

Analyze the Future Stock Trend of Pop Mart Based on the Time Series Model

Yiran Sha

*Mathematic Science Academy, University of Nottingham, Nottingham, United Kingdom
smyys13@nottingham.edu.cn*

Abstract. With the increasing sophistication of financial management techniques, the ability to predict stock prices has become a critical endeavor for investors and traders. The Autoregressive Integrated Moving Average (ARIMA) model stands out as a widely used method for forecasting stock price movements. This model, which belongs to the class of time series analysis, is particularly adept at estimating daily returns with associated confidence intervals. The present study aims to apply the ARIMA model to forecast the share prices of Pop Mart. The dataset encompasses the period from December 2024 to May 2025, focusing on the opening prices as key variables. The analysis reveals that the ARIMA model is capable of capturing the fluctuations in stock prices with a certain degree of accuracy. The Mean Absolute Percentage Error (MAPE) value of 6.86%, which is significantly below 10%, suggests that the predictive performance is commendable. Furthermore, the insights gleaned from this model offer investors and traders valuable guidance for interpreting market trends and assessing potential risks, thereby enhancing their decision-making capabilities.

Keywords: Time series analysis, stock price prediction, Autoregressive Integrated Moving Average (ARIMA) model

1. Introduction

The financial and stock markets are characterized by high volatility and unpredictability, posing significant challenges for investors in forecasting stock trends and prices. Precise forecasts are essential for enhancing profits and reducing risks, especially in both short-term trading and long-term investment plans. However, market volatility, influenced by economic, political, and psychological elements, continues to pose challenges to decision-making, even with improvements in analytical techniques. This underscores the necessity of developing robust predictive models to enhance market analysis. The ability to estimate future stock values with precision offers substantial economic benefits, making stock market forecasting a critical area of research for both academia and practitioners.

While numerous studies have explored stock market prediction, many traditional models fail to fully account for non-linear trends and external uncertainties, limiting their accuracy. The Autoregressive Integrated Moving Average (ARIMA) model. The ARIMA model includes autoregressive (AR) model, moving average (MA) model, and seasonal autoregressive integrated

moving average (SARIMA) model, widely used in time-series forecasting [1]. More precisely, forecasting in the realm of time series entails leveraging a particular model to predict forthcoming values of a variable that encapsulates aspects of a phenomenon, grounded in extant historical data [2]. When analyzing specific stocks of fast-growing consumer brands like Pop Mart, they are still not fully utilized. This study addresses this gap by applying the ARIMA model to predict Pop Mart's stock trends, combining theoretical rigor with practical relevance. The objectives include elucidating the ARIMA framework, detailing its construction and validation steps, and demonstrating its efficacy through a case study. Identifying the best model among the various combinations of non-seasonal and seasonal orders is crucial for achieving good forecasting performance with ARIMA models, given the rich variation they provide [3].

The paper is structured as follows: First, it provides a comprehensive overview of the ARIMA model and its applicability to stock market analysis. Next, it outlines the methodology, including data preprocessing, model selection, and accuracy verification. Finally, a case study of Pop Mart's stock is presented to validate the model's predictive performance, followed by conclusions and implications for investors.

2. Time series model

The order of integration, which indicates the number of differencing operations needed to achieve stationarity in a time series, is what the term "integrated" refers to [4]. Within this framework, the forthcoming value in a time series data set X_t is delineated mathematically by

$$(1 - \sum_{i=1}^p \varphi_i B^i)(1 - B)^d X_t = \left(1 + \sum_{j=1}^q \theta_j B^j\right) \epsilon_t \quad (1)$$

where B is the lag operator, p and q are the parameters of the autoregressive and moving average ARIMA (p, d, q), with d representing the order of integration.

