000775	2.16	0.031
Mean	0.001672	0.000775		

Residual Sums of Squares

DF	SS	MS
2676	3.09125	0.0011552

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	11.60	37.64	48.85	72.14
DF	6	18	30	42
P-Value	0.071	0.004	0.016	0.003

For this model, we can observe, the value of RSS is 3.09125.

So from all of the above Autoregressive and Moving Average models, we can see the AR(2) has the lowest RSS(3.09881) among other AR models and MA(5) has the lowest RSS(3.09125) among other MA models. Hence, for the ARIMA model, we are going to take Autoregressive order 2, Difference order 1, and Moving Average Order 5.

VI. ARIMA(2,1,5) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	0.042	0.140	0.30	0.766
AR 2	0.823	0.151	5.44	0.000
MA 1	-0.108	0.142	-0.77	0.444
MA 2	0.857	0.144	5.94	0.000
MA 3	0.1354	0.0367	3.69	0.000
MA 4	-0.0555	0.0204	-2.72	0.007
MA 5	-0.0116	0.0240	-0.48	0.630
Constant	0.000223	0.000121	1.84	0.065

Residual Sums of Squares

DF	SS	MS
2674	3.08746	0.0011546

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	11.60	37.64	48.85	72.14
DF	6	18	30	42
P-Value	0.071	0.004	0.016	0.003

The value of RSS for this model is 3.08746.

Since, the value of RSS for this ARIMA(2,1,5) is the lowest(3.08746) among all of the previous Autoregressive, Moving Average and ARIMA models, this is the most accurate model to predict future Bitcoin prices.

Forecasting:

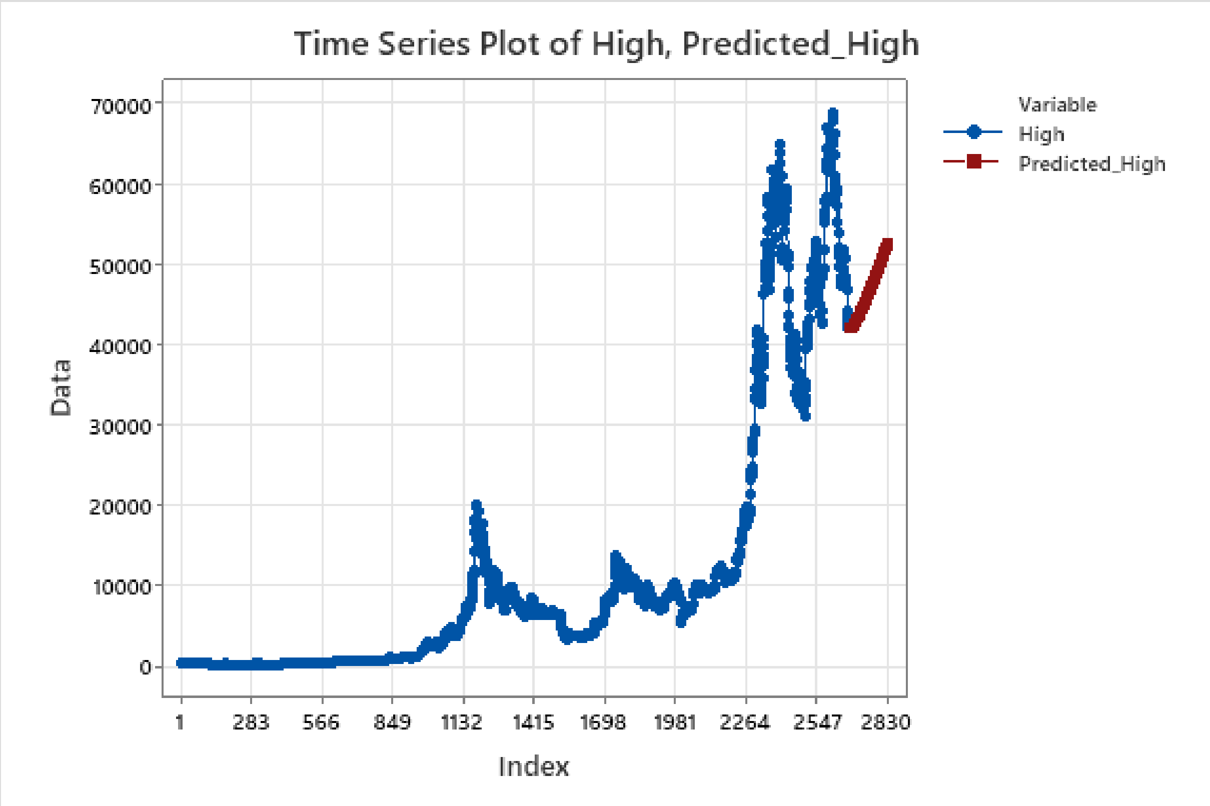
Using the forecast tool provided in Minitab, we are going to predict future values of "ln_High" and then taking exponential of the values, we are going to get respective future Bitcoin prices.

There are all total 150 predictions, we take exponential of the values to get actual Bitcoin "High" prices. The first 25 of the forecasts and their exponential values are:

Forecast
10.6462
10.646
10.6474
10.6469
10.6481
10.648
10.6492
10.6494
10.6507
10.6511
10.6523
10.653
10.6543
10.6551
10.6564
10.6573
10.6586
10.6597
10.6611
10.6622
10.6636
10.6648
10.6662
10.6675
10.669
10.6703

predicted_High
42032.56693
42024.16126
42083.03629
42062.00003
42112.50473
42108.29369
42158.85397
42167.28659
42222.1397
42239.03194
42289.7492
42319.36239
42374.41334
42408.32643
42463.4931
42501.72745
42557.01563
42603.8541
42663.54127
42710.49698
42770.33356
42821.68876
42881.68111
42937.46355
43001.91807
43057.85692

Plotting the predicted Bitcoin “High” Values with the actual “High” values:



As we can see from the Analysis, that ARIMA(2,1,5) is the most accurate model to predict future Bitcoin “High” prices among all other AR, MA and ARIMA model and from the graph, it is clear that there is going to be an increase in the price of Bitocoin meaning it is a good opportunity for Investors to invest and make profit over time.

2. Analysis for "Low" :

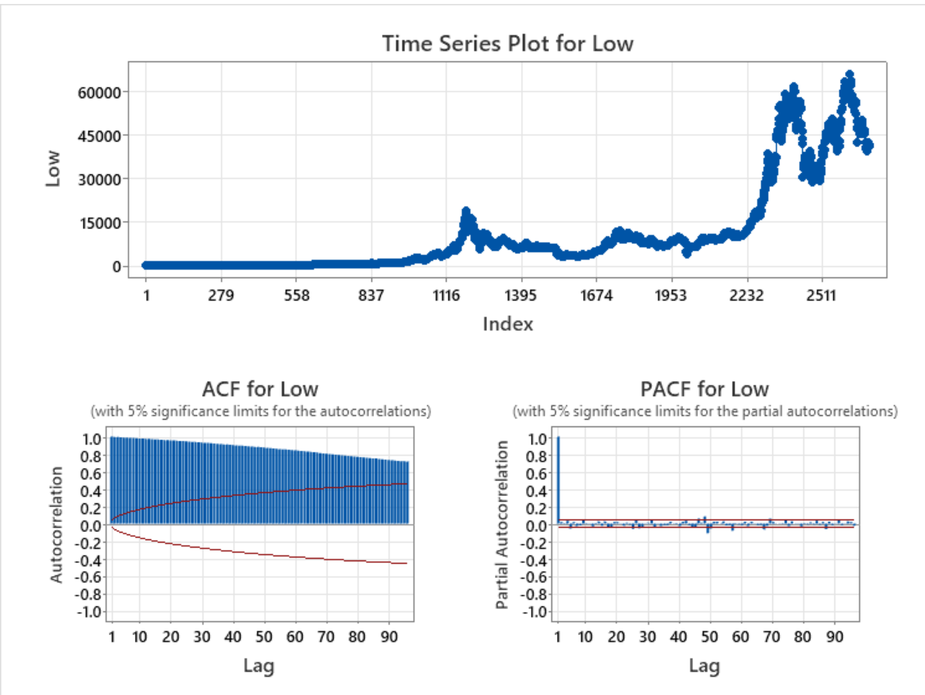
We have to perform Augmented Dickey-Fuller Test to check if the data is stationary or not.

Augmented Dickey-Fuller Test

Null hypothesis: Data are non-stationary

Alternative hypothesis: Data are stationary

Test Statistic	P-Value	Recommendation
-0.902223	0.787	Test statistic > critical value of -2.86263. Significance level = 0.05 Fail to reject null hypothesis. Consider differencing to make data stationary.



