FINAL PROJECT
LINEAR REGRESSION & TIME SERIES ANALYSIS

FORECASTING THE PT Bank Jago Tbk (ARTO) STOCK PRICE USING ARIMA MODEL

I GUSTI PUTU RANANTHA M.P. – 021201900008

CONTENT

65.74%

NUMBER OF INVESTOR INCREASE ON SEPTEMBER 2021
(CNBC INDONESIA,2021)

6,459.68

IHSG LEVEL ON OCTOBER 11, 2021

COVID-19 PANDEMIC

Made lot of aspect in Business had to adapt in digitalization

Indonesia will facing the Era of Digital Bank

47,722,913

USERS OF DIGIBANK IN 2021

74,785,062

PROECTION OF USERS ON DIGIBANK IN 2026 (JAYANI,2021)

Jago

PATRICK WALUJO

JERRY NG

PT BANK ARTOS INDONESIA TBK OFFICIALLY BECAME A PUBLIC COMPANY AFTER OFFERING ITS INITIAL SHARES TO THE PUBLIC ON JANUARY 12, 2016 AND LISTING ITS SHARES ON THE INDONESIA STOCK EXCHANGE (IDX) WITH THE ISSUER CODE ARTO

Rp 4,310

PRICE OF ARTO STOCK IN DESEMBER 2020

439.5%

INCREASE PRICE OF ARTO FROM JANUARY 2020 - NOVEMBER 2021

Rp 15,375/Share on November 30, 2021

13,154.3103%

INCREASE PRICE

ARTO STOCK PRICE FROM IPO

THIS SIGNIFICANCE PROGRESS

AMAZED ME A LOT AND MADE ME WANT TO MAKE A FORECASTING FOR ARTO FOR THE NEXT 8 WEEKS

LITERATURE REVIEW

- TIME SERIES
- STOCK

TIME SERIES

A time series is a single dataset of observations arranged in chronological order. The main difference from time series data is that the observations are related to a single quantity measured at multiple time points.

HOW TO ANALYZE TIME SERIES?

To analyze Time Series data, we must run several ways such as (Pandian, 2021):

- Collecting the data and cleaning it
- Preparing Visualization with respect to time vs key feature
- Observing the stationarity of the series
- Developing charts to understand its nature.
- Model building – AR, MA, ARMA and ARIMA
- Extracting insights from prediction

STOCK

DEFINITION

A sign of capital participation of a person or party (business entity) in a company or limited liability company (Perseroan Terbatas)

BENEFIT

- Dividend
- Capital Gain

RISK

- Capital Loss
- Liquidation Risk

METHOD

Box-Jenkins Method

Model of Time Series

Parameters Estimation

Forecasting

BOX-JENKINS METHOD

Considered as a stochastic model building and that it is an iterative approach that consists of the following 3 steps (Brownlee, 2017)

- Identification
- Estimation
- Diagnostic Checking

BOX-JENKINS METHOD

IDENTIFICATION

a. Differencing
- Unit Root Tests
- Avoid over differencing

b. Configuring AR and MA
- Autocorrelation Function (ACF)
- Partial Autocorrelation Function (PACF)

ESTIMATION

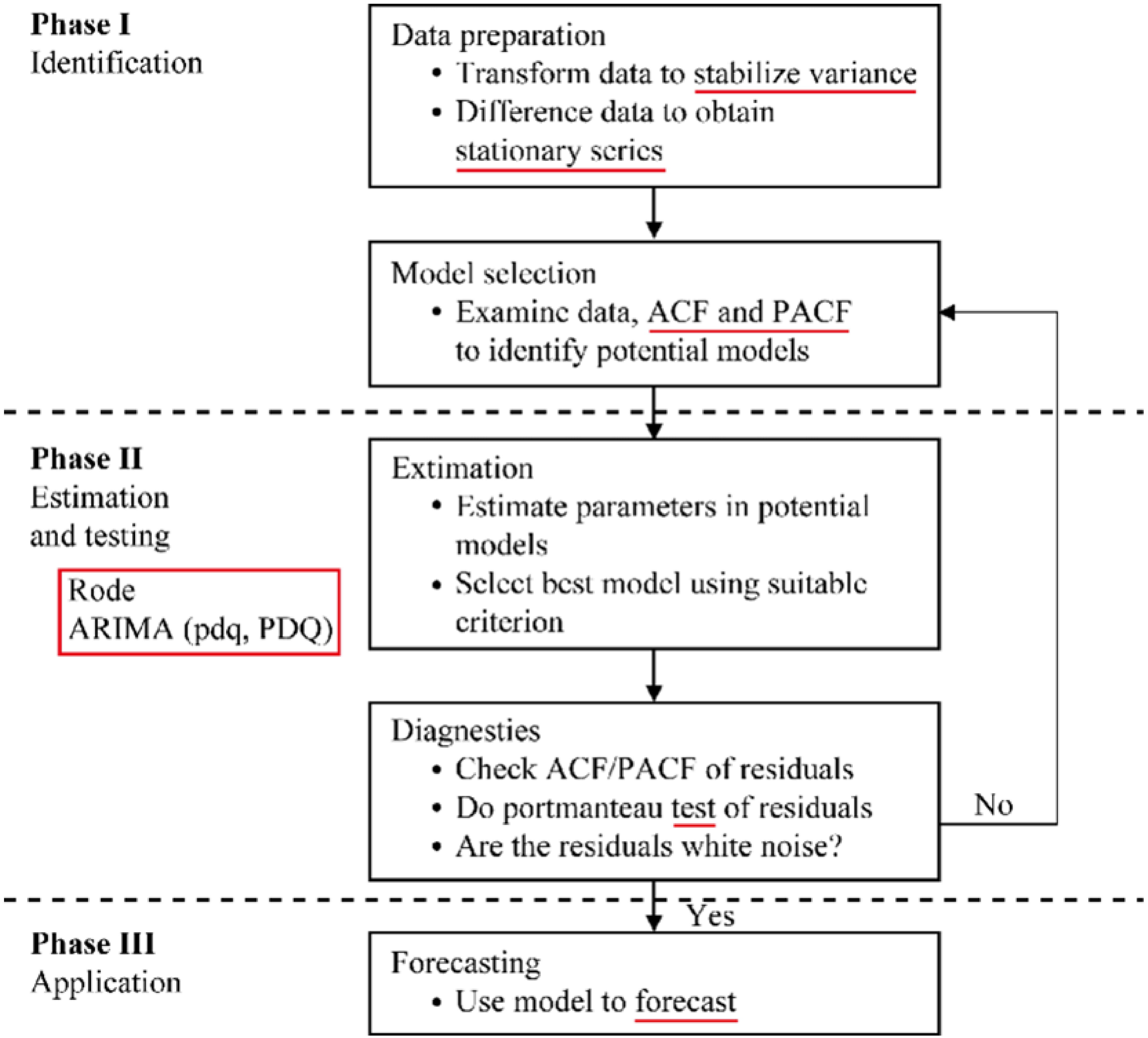
Estimation involves using numerical methods to minimize a loss or error term. In using estimation, we will worked with estimating model parameters for ARMA and ARIMA models

DIAGNOSTIC CHECKING

The idea of diagnostic checking is to look for evidence that the model is not a good fit for the data.

