Vol. 3 No. (2) (December 2025) 75-82

Journal of Data Insights

http://journalnew.unimus.ac.id/index.php/jodi

Forecasting Starbucks Stock Prices Using the ARIMA Method

Fellya Naza Nurcahyani [1], **Septiana Putri Milasari**[2], **Indah Manfaati Nur**[3]
[1,2,3]Department Statistics Study Program, Universitas Muhammadiyah Semarang, Indonesia

DOI: https://doi.org/10.26714/jodi.v3i2.309

Article Information

Article History:
Submitted: 15 January 2024
Revised: 20 November 2025
Approved: 20 December 2025

Keywords:
ARIMA; Saham; Starbucks.

Abstract

Starbucks is the largest coffee shop company in the world from the United States. This increase has become a trend in drinking coffee consumption among young people in a lifestyle while discussing. This indicates that the increase in the number of Starbucks stores is one of the drivers of Starbucks share prices among investors. Starbucks shares have the code SBUX as the issuer code. Starbucks Corporation is a coffee company and global coffeehouse chain. Satrbucks is an international company (MNCs) that anticipates various risks. The ARIMA forecasting method is different from other forecasting methods. This method uses an iterative approach to identify the most appropriate model from all possible existing models and this model can use all types of data. The ARIMA method was chosen for this research because this method is very suitable for short-term forecasting, where the products produced by the PT have a short expiration date. The result of the MAPE value is 3.218%, which means the accuracy is good because it is less than 10%.

✉ Corresponding :
E-mail: indahmnur@unimus.ac.id

e-ISSN: 2988 - 2109

INTRODUCTION

Starbucks is the largest coffeehouse company in the world and originates from the United States. The company opened its first store in Seattle, Washington. Over time, Starbucks has become a popular place for coffee enthusiasts of all ages—ranging from the young to the elderly—to relax and socialize. Its widespread popularity has driven rapid growth, reinforcing its position as the world's largest coffeehouse chain. In 2021, Starbucks operated 32,844 official stores worldwide (data.books). Indonesia ranks first in Southeast Asia, with 500 Starbucks stores spread across the country.

This growth also reflects the rising trend of coffee consumption among young people, especially as part of their lifestyle and social interactions. The increasing number of Starbucks outlets has contributed to greater investor interest, as store expansion is often associated with rising stock prices. However, in recent years, Starbucks' stock price has experienced a significant decline due to employee certification lawsuits and public reactions related to the Israel–Palestine conflict.

Forecasting is the process of estimating future needs in terms of quantity, quality, timing, and location to meet the demand for goods or services. Forecasting plays an essential role in helping predict future conditions. Based on its characteristics, forecasting techniques are generally classified into two main categories: qualitative forecasting and quantitative forecasting.

Starbucks stock is traded under the ticker symbol **SBUX**. Starbucks Corporation is a global coffee company and an international coffeehouse chain. As a multinational corporation (MNC), Starbucks faces a variety of risks. The company has not only transformed the business landscape but has also influenced American culture and impacted cultural trends around the world. In financial markets, stock prices fluctuate continuously due to supply and demand dynamics. These fluctuations require investors to analyze historical stock data before making investment decisions, as past data help in understanding potential price movements in future periods. Therefore, a forecasting model is needed to predict global Starbucks stock price fluctuations. According to Sari (2017), new investors are advised to consider Starbucks stock as a recommended investment because its forecasted price tends to remain relatively stable.

To address this issue, this study applies the **Autoregressive Integrated Moving Average (ARIMA)** method, which uses past and current values of the dependent variable to generate accurate short-term forecasts. ARIMA differs from other forecasting methods because it uses an iterative approach to identify the most suitable model from various possible alternatives, and it can be applied to many types of data. Time series models are particularly useful because future values are estimated based on historical patterns of the variable. ARIMA is chosen for this research because it is highly suitable for short-term forecasting, especially for data that exhibit short-term changes. The selected model is then validated using historical data to determine whether it accurately represents real conditions.

