Table 4.1 represents the summary statistics of the data.

From table 4.1, the minimum quarterly price is GHS 1.14 and the maximum price is 387.89. The 1st Quartile of the data is GHS 7.68 which indicates that 25% of the observations fall below GHS 7.68. The 3rd Quartile of the data is GHS 141.34, which indicates that 25% of the observation are above GHS 141.34. The average quarterly price of tomatoes is approximately GHS 92.93 and the median is GHS 30.10.The variation in the quarterly price of tomatoes is GHS 13581.26, with a standard deviation of GHS 116.54. The skewness test has an output of 1.21 which shows that the data is skewed to the right.

Table 4.1: Descriptive statistics for open bitcoin prices

Mean Med 1st Qu 3rd Qu Var Std Skew Max Min

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Mean | Median | 1st Qu. | 3rd Qu. | Variance | Std | Min | Max | Skewness |
| 92.93 | 30.10 | 7.68 | 141.34 | 13581.26 | 116.54 | 1.14 | 387.89 | 1.21 |

Table 4.1

4.3 Examine Patterns of Monthly Bitcoin Prices

This section examines the pattern the data takes. We analyze if there is a trend or seasonality and conduct the necessary test to prove these patterns.

4.3.1 Trend Analysis of Data

The results of the graphical analysis of the quarterly tomato price per crate in Ashanti area is presented in Figure 4.1. The data shows an upward trend in “quarterly data” according to years. There were no missing values in the data, however. The data depicts some outliers. This makes the data negatively skewed (skewed to the right). Analysis of the results reveals that trend and seasonality is evident in the observation. The tomato price trend did not reveal any significant trend from 1995 to 1997. However, a steadily increasing trend was observed between 2000 and 2010. Besides, a sharp significant increasing trend is observed between 2010 and 2020. (Fig 4.1) denotes that the data are non-stationary and seasonal. As a result, the data does not have a constant mean or variance.

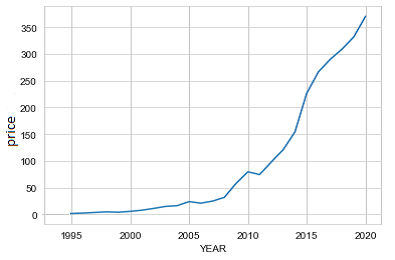


Figure 4.1 (Price Trend for Tomato from 1995 to 2020)

Hypothesis Test For Trend Analysis

MANN-KENDALL TEST

Table 4.2 shows the p-value obtained from the Mann-Kendall test when compared to an alpha of 0.05 reject the null hypothesis and concludes that the data contains a trend.

Table 4.2: Mann-Kendall Test of Data

Tau P-value

|  |  |
| --- | --- |
| Tau | P-Value |
| 3.272 | 2.2e-16 |

SEN-SLOPE TEST

Table 4.3 shows the p-value obtained from the Sen-slope test (2.2e-16) when compared to an alpha of 0.05 reject the null hypothesis and concludes that the slope of the data is not equal to zero and this indicates that, the data contains a trend. The Sen's slope value (2.552) indicates a positive trend in the data.

Table 4.3: Sen-Slope Test of Data

|  |  |
| --- | --- |
| Sen's slope | P-Value |
| 2.552 | 2.2e-16 |

4.3.2 Seasonality Test

From Fig 4.1, the data experienced an upward trend. The pattern of the graph is then

tested to see if a seasonal component exists in the data.