Two useful areas to investigate diagnostics are:

1. Overfitting
2. Residual Errors.


Phase I
Identification
Data preparation
• Transform data to stabilize variance
• Difference data to obtain stationary scrics
Modcl sclcction
• Examinc data, ACF and PACF to identify potential models
Phase II
Estimation and testing
Rode
ARIMA (pdq, PDQ)
Extimation
• Estimate parameters in potential models
• Sclcct bcst modcl using suitablc criterion
Diagnesties
• Check ACF/PACF of residuals
• Do portmanteau test of residuals
• Are the residuals white noise?
No
Yes
Phase III
Application
Forecasting
• Use model to forecast

BOX-JENKINS METHOD

The Box-Jenkins Model forecasts data using three principles: autoregression, differencing, and moving average. These three principles are known as p, d, and q, respectively. Each principle is used in the Box-Jenkins analysis; together, they are collectively shown as ARIMA (p, d, q).

BOX-JENKINS METHOD

The autoregression (p) process tests the data for its level of stationarity. If the data being used is stationary, it can simplify the forecasting process. If the data being used is non-stationary it will need to be differenced (d). The data is also tested for its moving average fit (which is done in part q of the analysis process). Overall, initial analysis of the data prepares it for forecasting by determining the parameters (p, d, and q), which are then applied to develop a forecast.

Model of Time Series

Autoregressive Model

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t$$

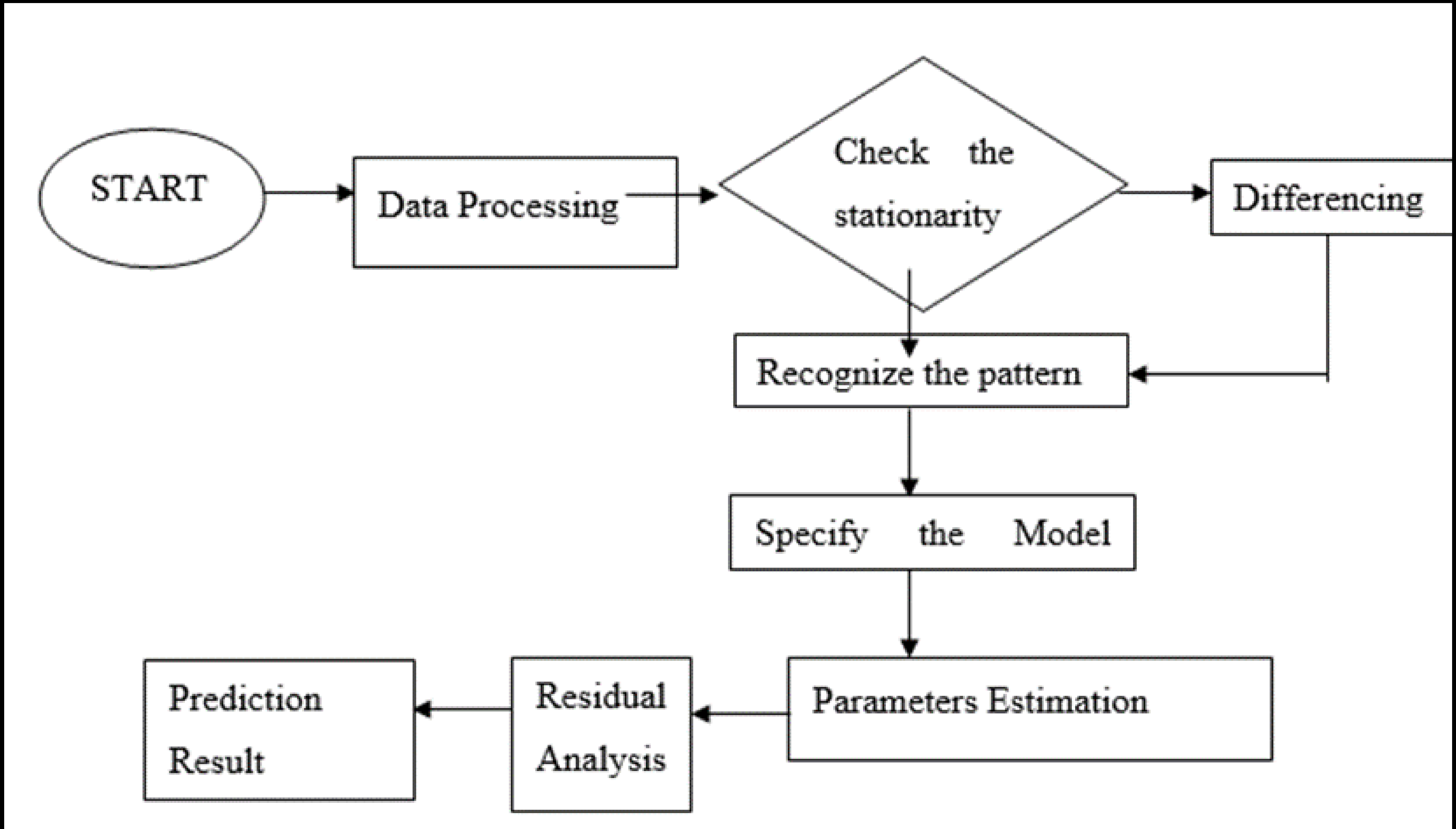
Moving Average

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

Autoregressive Moving Average (ARMA)

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

ARIMA
(Autoregressive Integrated Moving Average)



ARIMA
(Autoregressive Integrated Moving Average)

$$W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + \cdots + \phi_P W_{t-P} + \cdots + dW_{t-p-d} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \qquad (3.2.4.1)$$