METHOD

1. Autoregressive (AR) Model

   The Autoregressive (AR) model describes a condition in which the dependent variable is influenced by its own past values at previous time periods (Sugiarto & Harijono, 2000). In general, the Autoregressive (AR) model can be expressed as follows (Zhao et al., 2022):

$$Y_t = \theta_0 + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \cdots + \theta_p Y_{t-p} - e_t$$

   Where::

Y_t	: Stationary time series
θ_0	: Constant
$Y_{t-1}, \ldots, Y_{t-p}$	: Relevant past values
$\theta_1, \ldots, \theta_p$	: Coefficients or parameters of the autoregressive model
e_t	: Residual at time

2. Moving Average (MA) Model

In general, the Moving Average (MA) model can be expressed as follows (Zhao et al., 2022):

$$Y_t = \phi_0 + \phi_1 e_{t-1} - \phi_2 e_{t-2} - \cdots - \phi_n e_{t-q}$$

Dimana:

Y_t : Stationary time series

ϕ_0 : Constant

ϕ_n : Coefficients of the Moving Average model, representing weighted values; these coefficients may be positive or negative depending on the estimation results.

e_t : Past residuals used in the model, with qqq indicating the number of residual lags included.

The main difference between the Moving Average model and the Autoregressive model lies in the type of independent variable used. In the Autoregressive model, the independent variable is the lagged values of the dependent variable. In contrast, the Moving Average model uses past residuals as its independent variables. The order of the MA model (denoted as qqq) is determined by the number of residual lags included in the model.

3. Autoregressive Moving Average (ARMA) Model

Often, the characteristics of YYY cannot be explained solely by an AR process or an MA process; instead, they must be explained by a combination of both. A model that incorporates these two processes is known as the Autoregressive Moving Average (ARMA) model. The general form of this model is as follows:

$$Y_t = \gamma_0 + \partial_1 Y_{t-1} + \partial_2 Y_{t-2} + \cdots + \partial_n Y_{t-p} - \lambda_1 e_{t-1} - \lambda_2 e_{t-2} - \lambda_n e_{t-q}$$

Where YtY_tYt and ete_tet have the same meaning as previously described, γt\gamma_tγt represents the constant term, and δ\deltaδ and λ\lambdaλ are the model coefficients. If the model uses two lagged dependent variables and three lagged residuals, it is denoted as ARMA(2,3).

4. *Autoregressive Integrated Moving Average* (ARIMA)

The Autoregressive Integrated Moving Average (ARIMA) model is a forecasting method that does not incorporate independent variables. ARIMA relies entirely on past and current values of the dependent variable to generate accurate short-term forecasts. The ARIMA modeling process consists of three fundamental stages: identification, parameter estimation and testing, and diagnostic checking.
The general form of the ARIMA model is as follows (de Araújo Morais & da Silva Gomes, 2022; Wen et al., 2023):

$$y't = C + \emptyset y_{t-1} + \cdots + \emptyset p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots . + \theta_1 \varepsilon_{t-q} + \varepsilon 1$$

5. Data Stationarity

It is important to note that most time series data are non-stationary, and that the AR and MA components of the ARIMA model apply only to stationary time series. Stationarity means that the data do not show any upward or downward trend. In general, the data should appear roughly horizontal along the time axis. In other words, the fluctuations in the data occur around a constant mean, do not depend on time, and the variance of these fluctuations remains essentially constant throughout the period.

A non-stationary time series must be transformed into a stationary one by applying differencing. Differencing is the process of calculating the change or difference between observed values.

In the Box–Jenkins methodology, the ARIMA model is written using the notation ARIMA(p, d, q), where ppp represents the order of the Autoregressive (AR) component, ddd represents the degree of differencing, and qqq represents the order of the Moving Average (MA) component (Harijono & Sugiarto, 2000).

6. Evaluasi Akurasi Peramalan

The accuracy of the forecasting results is evaluated using the Mean Absolute Percentage Error (MAPE), which calculates the average error rate across all test data. MAPE is used as a metric to measure the accuracy of the forecasting model. The formula used to calculate MAPE is as follows (Yamacli & Yamacli, 2023):

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} |PE_t|$$

MAPE Criteria

Table 1. Uji *Mean Absolute Percentage Error* (MAPE)

No	MAPE Value	Description
1	<10%	Excellent
2	10%-20%	Good
3	20%-50%	Fair
4	>50%	Poor

7. Data Source

This study uses secondary data obtained from the website *investing.com*, consisting of 133 observations recorded from May 1, 2023, to November 7, 2023. The data used in this research are global SBUX stock prices.