By comparing the respective p-value in Table 4.4 with the default p-value of 0.05, the

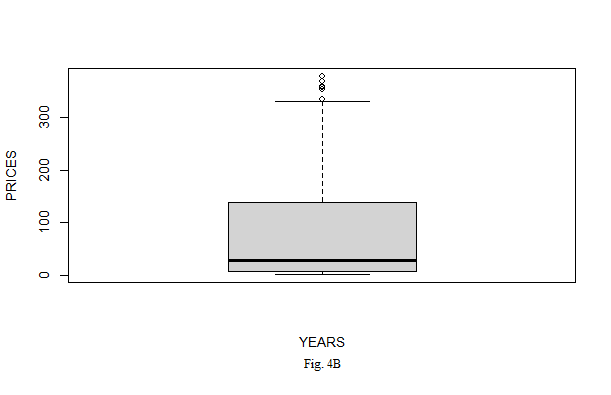
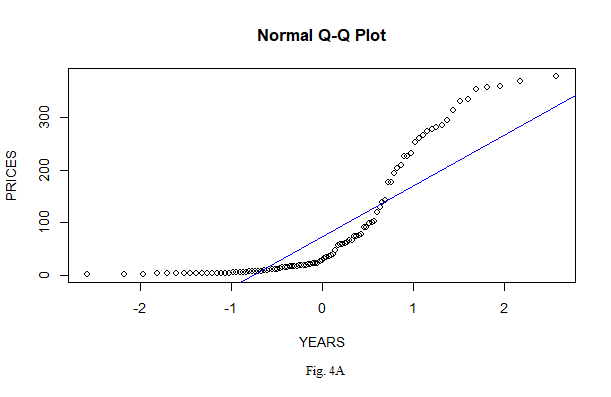
welch Seasonality test and Kruskal Wallis test had p-values of 1.91e-06 and 1.91e-06 respectively which are both greater than the default alpha value (0.05) hence, we fail to accept the null hypothesis and conclude that the data contains seasonal components hence, the need to difference the data.

Table 4.4: Seasonality Test of Data

|  |  |  |
| --- | --- | --- |
| Test | Test | P-value |
| Welch Seasonality | 12.86 | 1.91e-06 |
| Kruskal Wallis | 51.71 | 3.46e-11 |

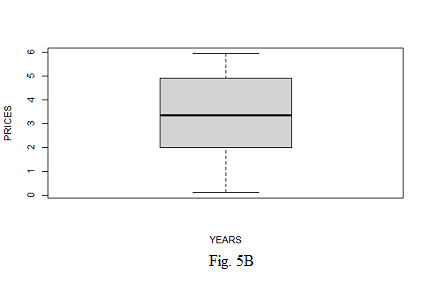
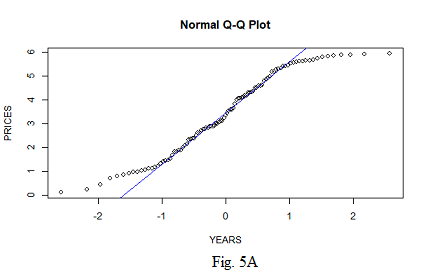
4.3.3 Box-Cox Transformation

From Fig 4.3A, we observe that majority of the points moves away from the line of best and Fig 4.3B also shows there exist some outliers therefore making the data not normal. λ also appears to be closer to 0 from the positive side. Further test was conducted and the specific value of λ = 0.153 therefore, we take the natural logarithm of the data to stabilize the variations.



4.3.4 Transformed Time Series Plot

From Fig 4.5, The transformed data now corrects both plots to depicts the features of normality.



4.4 Stationarity and Differencing

4.4.1 Stationarity

The data shows that there exists a trend in the data set. For this section, we perform three-unit root tests to determine objectively whether the monthly opening price of Bitcoin is stationary. The tests together with their hypothesis are given below.

From Table 4.5, it can be seen that the p-value (0.356) of the Augmented Dickey-Fuller (ADF) test is greater than the alpha level of (0.05); hence we fail to reject the null hypothesis () Subsequently, the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test has a p-value (0.01) which is less than alpha (0.05); hence we reject the null hypothesis (). The Phillips-Perron (PP) test has a p-value (0.010) which is lesser than the alpha value (0.05); hence we fail to reject the null hypothesis (). We conclude that the quarterly tomato price is not stationary since the ADF and KPSS test attests to it. As a result, to achieve stationarity, the series must be differenced.