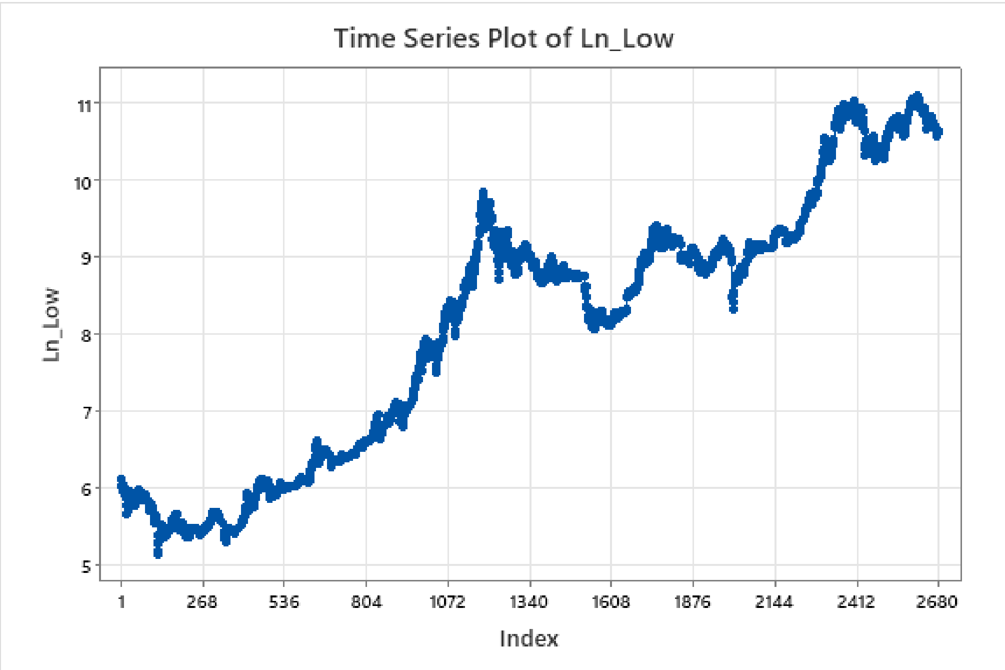
From the test, it is evident that the data is not stationary as the p-value > 0.05.

Log transforming the series:

We have to Log transform the data in an attempt to make it stationary.

Low
452.422
413.104
384.532
389.883
393.181
397.13
396.197
421.132
409.468
400.009
397.372
374.332
372.24
373.443
380.78

ln_Low
6.114615
6.023699
5.952027
5.965847
5.97427
5.984264
5.981912
6.042946
6.014859
5.991487
5.984873
5.925143
5.919539
5.922765
5.942222

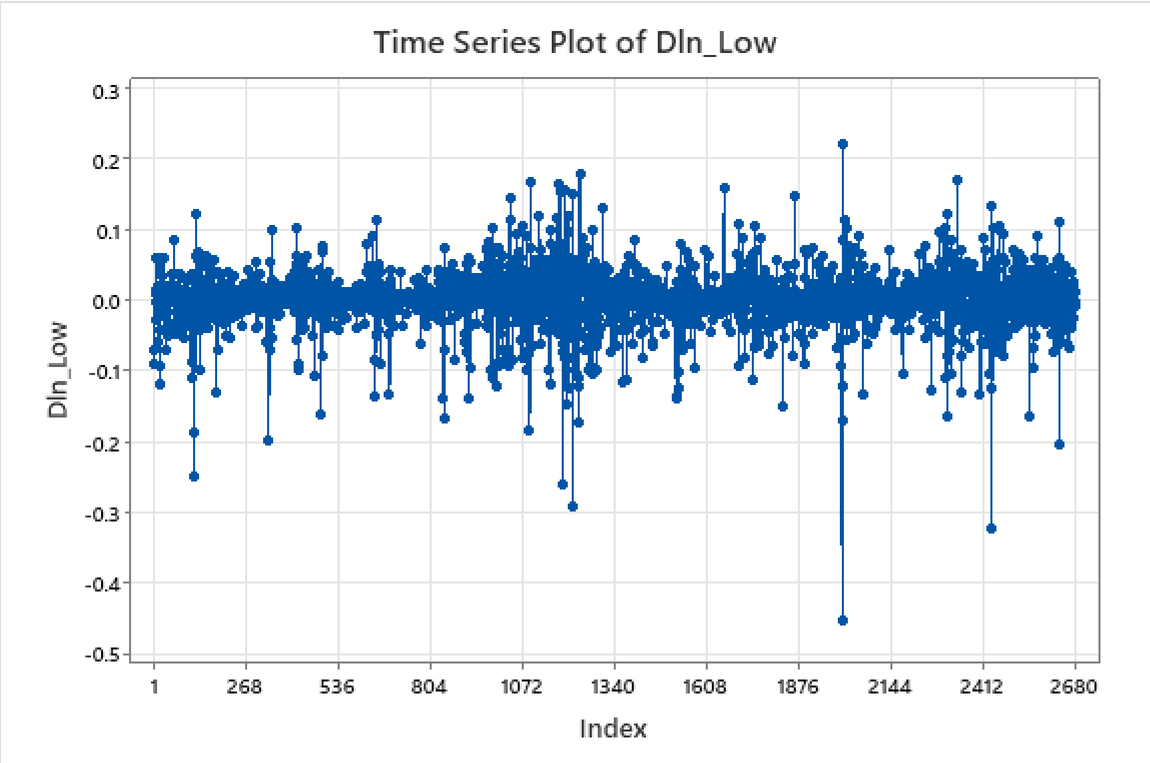


Now, we have to difference the data in order to remove any trend or seasonality.

Differencing:

ln_Low
6.114615
6.023699
5.952027
5.965847
5.97427
5.984264
5.981912
6.042946
6.014859
5.991487
5.984873
5.925143
5.919539
5.922765
5.942222

Dln_Low
-0.09092
-0.07167
0.01382
0.008423
0.009994
-0.00235
0.061035
-0.02809
-0.02337
-0.00661
-0.05973
-0.0056
0.003227
0.019456

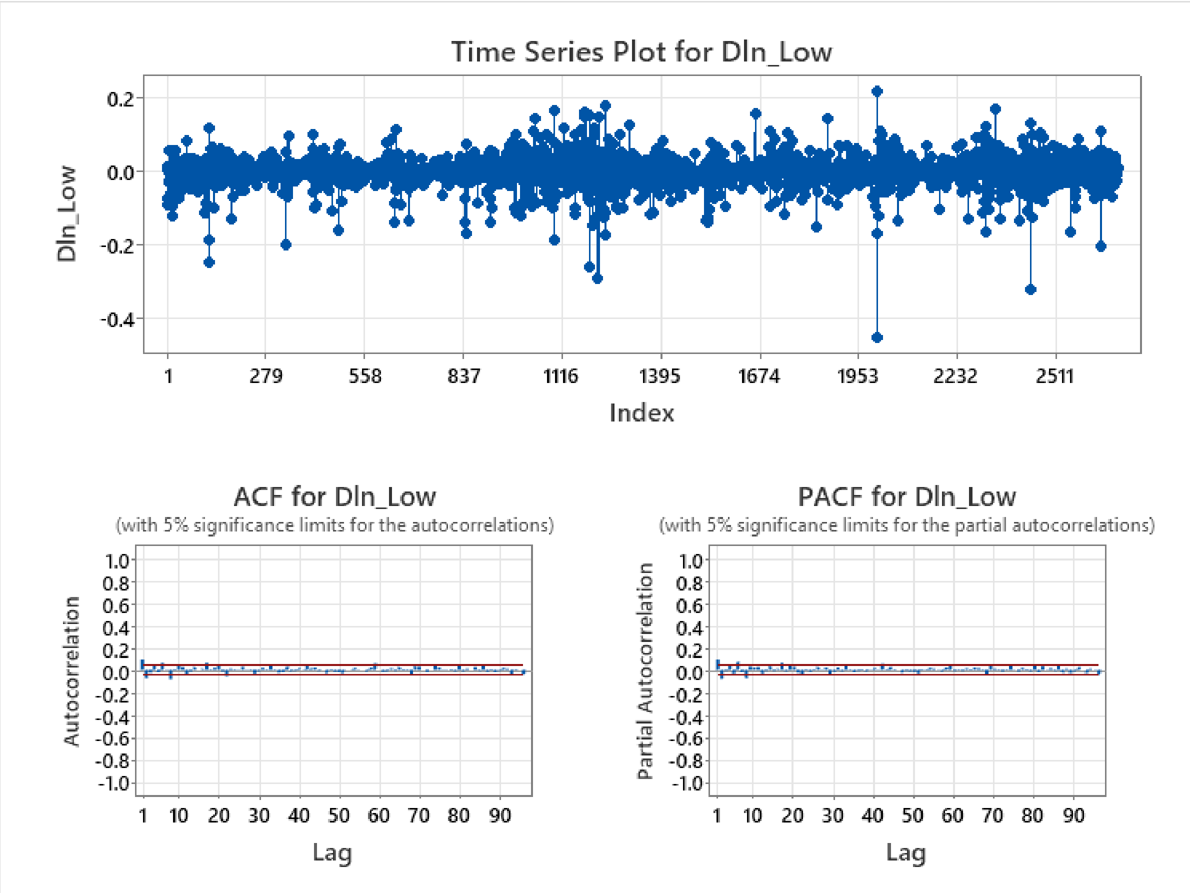


Now, we will perform Augmented Dickey Fuller test again to check the stationarity.