The formula for ARIMA (1,0,0) is

$$W_t = \phi_1 W_{t-1} + e_t \qquad (3.2.4.2)$$

The formula for ARIMA (0,1,0) is

$$W_t = W_{t-1} + e_t \qquad (3.2.4.3)$$

The formula for ARIMA (1,1,0) is

$$W_t = \phi_1 W_{t-1} + W_{t-1} + e_t \qquad (3.2.4.4)$$

The formula for ARIMA (0,1,1) is

$$W_t = W_{t-1} + e_t - \theta_1 e_{t-1} \qquad (3.2.4.5)$$

The formula for ARIMA (1,1,1) is

$$W_t = \phi_1 W_{t-1} + W_{t-1} + e_t - \theta_1 e_{t-1} \qquad (3.2.4.6)$$

3.3 Parameters Estimation

Least Square Estimation

For the first-order Autoregressive (AR) model, it is written as (Cryer & Chan, 2008, p. 154):

$$\boldsymbol{Y_t - \mu = \phi(Y_{t-1} - \mu) + e_t} \qquad \textbf{(3.3.2.1)}$$

$$(Y_t - \mu) - \phi(Y_{t-1} - \mu)$$

$$\boldsymbol{S_c(\phi, \mu) = \sum_{t=2}^{n}[(Y_t - \mu) - \phi(Y_{t-1} - \mu)]^2} \qquad \textbf{(3.3.2.2)}$$

3.3 Parameters Estimation

Least Square Estimation

For first order of Moving Average, we have

$$Y_t = e_t - \theta e_{t-1}$$

$$\boldsymbol{S_c(\theta) = \sum_{t=2}(e_t)^2 = (Y_t + \theta e_{t-1})^2} \qquad \textbf{(3.3.2.5)}$$

And for the first order of ARMA, we have

$$Y_t = \phi Y_{t-1} + e_t - \theta e_{t-1}$$

$$\boldsymbol{S_c(\theta) = \sum_{t=2}(e_t)^2 = (Y_t - \phi Y_{t-1} + \theta e_{t-1})^2} \qquad \textbf{(3.3.2.6)}$$

3.3 Parameters Estimation

Least Square Estimation

For Autoregressive (AR) models, the method of moments can be written as:

$$r_p = r_{p-1}\phi_1 + r_{p-2}\phi_2 + r_{p-3}\phi_3 + \cdots + \phi_p \quad (3.3.1.1)$$

and for the Moving Average (MA) models, for example, MA(1), the equation is as follows:

$$\hat{\theta} = \frac{-1+\sqrt{1-4r_1^2}}{2r_1} \quad (3.3.1.2)$$

3.4 Forecasting

The measuring instrument used to calculate prediction errors:

1. Mean Square Error (MSE)

$$MSE = \sum \frac{(At-Ft)^2}{n} \quad (3.4.1)$$

2. Mean Absolute Deviation (MAD)

$$MAD = \sum \frac{At-Ft}{n} \quad (3.4.2)$$

3. Mean Absolute Persentage Error (MAPE)

$$MAPE = \frac{100}{n} \sum At - \frac{Ft}{n}, \quad (3.4.3)$$

where,

At = Actual Demand in Period-t

Ft = Demand Forecast in Period-t

n = Number of Forecasting Periods Involved

RESULT & ANALYSIS

- Data Preparation and Processing

- Stationarity Check

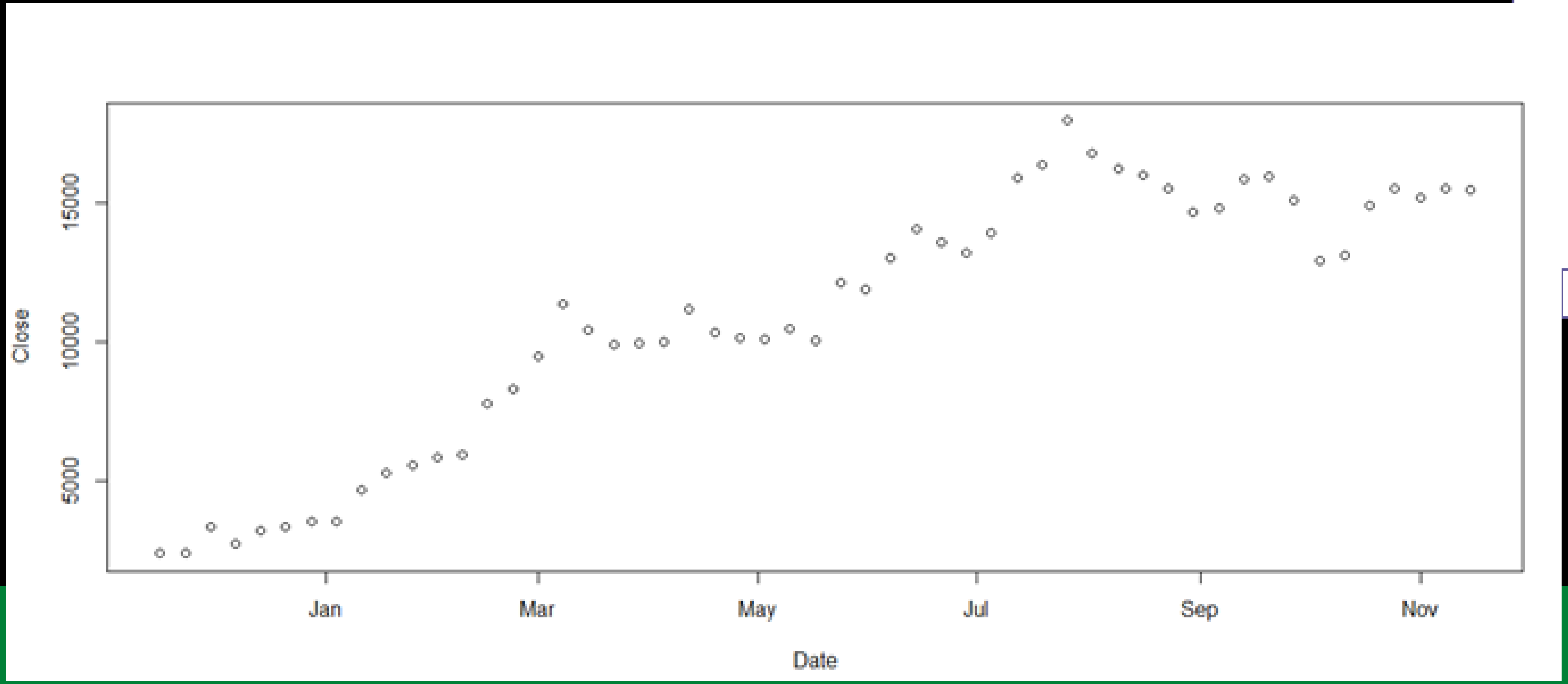
- Model Specification

-Parameter Estimation & Residual Analysis

- Forecasting

Data Preparation and Processing

```
#1 start
library(TSA)
library(forecast)
library(tseries)
library(ggplot2)

#2 data preparation
library(readxl)
ARTO_JK_Weekly_1_Year <- read_excel("D:/University/20211/
View(ARTO_JK_Weekly_1_Year)
plot(ARTO_JK_Weekly_1_Year)
```