*Table 4.5 (ADF, KPSS and PP Test for stationarity)*

|  |  |  |  |
| --- | --- | --- | --- |
| Test | Test Statistic | Lag parameter | P-value |
| ADF | -2.232 | 4 | 0.356 |
| KPSS | 2.104 | 4 | 0.010 |
| PP | -87.199 | 4 | 0.010 |

Plot of Differenced Data

Figure 4.6 is the time series plot for the differenced data. From the plot, there appears

to be no obvious trend. This is an indication that the differenced data has achieved

stationarity.

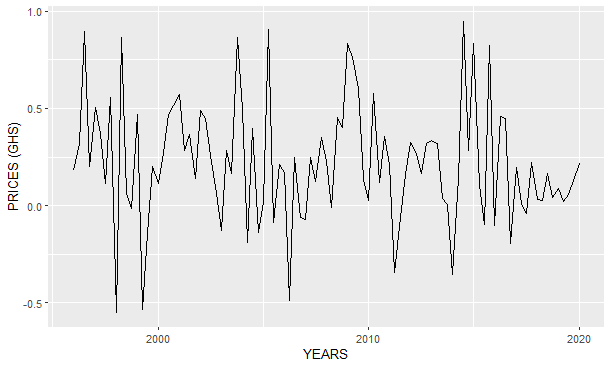


Figure 4.6 (Time Series Plot of Differenced Data)

4.4.3 Stationarity Test on The Differenced Data

From Table 4.6, it can be seen that the p-value (0.018) of the Augmented Dickey-Fuller (ADF) test is lesser than the alpha level of (0.05); hence we fail to accept the null hypothesis () Subsequently, the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test has a p-value (0.100) which is greater than the alpha value (0.05); hence we reject the null hypothesis () and having the p-value (0.010) for Phillips-Perron (PP) test being lesser than the alpha value (0.05), we fail to reject the null hypothesis (). We conclude that the quarterly tomato price is stationary since all the tests attests to it now.

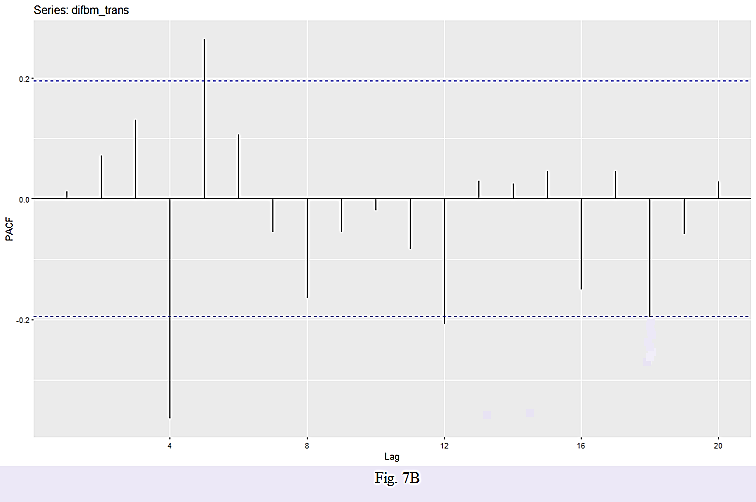
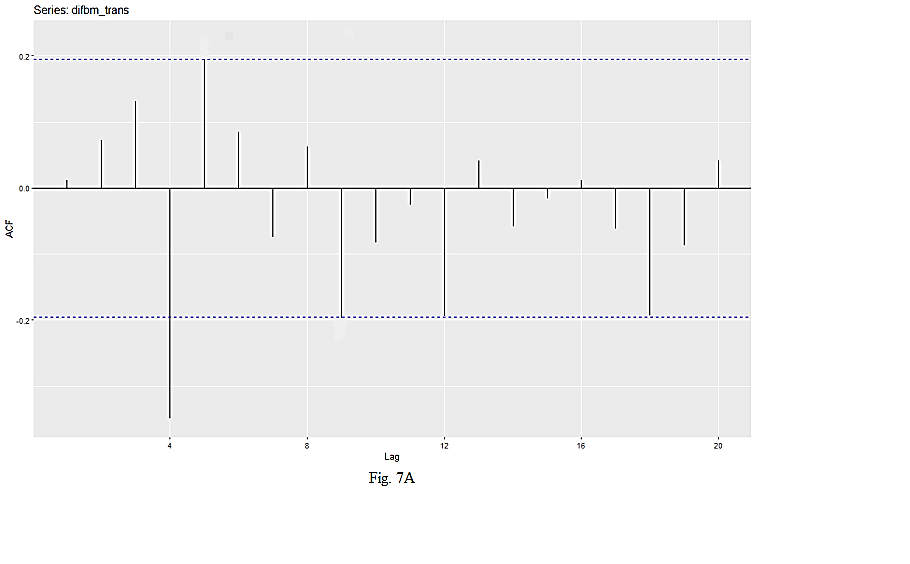
*Table 4.6 (ADF, KPSS and PP Test for stationarity)*