8. Metode

The data analysis procedures conducted to forecast SBUX stock prices using the ARIMA method are outlined as follows:

a. Describing the SBUX stock price data
b. Cleaning the dataset by removing any entries with zero values
c. Assessing the stationarity of the data
d. Conducting the Augmented Dickey–Fuller (ADF) test; if the p-value is greater than α (0.05), the data are classified as non-stationary and must be differenced until the p-value falls below α (0.05), indicating that the data have become stationary
e. Identifying the appropriate ARIMA model
Model identification is performed using the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots
f. Estimating the parameters of the model
The resulting ARIMA models are evaluated using the Akaike Information Criterion (AIC). The model with the lowest AIC and MAPE values is selected as the best-performing model
g. Conducting diagnostic testing
Diagnostic procedures include evaluating the residuals for normality using the Jarque–Bera test and testing for white noise (residual independence) using the Ljung–Box test

h. Forecasting SBUX stock prices for the next six periods

HASIL DAN PEMBAHASAN

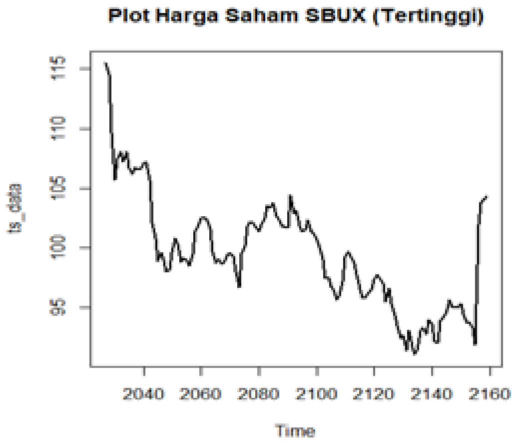

Figure 1. Plot of SBUX Stock Prices

The SBUX stock price series is not stationary, as it still exhibits a clear trend component. The plot shows an upward movement over time accompanied by fluctuations, indicating a stochastic trend pattern.

After identifying the pattern, the next step is to examine whether the data are stationary. If the data are non-stationary, differencing must be performed. If the data remain non-stationary after differencing, a transformation may be required. Stationarity can be assessed using the Augmented Dickey–Fuller (ADF) test. If the p-value is less than α (α = 0.05), the data can be considered stationary.

Table 1. Augmented Dickey–Fuller (ADF) Test

	Before diff	After diff
Dickey-fuller	-25,574	-51,063
lag-order	4	4
p-value	0,3455	0,01

The data are initially non-stationary, as indicated by the ADF p-value of 0.3455 prior to differencing. This value exceeds the significance level (α = 0.05), confirming that the series is not stationary. After differencing, the p-value decreases to 0.01, which is below α, indicating that the data have become stationary.

Table 2. Box–Ljung Test

Box-Ljung test	
x-squared	57,655
Df	1
p-value	0,0634

The p-value obtained from the Box–Ljung test is 0.06, which is greater than 0.05. This indicates that the data still exhibit autocorrelation.

Table 3. AIC Values

Model	AIC
ARIMA (1,1,1)	334,0810
ARIMA (0,1,1)	337,9141
ARIMA (1,1,0)	360,0744

The smallest AIC value is obtained from the ARIMA (1,1,1) model, with an AIC of 334.0810. Therefore, the ARIMA (1,1,1) model is selected as the most appropriate model for this analysis.

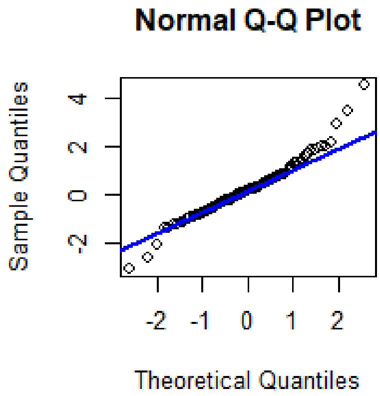

Figure 2. Normal Q–Q Plot

The data are normally distributed, as indicated by the Q–Q plot.