Augmented Dickey-Fuller Test

Null hypothesis: Data are non-stationary
Alternative hypothesis: Data are stationary

Test Statistic	P-Value	Recommendation
-10.4775	0.000	Test statistic <= critical value of -2.86263. Significance level = 0.05 Reject null hypothesis. Data appears to be stationary, not supporting differencing.



From the test, it is evident that that the data is now stationary, so now we can fit different models for prediction.

Model Fitting:

I. AR(1) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	0.0919	0.0192	4.78	0.000
Constant	0.001529	0.000757	2.02	0.043
Mean	0.001684	0.000834		

Residual Sums of Squares

DF	SS	MS
2680	4.11888	0.0015369

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	68.52	98.27	112.58	127.62
DF	10	22	34	46
P-Value	0.000	0.000	0.000	0.000

The value or RSS for this mode is 4.1188.

II. AR(2) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	0.0995	0.0193	5.17	0.000
AR 2	-0.0832	0.0193	-4.32	0.000
Constant	0.001661	0.000755	2.20	0.028
Mean	0.001689	0.000767		

Residual Sums of Squares

DF	SS	MS
2679	4.09048	0.0015269

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	44.11	72.72	86.56	103.75
DF	9	21	33	45
P-Value	0.000	0.000	0.000	0.000

The value of RSS is 4.09048.

III. MA(1) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
MA 1	-0.1078	0.0192	-5.62	0.000
Constant	0.001684	0.000838	2.01	0.045
Mean	0.001684	0.000838		

Residual Sums of Squares

DF	SS	MS
2680	4.11261	0.0015346

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	64.03	93.75	108.12	123.61
DF	10	22	34	46
P-Value	0.000	0.000	0.000	0.000

The value of RSS is 4.11261.

IV. MA(3) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
MA 1	-0.1020	0.0193	-5.28	0.000
MA 2	0.0675	0.0194	3.48	0.001
MA 3	0.0362	0.0193	1.88	0.061
Constant	0.001690	0.000753	2.24	0.025
Mean	0.001690	0.000753		

Residual Sums of Squares

DF	SS	MS
2678	4.08968	0.0015271

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	44.32	74.01	87.82	104.75
DF	8	20	32	44
P-Value	0.000	0.000	0.000	0.000

The value of RSS for this model is 4.08968.

V. MA(5) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
MA 1	-0.1058	0.0193	-5.48	0.000
MA 2	0.0751	0.0194	3.87	0.000
MA 3	0.0355	0.0194	1.83	0.067
MA 4	-0.0649	0.0194	-3.35	0.001
MA 5	0.0355	0.0193	1.84	0.066
Constant	0.001688	0.000771	2.19	0.029
Mean	0.001688	0.000771		

Residual Sums of Squares

DF	SS	MS
2676	4.06839	0.0015203

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	29.55	59.23	75.32	91.99
DF	6	18	30	42
P-Value	0.000	0.000	0.000	0.000

The value of RSS for this model is 4.06839.

So from all of the above Autoregressive and Moving Average models, we can see the AR(2) has the lowest RSS(4.09048) among other AR models and MA(5) has the lowest RSS(4.06389) among other MA models. Hence, for the ARIMA model, we are going to take Autoregressive order 2, Difference order 1, and Moving Average Order 5.

VI. ARIMA(2,1,5) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	-1.2473	0.0295	-42.31	0.000
AR 2	-0.9144	0.0223	-41.03	0.000
MA 1	-1.3579	0.0178	-76.10	0.000
MA 2	-0.97940	0.00480	-204.25	0.000
MA 3	0.0301	0.0281	1.07	0.285
MA 4	0.0863	0.0333	2.59	0.010
MA 5	0.0364	0.0204	1.79	0.074
Constant	0.00534	0.00238	2.25	0.025

Residual Sums of Squares

DF	SS	MS
2674	4.04153	0.0015114

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	30.49	51.54	62.30	76.92
DF	4	16	28	40
P-Value	0.000	0.000	0.000	0.000

The value of RSS is 4.04153.

Since, the value of RSS for this ARIMA(2,1,5) is the lowest(3.08746) among all of the previous Autoregressive, Moving Average and ARIMA models, this is the most accurate model to predict future Bitcoin prices.

Forecasting :

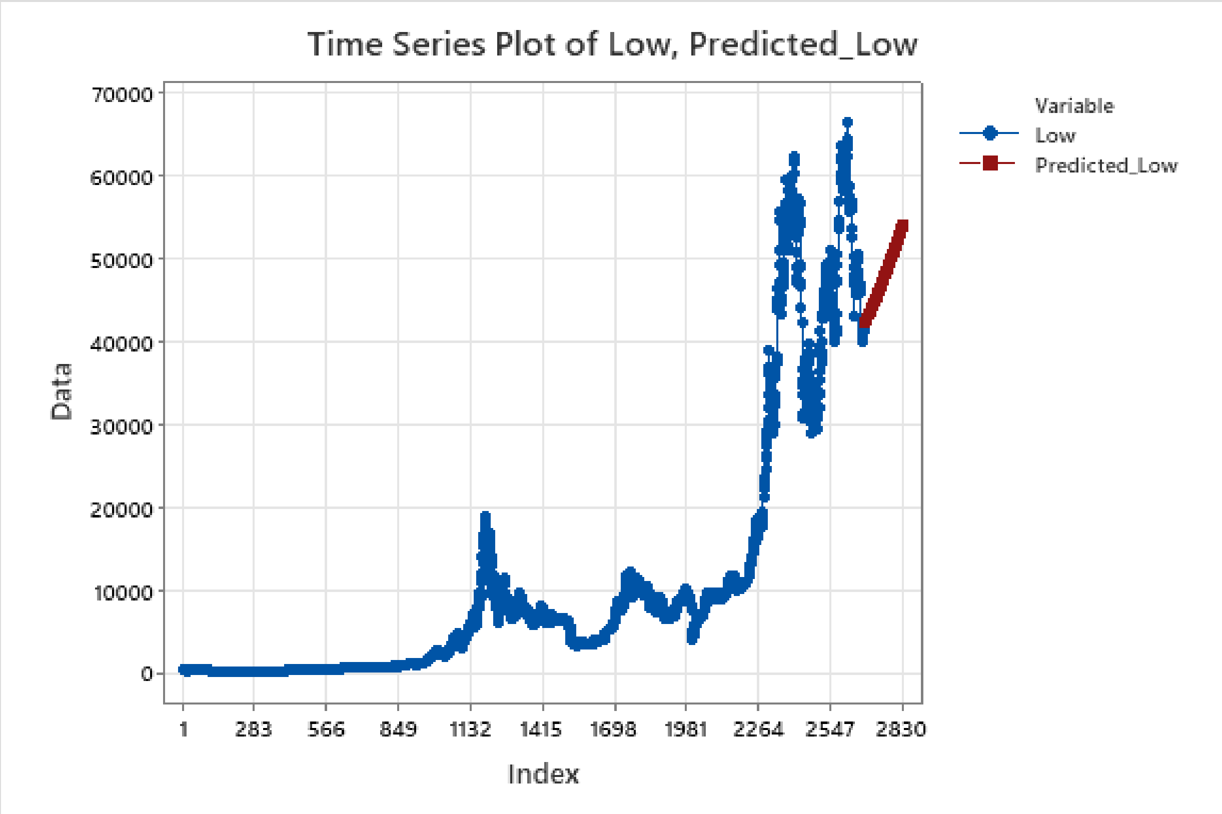
Using the forecast tool provided in Minitab, we are going to predict future values of "ln_Low" and then taking exponential of the values, we are going to get respective future Bitcoin prices.

There are all total 150 predictions, we take exponential of the values to get actual Bitcoin "Low" prices. The first 25 of the forecasts and their exponential values are:

Forecast
10.6501
10.6481
10.6462
10.6546
10.6506
10.6533
10.659
10.6548
10.6602
10.6626
10.66
10.6664
10.6662
10.6659
10.6718
10.67
10.6722
10.6764
10.6745
10.6784
10.6807
10.6796
10.6842
10.6848
10.6852

Predicted_Low
42196.81402
42112.50473
42032.56693
42387.12757
42217.9177
42332.0601
42574.04184
42395.60584
42625.16135
42727.5846
42616.63717
42890.25831
42881.68111
42868.81854
43122.49217
43044.9415
43139.74461
43321.31257
43239.08022
43408.04189
43507.99529
43460.16281
43660.54007
43686.74426
43704.22245

Plotting the predicted Bitcoin “Low” Values with the actual “Low” values:



As we can see from the Analysis, that ARIMA(2,1,5) is the most accurate model to predict future Bitcoin “Low” prices among all other AR, MA and ARIMA model and from the graph, it is clear that there is going to be an increase in the price of Bitocoin meaning it is a good opportunity for Investors to invest and make profit over time.