The author uses a variable date and closing price that will be used to forecast stock prices in the next 8 weeks. Data taken from November 16, 2020 to November 15, 2021

STATIONARITY CHECK

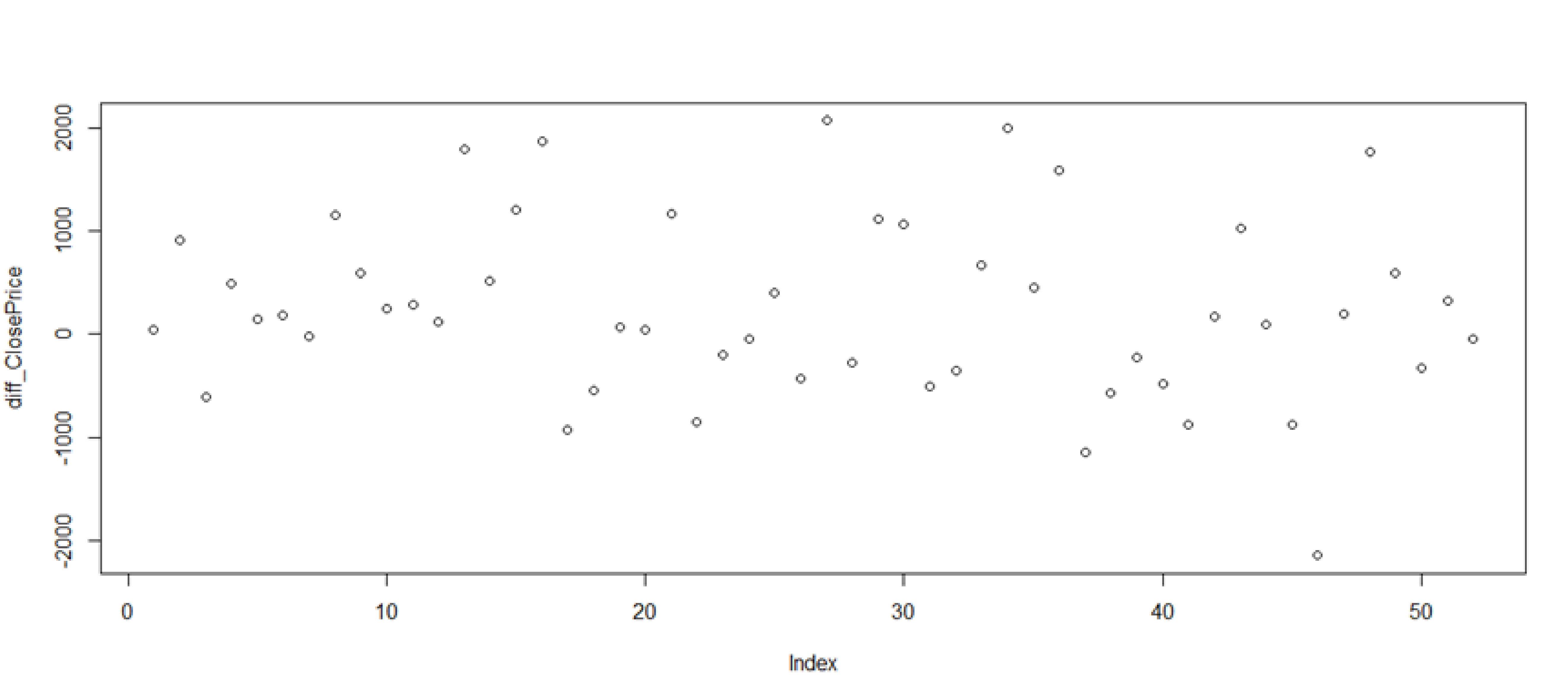
```
#3 check the staionarity
Date1 = ARTO_JK_Weekly_1_Year$Date
ClosePrice = ARTO_JK_Weekly_1_Year$`Close`
adf.test(ClosePrice)
acf(ClosePrice)
pacf(ClosePrice)

diff_ClosePrice <- diff(ClosePrice)
plot(diff_ClosePrice)
adf.test(diff_ClosePrice)
acf(diff_ClosePrice)
pacf(diff_ClosePrice)
```

STATIONARITY CHECK

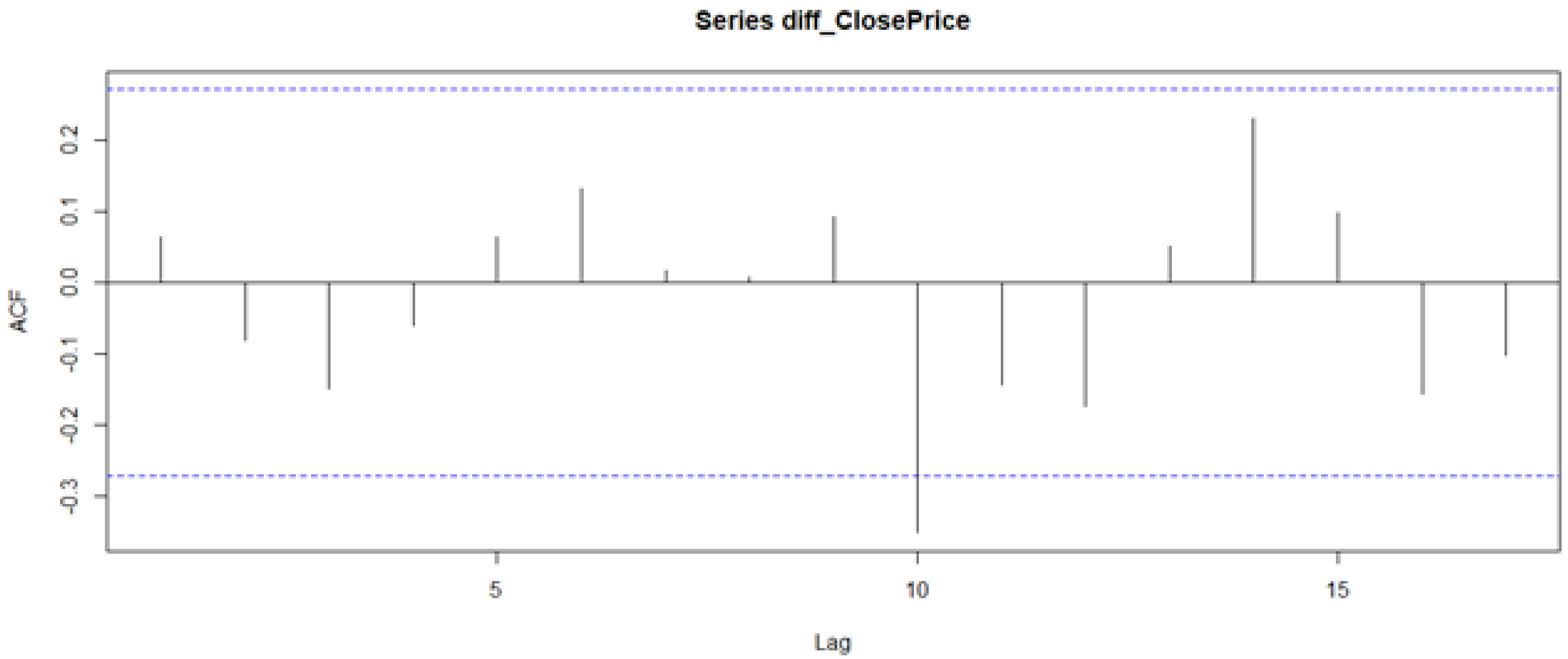
```
        Augmented Dickey-Fuller Test

data:  ClosePrice
Dickey-Fuller = -1.3168, Lag order = 3, p-value = 0.8494
alternative hypothesis: stationary
```

```
        Augmented Dickey-Fuller Test

data:  diff_ClosePrice
Dickey-Fuller = -4.2161, Lag order = 3, p-value = 0.01
alternative hypothesis: stationary
```