3. Model building

3.1. Data pretreatment

Prior to constructing the model, it is essential to assess whether the data is white noise and whether it exhibits stability. If the data is identified as a white noise series, it holds little significance for the research. Conversely, if the data is unstable, the final results may significantly deviate from the actual situation. To address this, the data must be differenced until it achieves stability [5]. Following differencing, the new dataset will undergo the Augmented Dickey-Fuller (ADF) test to confirm its stability. Should the ADF test yield a p-value greater than 0.05, indicating that the data remains unstable, further differencing and retesting are required. The parameter d in the integrated (I) component represents the number of differencing operations performed. Additionally, the Box-Ljung test will be employed to check for white noise. If the test result is less than 0.05, the null hypothesis will be rejected, suggesting that the time series is not white noise. Only after these checks are passed can the study proceed; otherwise, the time series is deemed unsuitable for analysis.

3.2. Estimation of model parameters

The initial estimation of the potential values for the parameters p and q in an ARIMA model can be done using the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF). p is used for Autoregressive, d is used for to make the time series data stationary, q is used for Moving Average, ARIMA is able to forecast stationary time series data [6]. For a more precise determination, the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are employed. The goal is to choose the model with the smallest AIC or BIC value. After determining the values for p , d , and q , the model's parameters will be estimated using Ordinary Least Squares (OLS) and Maximum Likelihood Estimation (MLE) methods.

3.3. Model testing and predicting

Once the predictive model has been constructed, it is imperative to evaluate the residuals through a white noise test and an overfitting test. These assessments are crucial to verify the model's efficacy and to confirm that the residuals adhere to white noise characteristics, which are essential for reliable forecasting. Additionally, upon finalizing the predictive model, the Mean Absolute Percentage Error (MAPE) metric will be implemented to assess its predictive accuracy.

4. Application

4.1. Target companies and time series charts

The article will analyze the publicly available stock price data of Pop Mart, spanning the period from December 2024 to May 2025. Utilizing R software, a time series plot of the daily opening prices for POP MART will be generated. Upon examination of the figure (referred to as Figure 1), it is evident that the data exhibits minor fluctuations and predominantly follows a stable upward trajectory.

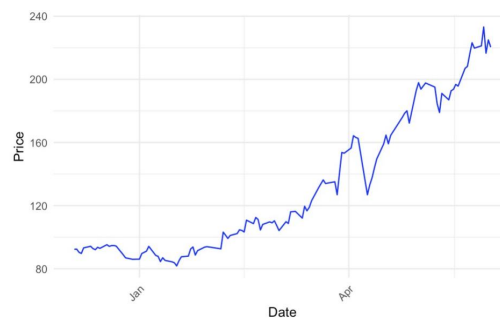


Figure 1. POP MART stock opening price chart from December 2024 to May 2025

4.2. Stability testing and differentiation

In order to achieve stationarity in the original time series data, differentiation is applied, followed by the ADF test to ascertain the smoothness of the resulting series. The data satisfies the stationarity condition, as evidenced by an ADF test p-value falling below the threshold of 0.05 after the application of two differencing operations. Concurrently, the white noise test produces a p-value that is less than 0.05, signifying that the data does not conform to a white noise process. This finding

validates the non-white noise characteristic of the series, thereby justifying the continuation of the analysis.

4.3. Model building

Before making predictions, this paper needs to first determine the various parameters of the model. The first step is to determine the p and q values in the model. Since ARIMA model requires time series to be stationary, ADF test was conducted to assess the stationarity of data [7]. Time series stationarity can be effectively assessed using ACF and PACF plots. A slow decay in the ACF plot indicates that the time series is non-stationary [8]. Based on the ACF and PACF methods, this paper can preliminarily determine the approximate range of p and q .

From the Figures 2 and 3, preliminary estimates for the parameters p and q can be deduced, with both approximately equal to 1. This inference is drawn from the observation that the coefficients at the first lag are statistically significant, as they lie outside the confidence intervals. In contrast, the coefficients at higher lags are all contained within their respective confidence intervals, indicating their lack of statistical significance. Typically, the ACF and PACF of a model manifest in one of two conditions: they either trail off or are truncated. When the ACF or PACF drops to zero following a specific lag—denoted as p or q —it is characterized as truncated. Conversely, when these functions exhibit a geometric decline toward zero as the lag order k ascends, this is referred to as a trailing pattern [9].