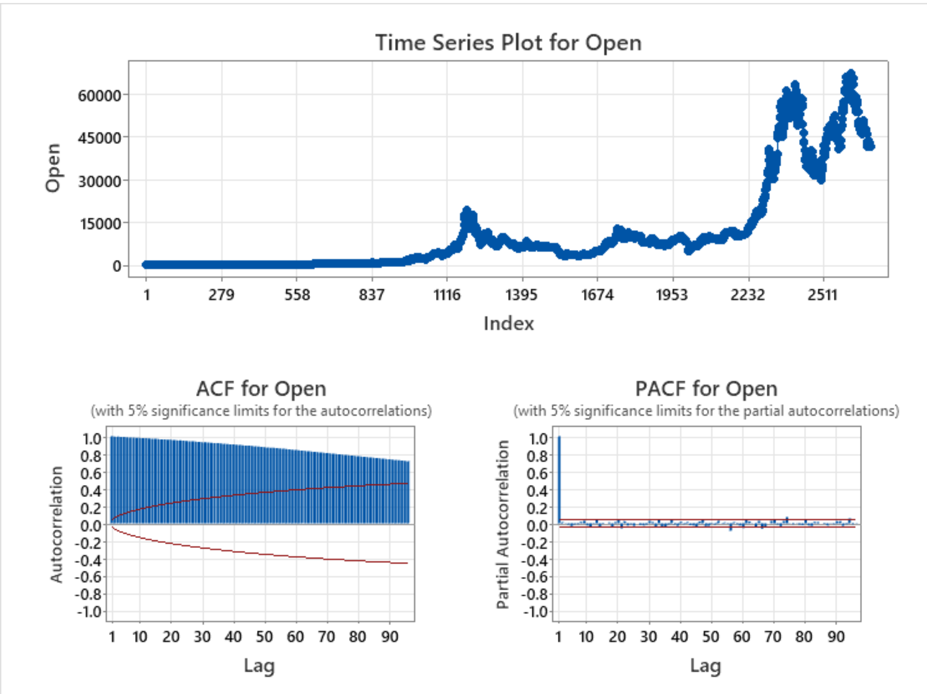
3. Analysis for “Open” :

We have to perform Augmented Dickey Fuller Test to check if the data is stationary or not.

Augmented Dickey-Fuller Test

Null hypothesis: Data are non-stationary
Alternative hypothesis: Data are stationary

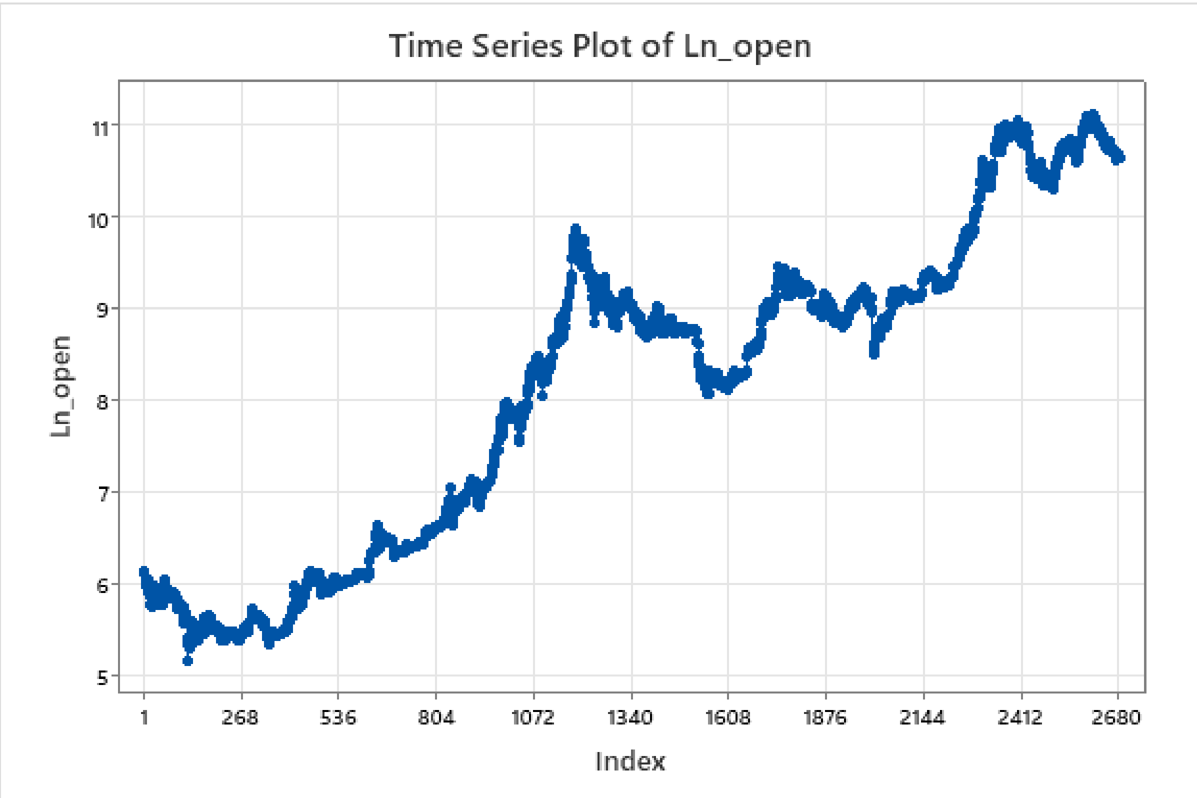
Test Statistic	P-Value	Recommendation
-0.929768	0.778	Test statistic > critical value of -2.86263. Significance level = 0.05 Fail to reject null hypothesis. Consider differencing to make data stationary.



As we can see that p-value > 0.05, we have to conclude that the data is not stationary and we have to perform Log transformation and Differencing to make it stationary.

Log Transforming the series:

Open	Ln_Open
465.864	6.143894
456.86	6.124377
424.103	6.049976
394.673	5.978058
408.085	6.011475
399.1	5.989212
402.092	5.996681
435.751	6.077071
423.156	6.047741
411.429	6.019636
403.556	6.000315
399.471	5.990141
376.928	5.932054
376.088	5.929823
387.427	5.959527

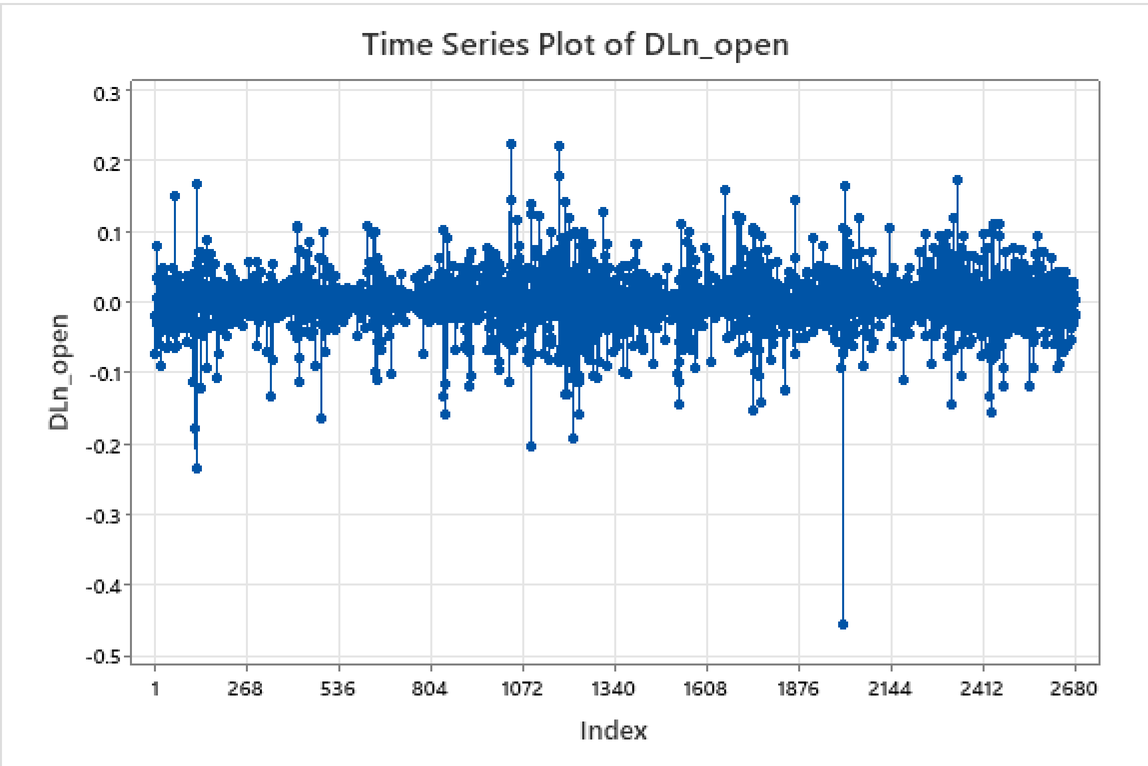


Now, we will perform differencing to remove any trend and seasonality.