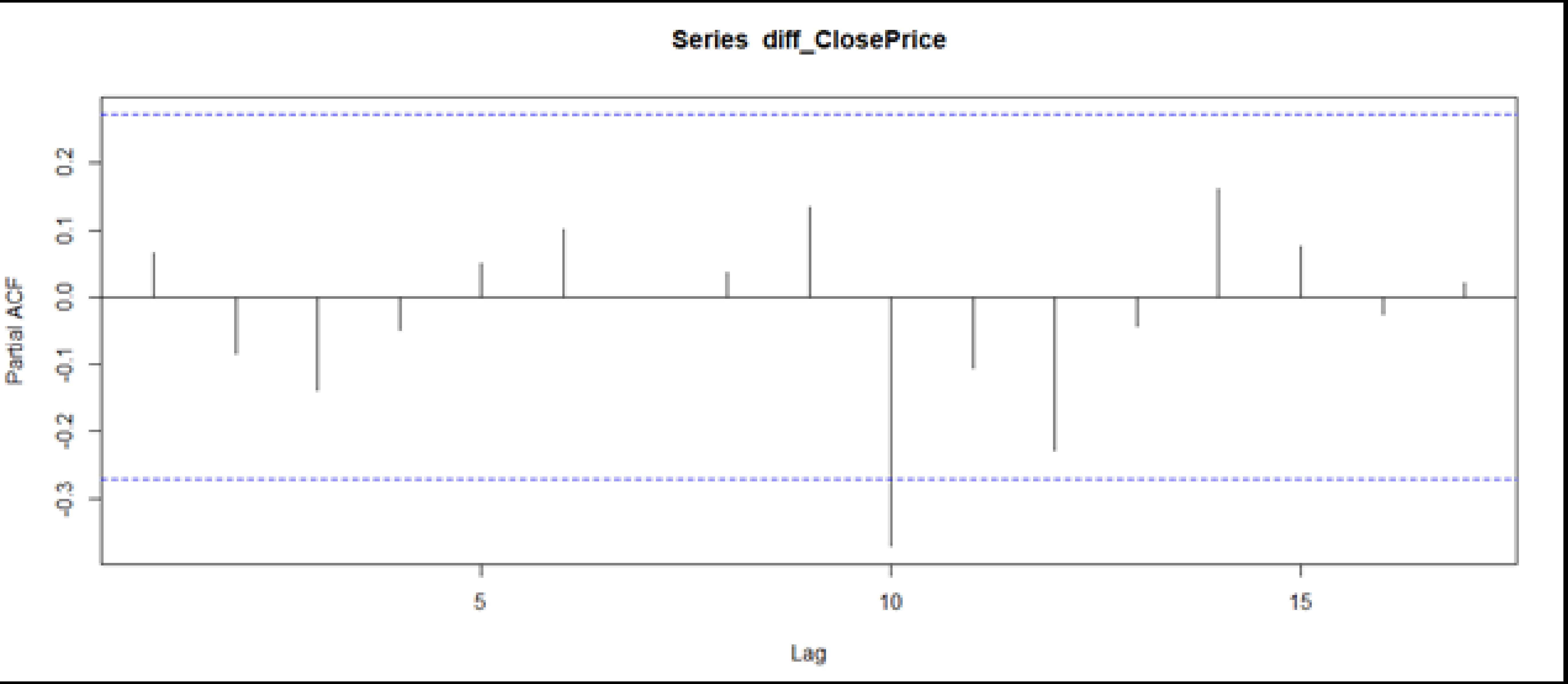
STATIONARITY CHECK



MODEL SPECIFICATION



MODEL SPECIFICATION



MODEL SPECIFICATION

ARIMA MODEL	p	d	q
ARIMA MODEL 1	0	1	0
ARIMA MODEL 2	1	1	0
ARIMA MODEL 3	2	1	0
ARIMA MODEL 4	3	1	0
ARIMA MODEL 5	4	1	0
ARIMA MODEL 6	5	1	0
ARIMA MODEL 7	6	1	0
ARIMA MODEL 8	7	1	0
ARIMA MODEL 9	8	1	0
ARIMA MODEL 10	9	1	0
ARIMA MODEL 11	10	1	0
ARIMA MODEL 12	0	1	1
ARIMA MODEL 13	1	1	1
ARIMA MODEL 14	2	1	1
ARIMA MODEL 15	3	1	1
ARIMA MODEL 16	4	1	1
ARIMA MODEL 17	5	1	1
ARIMA MODEL 18	6	1	1
ARIMA MODEL 19	7	1	1
ARIMA MODEL 20	8	1	1
ARIMA MODEL 21	9	1	1
ARIMA MODEL 22	10	1	1
ARIMA MODEL 23	0	1	2
ARIMA MODEL 24	1	1	2
ARIMA MODEL 25	2	1	2
ARIMA MODEL 26	3	1	2
ARIMA MODEL 27	4	1	2
ARIMA MODEL 28	5	1	2
ARIMA MODEL 29	6	1	2
ARIMA MODEL 30	7	1	2
ARIMA MODEL 31	8	1	2
ARIMA MODEL 32	9	1	2
ARIMA MODEL 33	10	1	2
ARIMA MODEL 34	0	1	3
ARIMA MODEL 35	1	1	3
ARIMA MODEL 36	2	1	3
ARIMA MODEL 37	3	1	3
ARIMA MODEL 38	4	1	3
ARIMA MODEL 39	5	1	3
ARIMA MODEL 40	6	1	3