|  |  |  |  |
| --- | --- | --- | --- |
| Test | Test Statistic | Lag parameter | P-value |
| ADF | -4.274 | 4 | 0.010 |
| KPSS | 0.150 | 4 | 0.100 |
| PP | -100.26 | 4 | 0.010 |

4.5 Model Identification And Selection

4.5.1 ACF And PACF Of Differenced Data

From Fig 4.7A, the ACF becomes significant only at lag 4. Similarly, for the PACF plot Fig 4.7B, the PACF is slightly significant at lag 4,5 and 12 suggesting the presence of an AR component. Both the ACF and PACF plots reveal little information about the appropriate order of the data hence we fit competing models and select the model with the least information criterion.



4.5.2 Information Criteria Approach

From Table 4.7, the model with the least AIC and AICc is ARIMA [3,0,0][0,0,1] with values of 77.7 and 78.9respectively. The model with the least BIC is ARIMA[1,0,0][0,0,1] with a value of 91.93.

We subject both models to residual diagnostics test.

|  |  |  |  |
| --- | --- | --- | --- |
| MODEL | AIC | BIC | AICc |
| ARIMA [0,0,0][0,0,1] | 86.42 | 94.27 | 86.67 |
| ARIMA [1,0,0][0,0,1] | 81.47 | 91.93 | 81.89 |
| ARIMA[3,0,0][0,0,1] | 80.16 | 95.85 | 81.05 |
| ARIMA [3,0,1][0,0,1] | 77.7 | 96.01 | 78.9 |

4.6 Residual Diagnostics

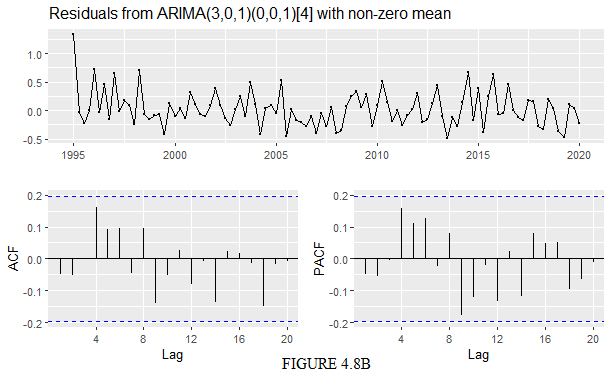
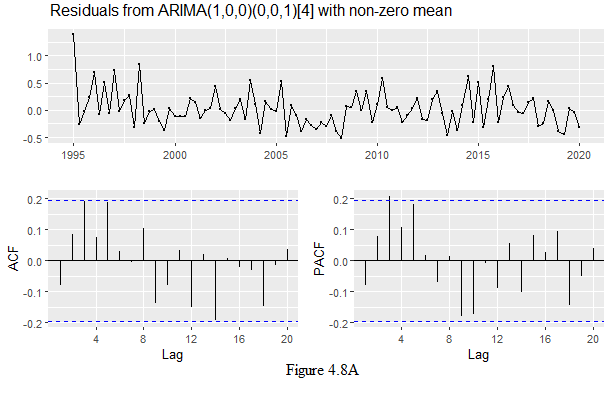
4.6.1 Residual Diagnostic for ARIMA [1,0,0][0,0,1]

From Figure 4.8A, from the ACF plot, it can be observed that there exist spikes that falls on the threshold line which are not really significant but the PACF do have a significant spike hence, it can be deduced that the errors are correlated.