Differencing:

Ln_Open
6.143894
6.124377
6.049976
5.978058
6.011475
5.989212
5.996681
6.077071
6.047741
6.019636
6.000315
5.990141
5.932054
5.929823
5.959527

DLn_Open
-0.0195168
-0.0744006
-0.0719188
0.03341788
-0.0222634
0.00746892
0.08039005
-0.0293301
-0.0281045
-0.0193212
-0.0101741
-0.058087
-0.002231
0.02970426

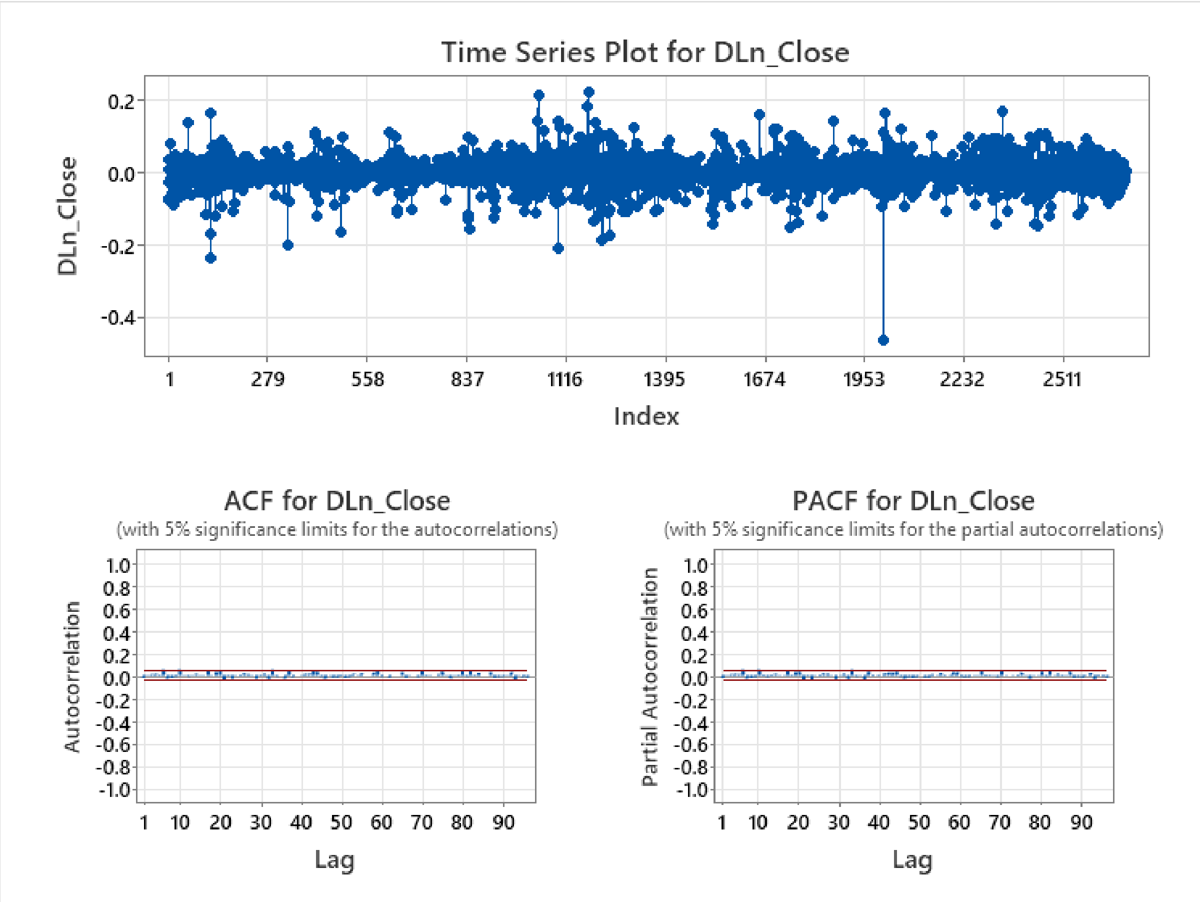


We need to test the stationarity of this data by Augmented Dickey Fuller Test.

Augmented Dickey-Fuller Test

Null hypothesis: Data are non-stationary
Alternative hypothesis: Data are stationary

Test Statistic	P-Value	Recommendation
-15.5423	0.000	Test statistic <= critical value of -2.86262. Significance level = 0.05 Reject null hypothesis. Data appears to be stationary, not supporting differencing.



As we can see that the p-value of this test is 0.00<0.05, hence this data is now stationary and we can fit various models to forecast.

Model Fitting :

I. AR(1) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	-0.0200	0.0193	-1.03	0.301
Constant	0.001710	0.000753	2.27	0.023
Mean	0.001676	0.000738		

Residual Sums of Squares

DF	SS	MS
2680	4.07314	0.0015198

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	18.95	37.70	53.90	71.48
DF	10	22	34	46
P-Value	0.041	0.020	0.016	0.009

For this model, the value of RSS is 4.07314.

II. AR(2) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	-0.0198	0.0193	-1.02	0.306
AR 2	0.0094	0.0193	0.49	0.625
Constant	0.001693	0.000753	2.25	0.025
Mean	0.001676	0.000745		

Residual Sums of Squares

DF	SS	MS
2679	4.07278	0.0015203

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	18.93	37.55	53.71	71.26
DF	9	21	33	45
P-Value	0.026	0.015	0.013	0.008

The value of RSS for this model is 4.07278.

III. MA(1) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
MA 1	0.0196	0.0193	1.01	0.311
Constant	0.001676	0.000738	2.27	0.023
Mean	0.001676	0.000738		

Residual Sums of Squares

DF	SS	MS
2680	4.07318	0.0015198

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	18.96	37.71	53.93	71.50
DF	10	22	34	46
P-Value	0.041	0.020	0.016	0.009

The value of RSS for this model is 4.07318.

IV. MA(3) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
MA 1	0.0200	0.0193	1.04	0.300
MA 2	-0.0095	0.0193	-0.49	0.623
MA 3	-0.0074	0.0193	-0.38	0.701
Constant	0.001676	0.000751	2.23	0.026
Mean	0.001676	0.000751		

Residual Sums of Squares

DF	SS	MS
2678	4.07252	0.0015207

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	18.95	37.48	53.72	71.35
DF	8	20	32	44
P-Value	0.015	0.010	0.009	0.006

The value of RSS is 4.07252.

V. MA(5) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
MA 1	0.0216	0.0193	1.12	0.263
MA 2	-0.0095	0.0193	-0.49	0.622
MA 3	-0.0087	0.0193	-0.45	0.653
MA 4	-0.0133	0.0193	-0.69	0.493
MA 5	-0.0163	0.0193	-0.84	0.400
Constant	0.001673	0.000773	2.17	0.030
Mean	0.001673	0.000773		

Residual Sums of Squares

DF	SS	MS
2676	4.07065	0.0015212

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	19.53	37.49	53.94	70.34
DF	6	18	30	42
P-Value	0.003	0.005	0.005	0.004

For this model, the RSS value is 4.07065.

So from all of the above Autoregressive and Moving Average models, we can see the AR(2) has the lowest RSS(4.07278) among other AR models and MA(5) has the lowest RSS(4.07065) among other MA models. Hence, for the ARIMA model, we are going to take Autoregressive order 2, Difference order 1, and Moving Average Order 5.

VI. ARIMA(2,1,5) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	-1.5677	0.0382	-41.04	0.000
AR 2	-0.718	0.165	-4.35	0.000
MA 1	-1.5477	0.0362	-42.73	0.000
MA 2	-0.694	0.161	-4.31	0.000
MA 3	-0.0065	0.0381	-0.17	0.864
MA 4	-0.0167	0.0362	-0.46	0.644
MA 5	0.0102	0.0220	0.46	0.644
Constant	0.00551	0.00244	2.26	0.024

Residual Sums of Squares

DF	SS	MS
2674	4.06402	0.0015198

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	15.50	33.03	49.70	67.19
DF	4	16	28	40
P-Value	0.004	0.007	0.007	0.005

The RSS for this model is 4.06402.

Since, the value of RSS for this ARIMA(2,1,5) is the lowest(4.06402) among all of the previous Autoregressive, Moving Average and ARIMA models, this is the most accurate model to predict future Bitcoin prices.