MODEL SPECIFICATION

ARIMA MODEL	p	d	q
ARIMA MODEL 41	7	1	3
ARIMA MODEL 42	8	1	3
ARIMA MODEL 43	9	1	3
ARIMA MODEL 44	10	1	3
ARIMA MODEL 45	0	1	4
ARIMA MODEL 46	1	1	4
ARIMA MODEL 47	2	1	4
ARIMA MODEL 48	3	1	4
ARIMA MODEL 49	4	1	4
ARIMA MODEL 50	5	1	4
ARIMA MODEL 51	6	1	4
ARIMA MODEL 52	7	1	4
ARIMA MODEL 53	8	1	4
ARIMA MODEL 54	9	1	4
ARIMA MODEL 55	10	1	4
ARIMA MODEL 56	0	1	5
ARIMA MODEL 57	1	1	5
ARIMA MODEL 58	2	1	5
ARIMA MODEL 59	3	1	5
ARIMA MODEL 60	4	1	5

ARIMA MODEL	p	d	q
ARIMA MODEL 60	4	1	5
ARIMA MODEL 61	5	1	5
ARIMA MODEL 62	6	1	5
ARIMA MODEL 63	7	1	5
ARIMA MODEL 64	8	1	5
ARIMA MODEL 65	9	1	5
ARIMA MODEL 66	10	1	5
ARIMA MODEL 67	0	1	6
ARIMA MODEL 68	1	1	6
ARIMA MODEL 69	2	1	6
ARIMA MODEL 70	3	1	6
ARIMA MODEL 71	4	1	6
ARIMA MODEL 72	5	1	6
ARIMA MODEL 73	6	1	6
ARIMA MODEL 74	7	1	6
ARIMA MODEL 75	8	1	6
ARIMA MODEL 76	9	1	6
ARIMA MODEL 77	10	1	6
ARIMA MODEL 78	0	1	7
ARIMA MODEL 79	1	1	7
ARIMA MODEL 80	2	1	7

MODEL SPECIFICATION

ARIMA MODEL	p	d	q
ARIMA MODEL 81	3	1	7
ARIMA MODEL 82	4	1	7
ARIMA MODEL 83	5	1	7
ARIMA MODEL 84	6	1	7
ARIMA MODEL 85	7	1	7
ARIMA MODEL 86	8	1	7
ARIMA MODEL 87	9	1	7
ARIMA MODEL 88	10	1	7
ARIMA MODEL 89	0	1	8
ARIMA MODEL 90	1	1	8
ARIMA MODEL 91	2	1	8
ARIMA MODEL 92	3	1	8
ARIMA MODEL 93	4	1	8
ARIMA MODEL 94	5	1	8
ARIMA MODEL 95	6	1	8
ARIMA MODEL 96	7	1	8
ARIMA MODEL 97	8	1	8
ARIMA MODEL 98	9	1	8
ARIMA MODEL 99	10	1	8
ARIMA MODEL 100	0	1	9

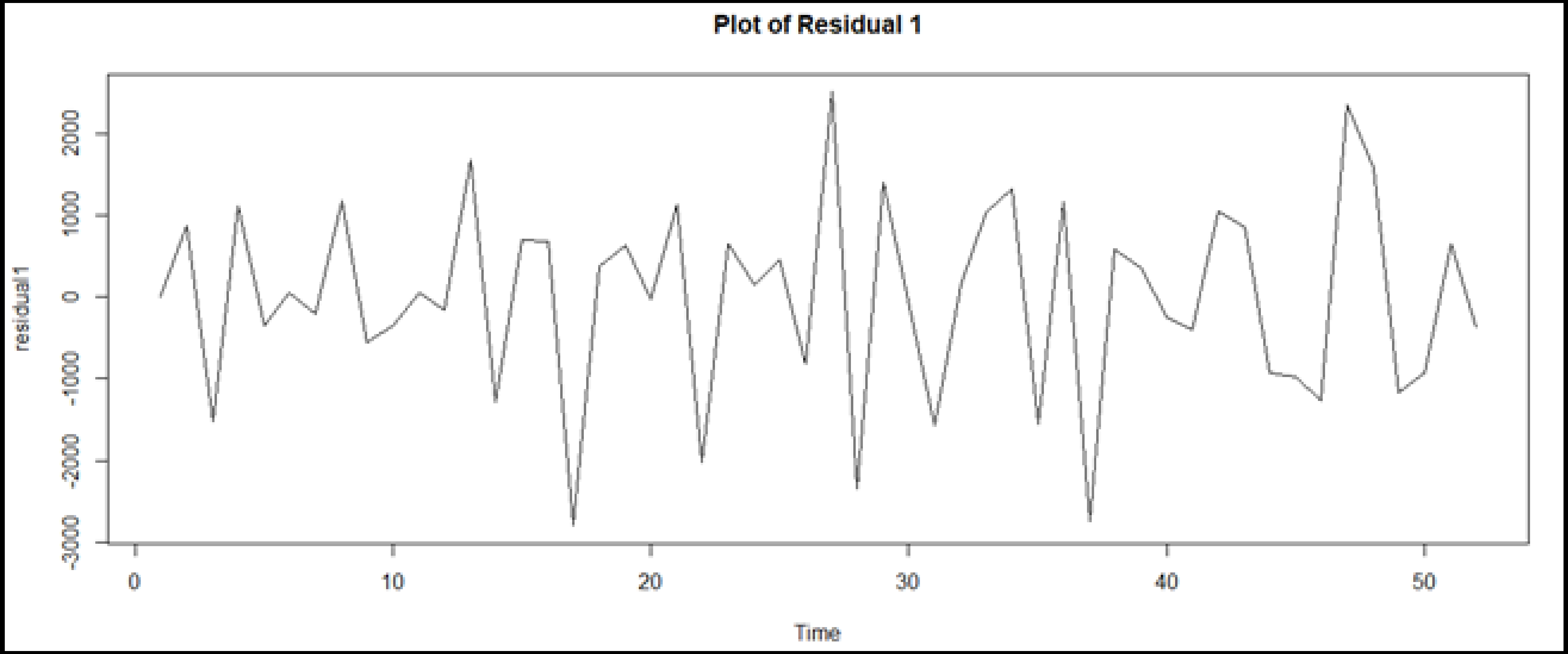
ARIMA MODEL	p	d	q
ARIMA MODEL 101	1	1	9
ARIMA MODEL 102	2	1	9
ARIMA MODEL 103	3	1	9
ARIMA MODEL 104	4	1	9
ARIMA MODEL 105	5	1	9
ARIMA MODEL 106	6	1	9
ARIMA MODEL 107	7	1	9
ARIMA MODEL 108	8	1	9
ARIMA MODEL 109	9	1	9
ARIMA MODEL 110	10	1	9
ARIMA MODEL 111	0	1	10
ARIMA MODEL 112	1	1	10
ARIMA MODEL 113	2	1	10
ARIMA MODEL 114	3	1	10
ARIMA MODEL 115	4	1	10
ARIMA MODEL 116	5	1	10
ARIMA MODEL 117	6	1	10
ARIMA MODEL 118	7	1	10
ARIMA MODEL 119	8	1	10
ARIMA MODEL 120	9	1	10
ARIMA MODEL 121	10	1	10