4.6.2 Residual Diagnostic for ARIMA [3,0,1][0,0,1]

From Figure 4.8B, the plot of the residuals shows no sign of a changing mean. This indicates that mean of the errors is constant. From the ACF plot, it can be observed that there are no signs of a significant spike and so it is for the PACF plot. This is an indication that the errors are uncorrelated.

Conclusively; ARIMA [3,0,1][0,0,1] appears more appropriate relative to ARIMA [1,0,0][0,0,1] because, ARIMA [3,0,1][0,0,1] had its spikes relatively further away from the theshold than ARIMA [1,0,0][0,0,1].



4.6.3 Ljung-Box and Shapiro-Wilk Test

From Table 4.8, For the Ljung-Box test, the p-value (0.091) is greater than alpha (0.05)

hence we fail to reject the null hypothesis and conclude that the ARIMA [3,0,1][0,0,1] model is

appropriate. For the Shapiro-Wilk Normality test, the p-value (0.424) is greater than

alpha (0.05), hence we fail to reject the null hypothesis and conclude that the errors of

the ARIMA [3,0,1][0,0,1] model are normally distributed.

.

|  |  |  |  |
| --- | --- | --- | --- |
| MODEL | TEST | TEST STATISTICS | P-VALUE |
| ARIMA [3,0,1][0,0,1] | Ljung-Box test | 6.490 | 0.091 |
| Shapiro-Wilk Normality Test | 0.972 | 0.424 |

4.7 Forecasting

4.7.1 Coefficient of ARIMA [3,0,1][0,0,1]

Table 4.9 shows the coefficients of the model ARIMA [3,0,1][0,0,1] which was used to generate

a mathematical model which can be used to predict the quarterly prices of tomatoes in Ashanti Region for the next six(6) years.

.

Table 4.9: Coefficients of *:* ARIMA (3,0,1) (0,0,1) [4]

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | *AR (1)* | *AR (2)* | *AR (3)* | *MA (1)* | *SMA (1)* |
| *Coefficient* | *-0.638* | *0.379* | *0.290* | *0.967* | *-0.959* |
| *S.E.* | *0.116* | *0.140* | *0.116* | *0.121* | *0.200* |
| *Best competing model**: ARIMA (*3,0,1*) (0,0,1) [4]* | | | | |  |

MATHEMATICAL MODEL

From Table 4.7, the model with the least AIC and AICc is ARIMA(5,0,0) with values of 105.07 and 106.28 respectively. The model with the least BIC is ARIMA(3,0,2) with a value of 110.20.

We subject both models to residual diagnostics test.

|  |  |  |  |
| --- | --- | --- | --- |
| MODEL | AIC | BIC | AICc |
| ARIMA(5,0,0) | 105.07 | 123.38 | 106.28 |
| ARIMA(4,0,1) | 109.31 | 127.62 | 110.52 |
| ARIMA(3,0,1) | 106.21 | 121.90 | 107.10 |
| ARIMA(3,0,2) | 108.99 | 110.20 | 127.30 |