Forecasting :

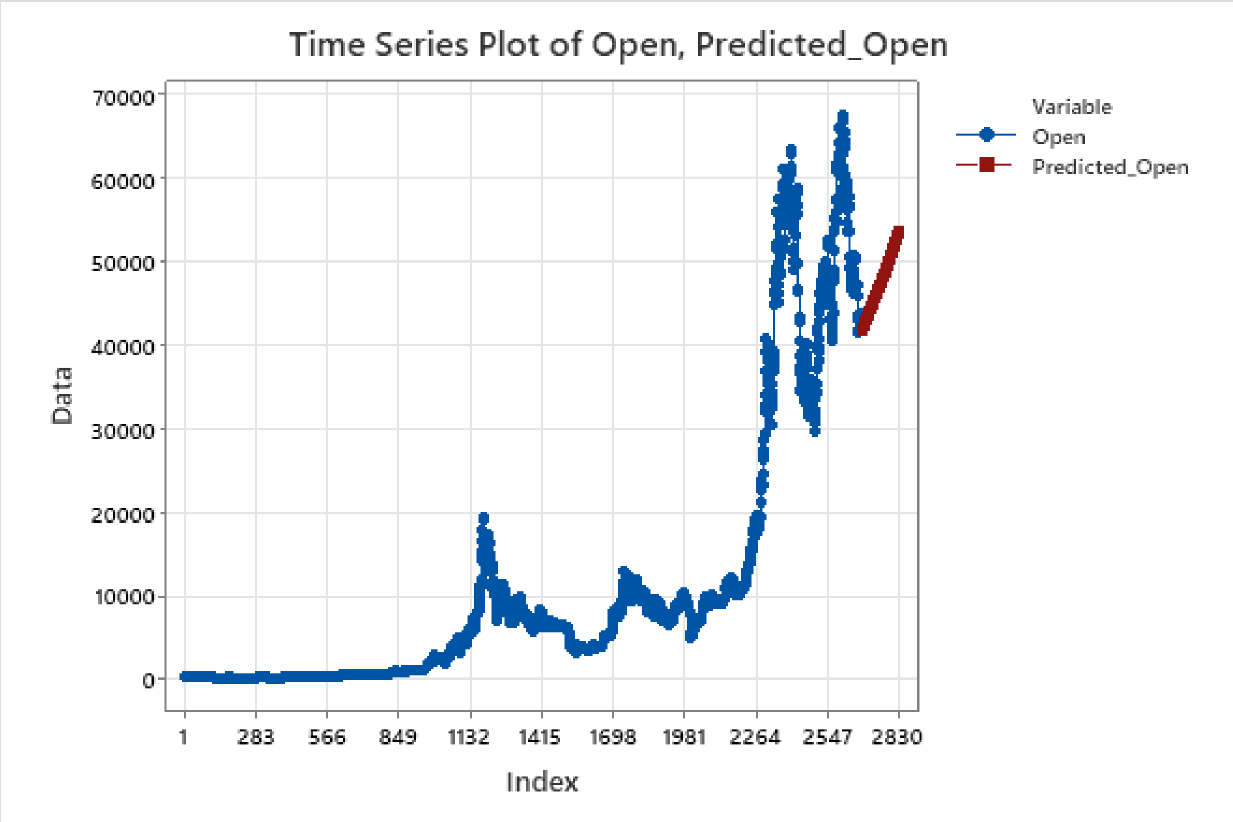
Using the forecast tool provided in Minitab, we are going to predict future values of “ln_Open” and then taking exponential of the values, we are going to get respective future Bitcoin prices.

There are all total 150 predictions, we take exponential of the values to get actual Bitcoin “Open” prices. The first 25 of the forecasts and their exponential values are:

Forecast
10.6417
10.6428
10.6448
10.6461
10.6483
10.6494
10.6516
10.6529
10.6548
10.6564
10.658
10.6598
10.6613
10.6632
10.6647
10.6665
10.6681
10.6698
10.6715
10.6731
10.6748
10.6765

predicted_Open
41843.84533
41889.89888
41973.76251
42028.36389
42120.92807
42167.28659
42260.15674
42315.13066
42395.60584
42463.4931
42531.48908
42608.1147
42672.07483
42753.22884
42817.40681
42894.54755
42963.23376
43036.33337
43109.55736
43178.58786
43252.05389
43325.64492

Plotting the predicted Bitcoin “Open” Values with the actual “Open” values:



As we can see from the Analysis, that ARIMA(2,1,5) is the most accurate model to predict future Bitcoin “Open” prices among all other AR, MA and ARIMA model and from the graph, it is clear that there is going to be an increase in the price of Bitocoin meaning it is a good opportunity for Investors to invest and make profit over time.

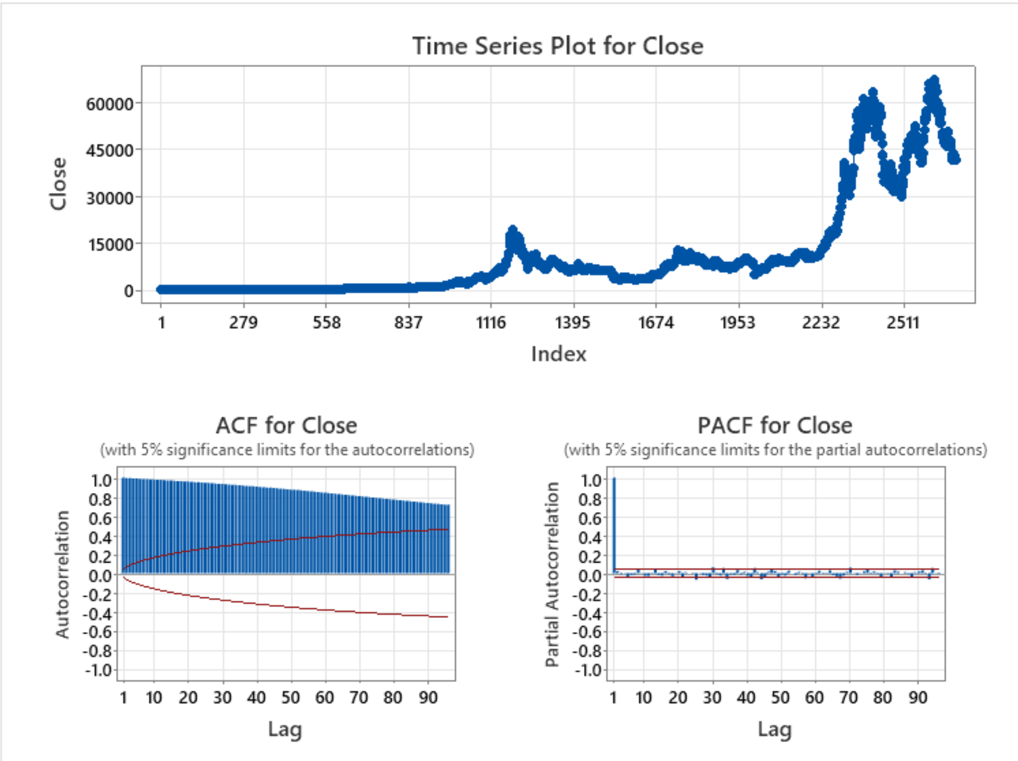
4. Analysis for "Close" :

We have to perform Augmented Dickey-Fuller Test to check if the data is stationary or not.

Augmented Dickey-Fuller Test

Null hypothesis: Data are non-stationary
Alternative hypothesis: Data are stationary

Test Statistic	P-Value	Recommendation
-0.920124	0.781	Test statistic > critical value of -2.86263. Significance level = 0.05 Fail to reject null hypothesis. Consider differencing to make data stationary.

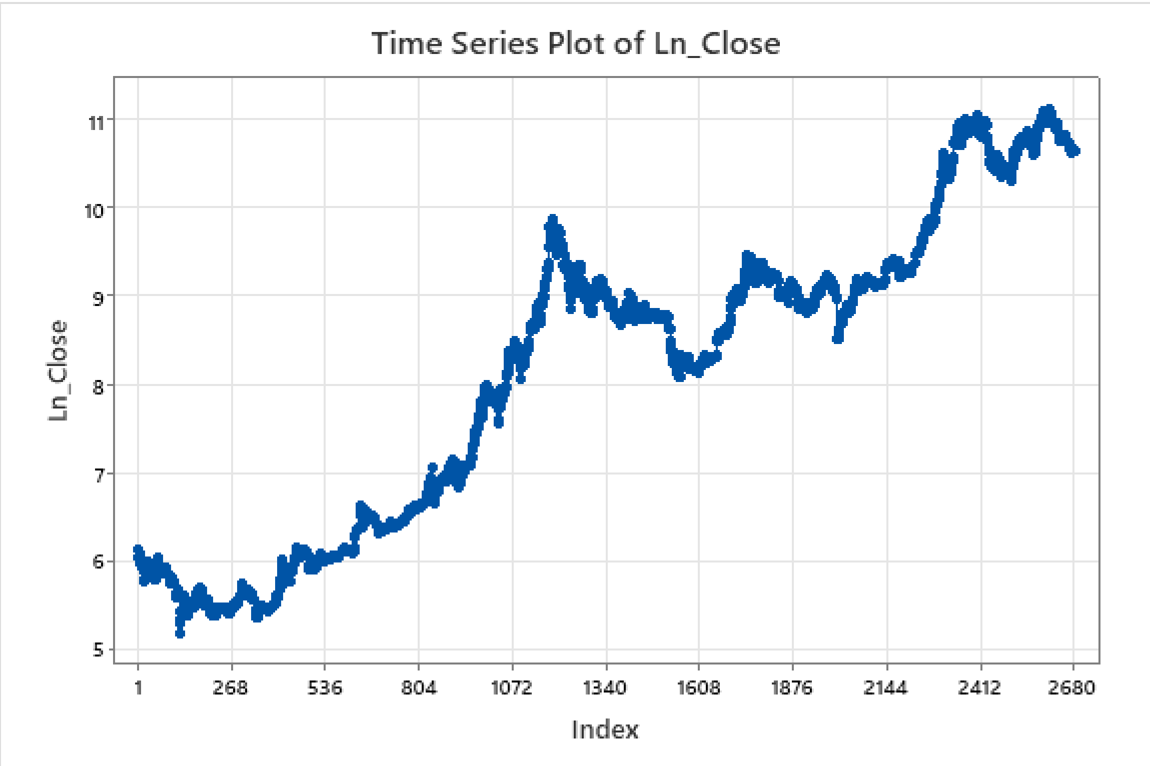


As we can see that p-value > 0.05, we have to conclude that the data is not stationary and we have to perform Log transformation and Differencing to make it stationary.