MODEL SPECIFICATION

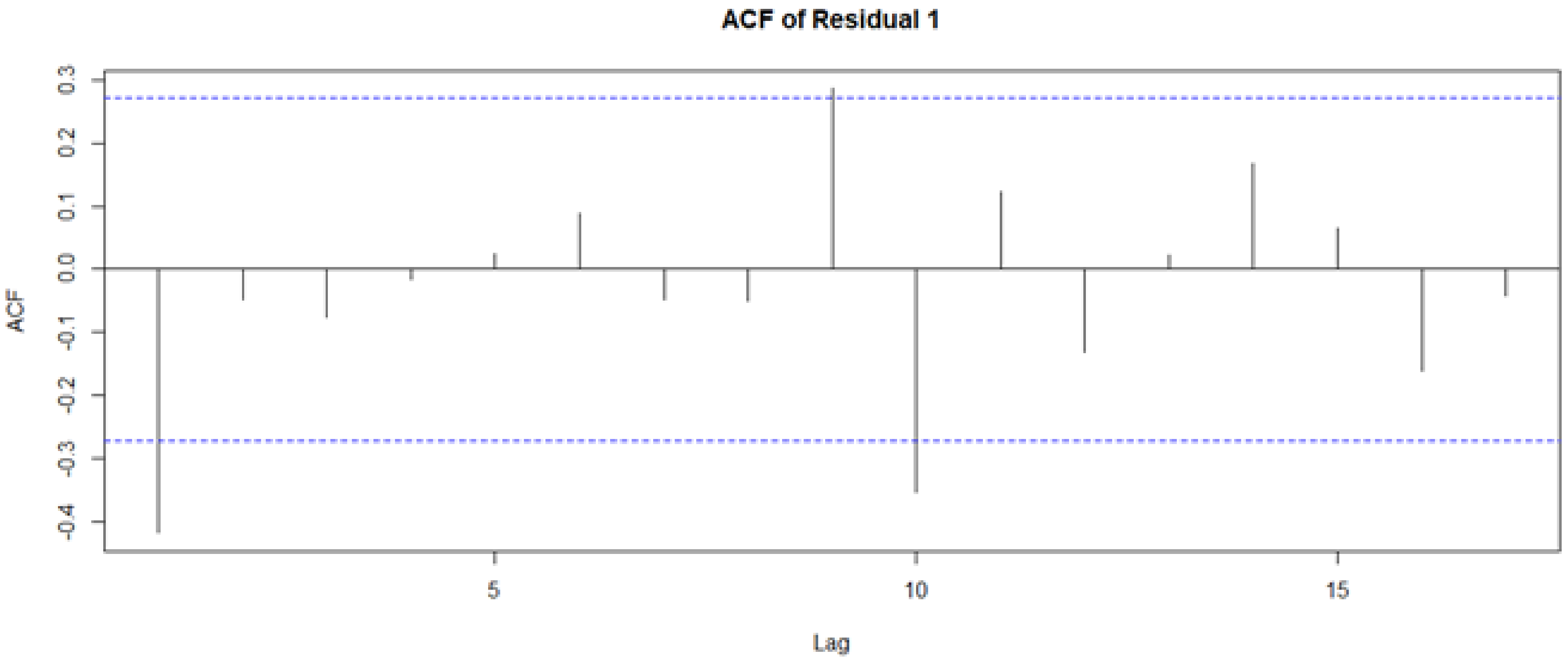
PARAMETER ESTIMATION & RESIDUAL ANALYSIS

```
> auto.arima(ClosePrice,ic="aic",trace = TRUE)

 Fitting models using approximations to speed things up...

 ARIMA(2,1,2)            with drift         : Inf
 ARIMA(0,1,0)            with drift         : 842.1547
 ARIMA(1,1,0)            with drift         : 844.882
 ARIMA(0,1,1)            with drift         : 843.8933
 ARIMA(0,1,0)                               : 844.2862
 ARIMA(1,1,1)            with drift         : Inf

 Now re-fitting the best model(s) without approximations...

 ARIMA(0,1,0)            with drift         : 855.7766

 Best model: ARIMA(0,1,0)            with drift

Series: ClosePrice
ARIMA(0,1,0) with drift

Coefficients:
         drift
      250.8644
s.e.  120.9769

sigma^2 estimated as 775963:  log likelihood=-425.89
AIC=855.78   AICc=856.02   BIC=859.68
```