4.6 Residual Diagnostics

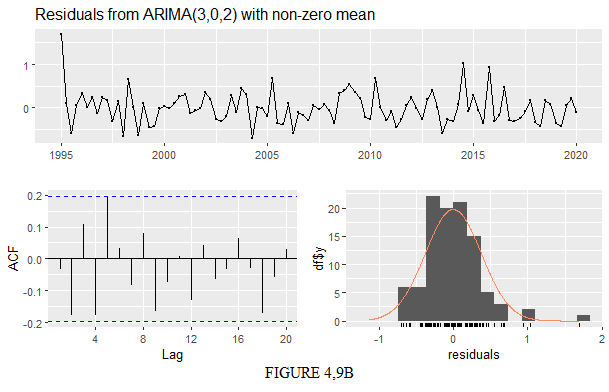
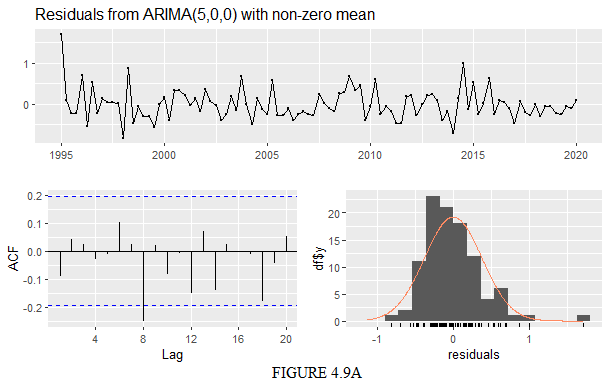
4.6.1 Residual Diagnostic for ARIMA(5,0,0)

From Figure 4.9A, from the ACF plot, it can be observed that there exist a significant spike at lag 8 hence, it can be deduced that the errors are correlated. The histogram also does not completely show normality within the residuals of the model and with the model having a Ljung-Box test p-value (0.003386) which is less than the default alpha value (0.05), we conclude that the model is not appropriate.

4.6.2 Residual Diagnostic for ARIMA(3,0,2)

From Figure 4.9B, the plot of the residuals shows no sign of a changing mean. This indicates that mean of the errors is constant. From the ACF plot, it can be observed that there are no signs of a significant spike. This is an indication that the errors are uncorrelated. Moreover, the model had a Ljung-Box test p-value (0.0001021) which is less than the default alpha value (0.05), we conclude that the model is not much appropriate.

Conclusively; ARIMA (3,0,2) appears more appropriate relative to ARIMA (5,0,0) because ARIMA (3,0,2) had its spikes relatively further away from the theshold than ARIMA (5,0,0).



4.7 Forecasting

4.7.1 Coefficient of ARIMA [3,0,1][0,0,1]

Table 4.7 is a summary of the forecasted values from 1995 to 2020. The forecasted values are presented in their logarithmic forms. The table comprises point forecasts with 95% confidence intervals.

Table 4.7 (Forecast Value by ARIMA (3,0,1) (0,0,1) [4] From 2020 to 2025)

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | 95% Confidence Interval | |
| YEAR QTR | Point Forecast  (GHC) | Lower Limit | Upper Limit  (GHC) |
| 2020 Q3 | 351.00 |  |  |
| 2020 Q4 | 358.47 | 311.70 | 405.25 |
| 2021 Q1 | 343.53 | 295.50 | 391.55 |
| 2021 Q2 | 451.53 | 398.59 | 504.47 |
| 2021 Q3 | 380.71 | 322.00 | 439.41 |
| 2021 Q4 | 376.51 | 310.82 | 442.19 |
| 2022 Q1 | 373.68 | 302.32 | 445.04 |
| 2022 Q2 | 460.27 | 384.45 | 536.10 |
| 2022 Q3 | 407.13 | 321.96 | 492.30 |
| 2022 Q4 | 406.82 | 317.21 | 496.43 |
| 2023 Q1 | 401.07 | 307.05 | 495.09 |
| 2023 Q2 | 484.61 | 385.24 | 583.98 |
| 2023 Q3 | 430.41 | 319.86 | 540.97 |
| 2023 Q4 | 430.36 | 313.53 | 547.18 |
| 2024 Q1 | 425.09 | 302.40 | 547.78 |
| 2024 Q2 | 508.87 | 379.63 | 638.10 |
| 2024 Q3 | 454.67 | 313.51 | 595.81 |
| 2024 Q4 | 454.54 | 306.03 | 603.04 |
| 2025 Q1 | 449.23 | 293.76 | 604.70 |
| 2025 Q2 | 533.00 | 369.94 | 696.05 |
| 2025 Q3 | 478.81 | 303.16 | 654.46 |
| 2025 Q4 | 478.69 | 294.72 | 662.65 |