Log transforming the series :

Close
457.334
424.44
394.796
408.904
398.821
402.152
435.791
423.205
411.574
404.425
399.52
377.181
375.467
386.944
383.615
375.072

Ln_Close
6.125414
6.050771
5.978369
6.01348
5.988513
5.99683
6.077163
6.047857
6.019989
6.002466
5.990264
5.932725
5.928171
5.95828
5.949639
5.927118

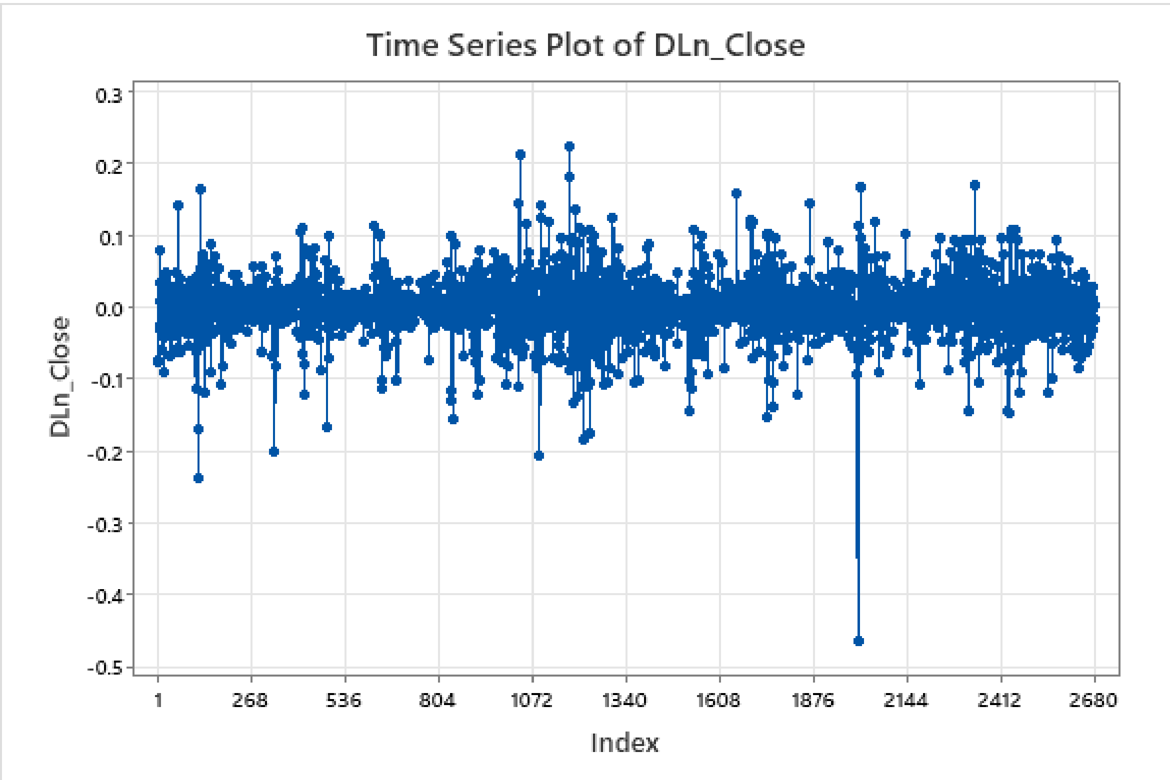


Now, we will use differencing to remove any trend or seasonality.

Differencing :

Ln_Close
6.125414
6.050771
5.978369
6.01348
5.988513
5.99683
6.077163
6.047857
6.019989
6.002466
5.990264
5.932725
5.928171
5.95828
5.949639
5.927118

DLn_Close
-0.074643352
-0.072401507
0.03511124
-0.02496766
0.008317417
0.08033259
-0.029306072
-0.027867819
-0.01752257
-0.012202477
-0.057538621
-0.004554567
0.030109369
-0.00864056
-0.022521436

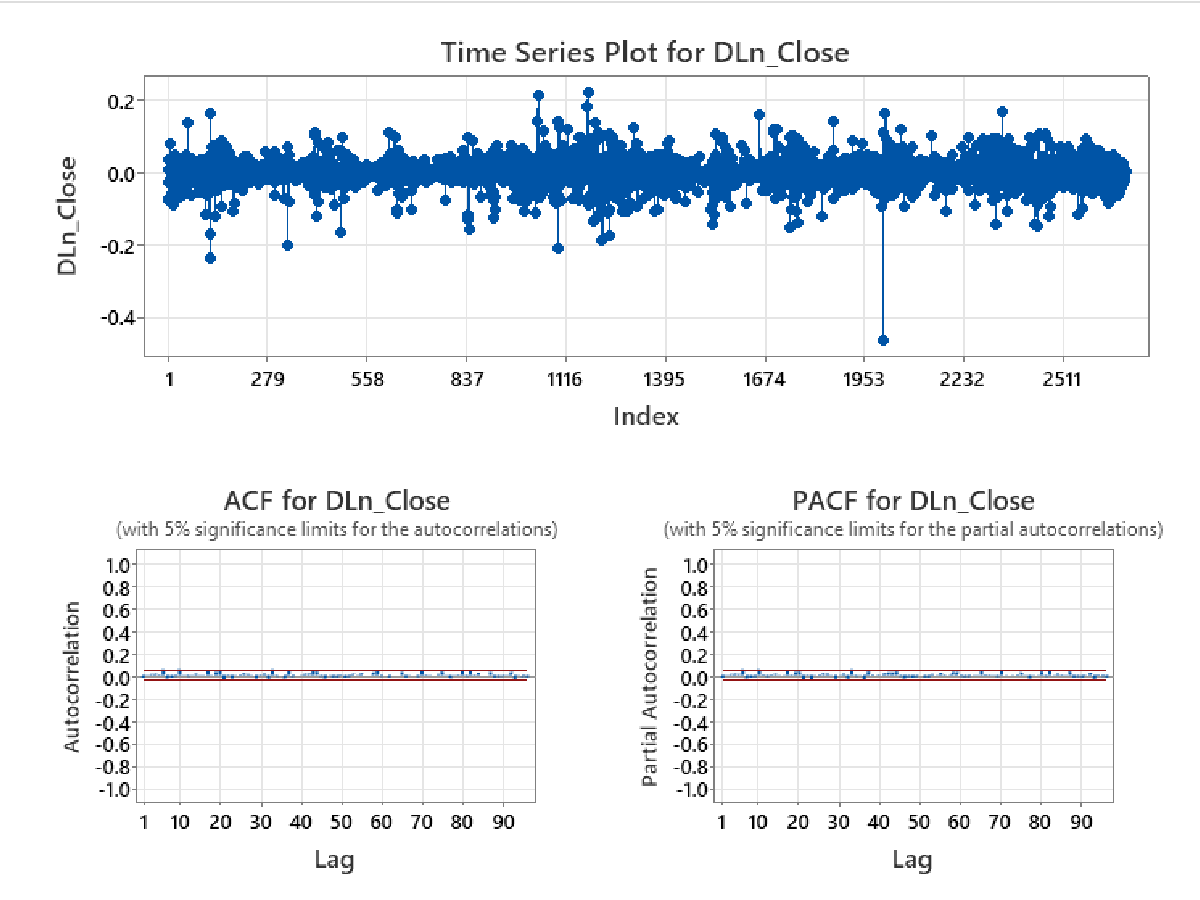


We need to test the stationarity of this data by Augmented Dickey Fuller Test.

Augmented Dickey-Fuller Test

Null hypothesis: Data are non-stationary
Alternative hypothesis: Data are stationary

Test Statistic	P-Value	Recommendation
-15.5423	0.000	Test statistic <= critical value of -2.86262. Significance level = 0.05 Reject null hypothesis. Data appears to be stationary, not supporting differencing.



As we can see that the p-value of this test is 0.00<0.05, hence this data is now stationary and we can fit various models to forecast.

Model Fitting :

I. AR(1):

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	-0.0217	0.0193	-1.12	0.261
Constant	0.001722	0.000756	2.28	0.023
Mean	0.001685	0.000740		

Residual Sums of Squares

DF	SS	MS
2680	4.11277	0.0015346

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	21.06	38.73	54.94	70.97
DF	10	22	34	46
P-Value	0.021	0.015	0.013	0.011

The value of RSS is 4.11277.

II. AR(2) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	-0.0216	0.0193	-1.12	0.263
AR 2	0.0036	0.0193	0.19	0.853
Constant	0.001715	0.000757	2.27	0.023
Mean	0.001685	0.000743		

Residual Sums of Squares

DF	SS	MS
2679	4.11272	0.0015352

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	21.12	38.73	54.93	70.93
DF	9	21	33	45
P-Value	0.012	0.011	0.010	0.008

The value of RSS is 4.11272.