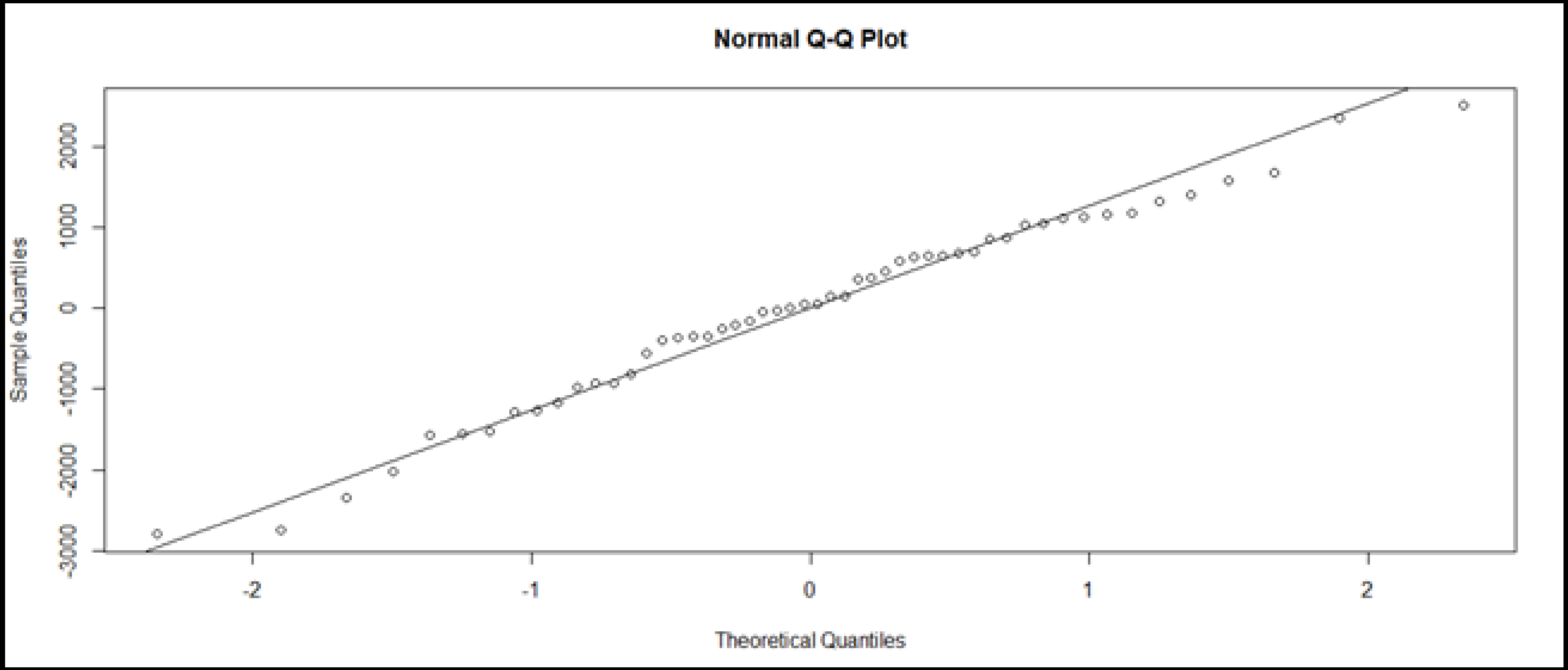
PARAMETER ESTIMATION & RESIDUAL ANALYSIS



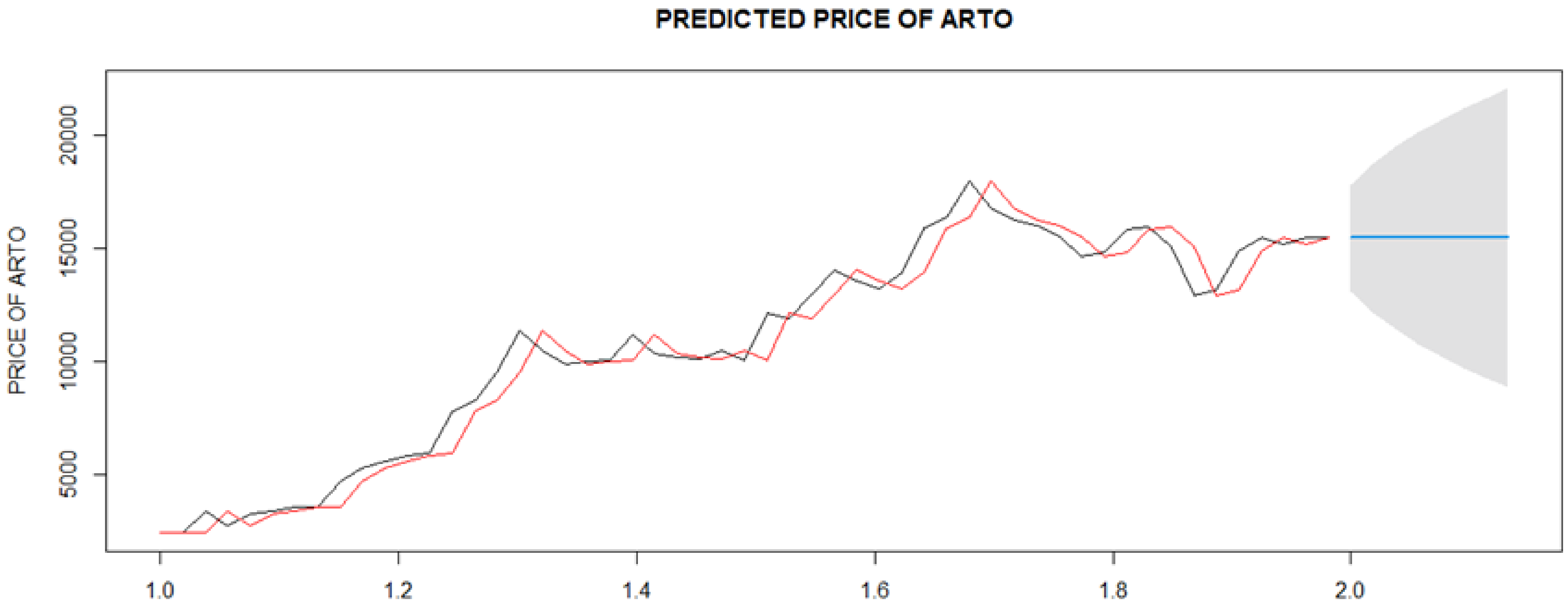
PARAMETER ESTIMATION & RESIDUAL ANALYSIS



PARAMETER ESTIMATION & RESIDUAL ANALYSIS



FORECASTING



FORECASTING

Date	Predicted Price	Lower Limit	Upper Limit	Actual Price
11/22/2021	15450	13111.843	17788.16	15,500.00
11/29/2021	15450	12143.347	18756.65	15,375.00
12/6/2021	15450	11400.193	19499.81	15,400.00
12/13/2021	15450	10773.686	20126.31	
12/20/2021	15450	10221.722	20678.28	
12/27/2021	15450	9722.709	21177.29	
1/3/2022	15450	9263.818	21636.18	
1/10/2022	15450	8836.693	22063.31	

FORECASTING

CONCLUSION

The conclusion of this project is that we use the Auto Arima function and get the ARIMA model (0,1,0) as the best model. This model also shows projections in the future that are not significantly down. This project is also expected to help make predictions of stock prices, especially PT Bank Jago Tbk. (ARTO)

If you can’t explain data simply, you don’t understand it well enough.

ALBERT EINSTEIN

THANKYOU !