Table 4. Forecasted Values

No	Period	Forecasting
1	08/11/2023	100,713
2	09/11/2023	99,818
3	10/11/2023	99,596
4	13/11/2023	99,541
5	14/11/2023	99,528
6	15/11/2023	99,525
7	16/11/2023	99,524
8	17/11/2023	99,523
9	20/11/2023	99,523
10	21/11/2023	99,523
11	22/11/2023	99,523
12	23/11/2023	99,523

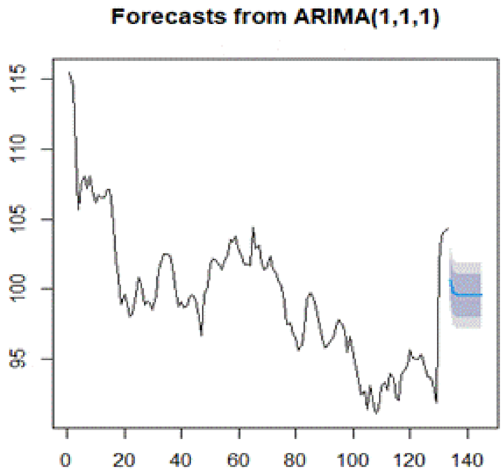

Figure 3. Forecast Plot of SBUX Stock Prices

The ARIMA (1,1,1) model forecasting results yielded an error accuracy value with a MAPE indicator of 3.218%. This value is less than 10%, indicating excellent forecasting capability using the ARIMA (1,1,1) model.

CONCLUSION

Based on the analysis conducted, it can be concluded that the SBUX stock price data from May 1, 2023, to November 7, 2023, exhibit noticeable fluctuations. The stock prices show a trend pattern prior to differencing. The ARIMA (1,1,1) model is identified as the best model for forecasting SBUX stock prices, as it has the lowest AIC value and produces a forecasting error rate of less than 10%

DAFTAR PUSTAKA

Abror Gustiansyah, M., Rizki, A., Apriyanti, B., Maulidia, K., Julio Roger Roa, R., Al Hadi, O., Ghoni Trisno Hidayatulloh, N., Andriyani Lestari Ningsih, W., Putri Ratnasari, A., & Angraini, Y. (2023). Aplikasi Model ARIMA dalam Peramalan Data Harga Emas Dunia Tahun 2010-2022. Jurnal Statistika Dan Aplikasinya, 7(1).

Buchori, M., & Sukmono, T. (2018). Peramalan Produksi Menggunakan Metode Autoregressive Integrated Moving Average (ARIMA) di PT. XYZ. PROZIMA (Productivity, Optimization and Manufacturing System Engineering), 2(1), 27–33. https://doi.org/10.21070/prozima.v2i1.1290

de Araújo Morais, L. R., & da Silva Gomes, G. S. (2022). Forecasting daily Covid-19 cases in the world with a hybrid ARIMA and neural network model. Applied Soft Computing, 126. https://doi.org/10.1016/j.asoc.2022.109315

Fattah, H. A. H. (2014). Risiko Bisnis Yang Dihadapi Starbucks. In Jurnal Media Wahana Ekonomika (Vol. 10, Issue 4). http://www.swa.co.id/sekunder/resensi/pemasaran/

Sastri Nenda, M., Harto, S., Riau Kampus Bina Widya Jl R Soebrantas Km, U. H., & Baru Pekanbaru, S. (2017). Dampak Investasi Starbucks Coffee Terhadap Penjualanproduksi Kopi Lokal Di Indonesia. In JOM FISIP (Vol. 4, Issue 1).

Schima Wulandari, S., & Yurinanda, S. (2021). Penerapan Metode ARIMA Dalam Memprediksi Fluktuasi Harga Saham PT Bank Central Asia Tbk. www.finance.yahoo.com.

Wen, T., Liu, Y., Bai, Y. he, & Liu, H. (2023). Modeling and forecasting CO2 emissions in China and its regions using a novel ARIMA-LSTM model. Heliyon, 9(11). https://doi.org/10.1016/j.heliyon.2023.e21241

Yamacli, D. S., & Yamacli, S. (2023). Estimation of the unemployment rate in Turkey: A comparison of the ARIMA and machine learning models including Covid-19 pandemic periods. Heliyon, 9(1). https://doi.org/10.1016/j.heliyon.2023.e12796

Zhao, L., Li, Z., & Qu, L. (2022). Forecasting of Beijing PM2.5 with a hybrid ARIMA model based on integrated AIC and improved GS fixed-order methods and seasonal decomposition. Heliyon, 8(12). https://doi.org/10.1016/j.heliyon.2022.e12239