III. MA(1) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
MA 1	0.0215	0.0193	1.11	0.265
Constant	0.001685	0.000740	2.28	0.023
Mean	0.001685	0.000740		

Residual Sums of Squares

DF	SS	MS
2680	4.11278	0.0015346

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	21.06	38.73	54.95	70.98
DF	10	22	34	46
P-Value	0.021	0.015	0.013	0.010

The value of RSS for this model is 4.11278.

IV. MA(3) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
MA 1	0.0220	0.0193	1.14	0.255
MA 2	-0.0040	0.0193	-0.20	0.838
MA 3	-0.0110	0.0193	-0.57	0.570
Constant	0.001685	0.000751	2.24	0.025
Mean	0.001685	0.000751		

Residual Sums of Squares

DF	SS	MS
2678	4.11215	0.0015355

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	21.02	38.56	54.88	71.00
DF	8	20	32	44
P-Value	0.007	0.008	0.007	0.006

The value of RSS for this model is 4.11215.

V. MA(5) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
MA 1	0.0231	0.0193	1.20	0.231
MA 2	-0.0033	0.0193	-0.17	0.864
MA 3	-0.0120	0.0193	-0.62	0.534
MA 4	-0.0149	0.0193	-0.77	0.441
MA 5	-0.0110	0.0193	-0.57	0.568
Constant	0.001683	0.000771	2.18	0.029
Mean	0.001683	0.000771		

Residual Sums of Squares

DF	SS	MS
2676	4.11073	0.0015361

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	19.53	37.49	53.94	70.34
DF	6	18	30	42
P-Value	0.003	0.005	0.005	0.004

For this model, the RSS is 4.11073.

So from all of the above Autoregressive and Moving Average models, we can see the AR(2) has the lowest RSS(4.11272) among other AR models and MA(5) has the lowest RSS(4.11073) among other MA models. Hence, for the ARIMA model, we are going to take Autoregressive order 2, Difference order 1, and Moving Average Order 5.

VI. ARIMA(2,1,5) :

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	-1.6975	0.0181	-93.91	0.000
AR 2	-0.8227	0.0489	-16.82	0.000
MA 1	-1.66976	0.00889	-187.73	0.000
MA 2	-0.7715	0.0269	-28.71	0.000
MA 3	0.0168	0.0384	0.44	0.661
MA 4	-0.0237	0.0385	-0.62	0.538
MA 5	-0.0018	0.0214	-0.08	0.934
Constant	0.00593	0.00259	2.29	0.022

Residual Sums of Squares

DF	SS	MS
2674	4.09752	0.0015324

Back forecasts excluded

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	16.64	32.64	47.25	62.99
DF	4	16	28	40
P-Value	0.002	0.008	0.013	0.012

The value of RSS is 4.09752.

Since, the value of RSS for this ARIMA(2,1,5) is the lowest(4.09752) among all of the previous Autoregressive, Moving Average and ARIMA models, this is the most accurate model to predict future Bitcoin prices.

Forecasting :

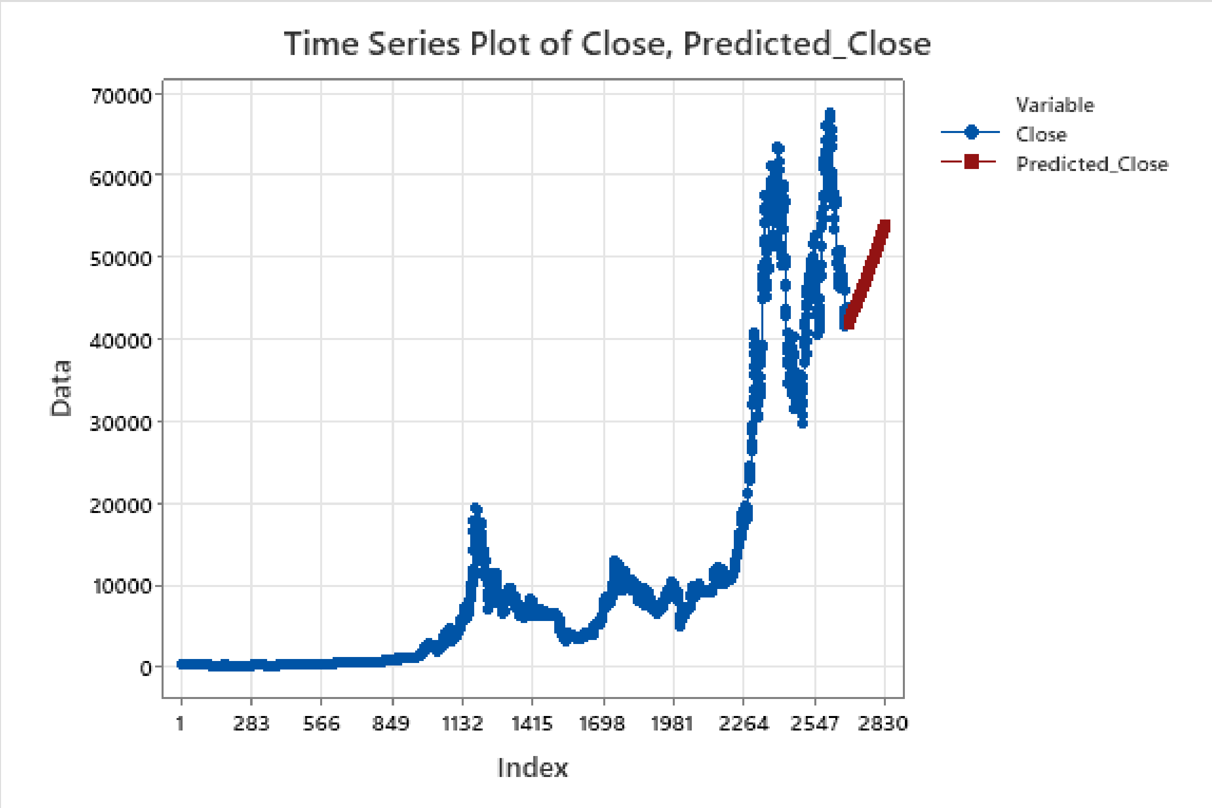
Using the forecast tool provided in Minitab, we are going to predict future values of "ln_Close" and then taking exponential of the values, we are going to get respective future Bitcoin prices.

There are all total 150 predictions, we take exponential of the values to get actual Bitcoin "Close" prices. The first 25 of the forecasts and their exponential values are:

Forecast
10.6451
10.6467
10.6484
10.6501
10.6518
10.6535
10.6552
10.6569
10.6585
10.6602
10.6619
10.6636
10.6653
10.667
10.6687
10.6703
10.672
10.6737
10.6754
10.6771
10.6788
10.6805
10.6821
10.6838
10.6855

Predicted_Close
41986.35653
42053.58847
42125.14038
42196.81402
42268.60961
42340.52736
42412.56747
42484.73016
42552.76014
42625.16135
42697.68576
42770.33356
42843.10496
42916.00018
42989.01943
43057.85692
43131.11753
43204.50279
43278.01291
43351.6481
43425.40858
43499.29456
43568.94914
43643.07935
43717.33568

Plotting the predicted Bitcoin “Close” Values with the actual “Close” values:



As we can see from the Analysis, that ARIMA(2,1,5) is the most accurate model to predict future Bitcoin “Low” prices among all other AR, MA and ARIMA model and from the graph, it is clear that there is going to be an increase in the price of Bitocoin meaning it is a good opportunity for Investors to invest and make profit over time.

• Conclusion :

In conclusion, my project focused on forecasting Bitcoin Prices using time series analysis which is going to help traders, investors who seek to gain profit by investing.

Throughout this project, it is evident that ARIMA models is a valuable tool for predicting Bitcoin prices by leveraging historical data and capturing important patterns and dependencies. By considering autoregressive, moving average and differencing components, the models accounted for trends, seasonality and the relationship between past and present observations.

Ultimately, this project contributed to a better understanding of the potential trends and movements in Bitcoin's prices. The forecast model generated by the ARIMA model can be valuable for investors, traders, and other stakeholders in making informed decisions, develop effective strategies, and decrease risks associated with Bitcoin investments. This project provides a valuable tool for navigating the dynamic and ever-changing landscape of the Bitcoin market, by optimizing investor's decision making processes and helping them gain profit over time.

• Bibliography :

The sources from where I gathered knowledge to make this project are:

I. https://www.wikipedia.org/

II. Fundamental of statistics vol II by Goon, A.M., Gupta, M.K. and Dasgupta, B.

III. Fundamental of Applied statistics by Gupta, S.C. and Kapoor, V.K.

• Acknowledgement :

I would like to express my gratitude to our Department faculty, Mr. Nilkanta Mukherjee, Mrs. Sutapa Biswas and Ms. Riddhi Das Majumder for their invaluable guidance, expertise and unwavering support throughout the entire process. Their insights, feedback and encouragement have been instrumental in shaping this project. I would also like to thank my friends and family who have stood by me and helped me in this project. Thank you all once again for your support, encouragement and contributions.