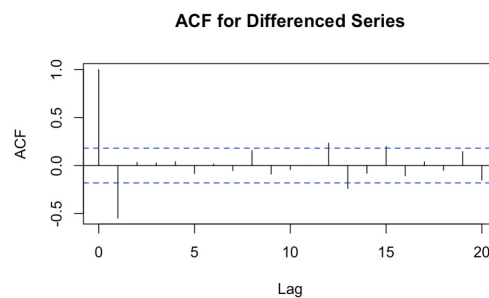


Figure 2. ACF plot of second-order differencing data

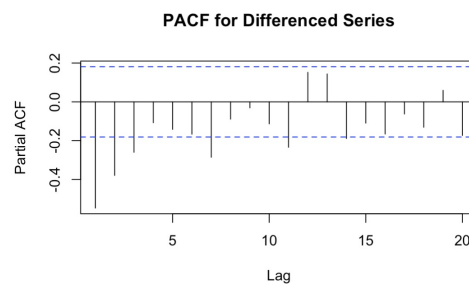


Figure 3. PACF plot of second-order differencing data

According to the AIC and BIC methods, in this paper, the values of p and q are both set between 0 and 5, and then AIC and BIC are selected to achieve the minimum combination simultaneously. Therefore, it was finally determined to use the ARIMA (0,2,1) model. After constructing the model, this paper needs to test the residual terms of the model to verify whether the residual terms are white noise. If the residual terms are white noise, it indicates that the constructed model is reasonable;

otherwise, further adjustments are required. From Table 1, it can be clearly seen that all p-values are greater than 0.05, indicating that the residual terms are white noise and the model can be used for final prediction.

Table 1. Residual test of the model

Lagging Order	X-squared	df	p-value
6	6.51	6	0.39
12	13.19	12	0.36
18	27.87	18	0.06

Finally, to verify whether the order of the model is reasonable, this paper will conduct an overfitting test on the model to avoid bias in the final results due to an excessively high order. As shown in Table 2, the result of the overfitting test is 0.00, which is less than 0.05. Therefore, the model does not have an overfitting problem, and the final prediction can be made.

Table 2. Overfitting test of the model

Parameter estimate	s.e.	p-value
-0.99	0.02	0.00

4.4. Model forecast

Subsequently, the model is employed for forecasting purposes. In this study, predictions for the opening stock prices of Pop Mart for the subsequent 7 days are derived (as shown in Figure 4). In Figure 4, the light grey region delineates the 80% confidence interval, while the blue-grey area represents the 95% confidence interval.

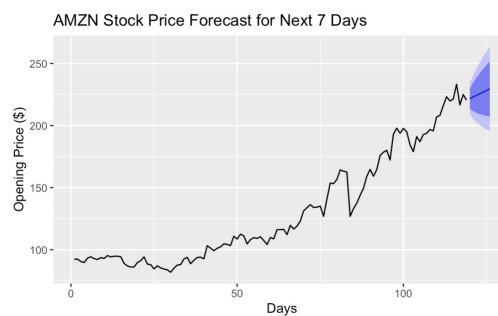


Figure 4. Graph of model forecasting results

Then Table 3 lists the specific prediction data and results under 80% and 95% confidence intervals.

Table 3. POP MART's stock forecast value and confidence interval for the next 7 days

Date	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2025.06.03	221.70	213.65	229.76	209.38	234.02
2025.06.04	223.00	211.53	234.48	205.45	240.55
2025.06.05	224.30	210.14	238.46	202.65	245.96
2025.06.06	225.60	209.13	242.08	200.41	250.80
2025.06.09	226.91	208.35	245.46	198.53	255.28
2025.06.10	228.21	207.73	248.68	196.89	259.52
2025.06.11	229.51	207.23	251.78	195.44	263.58

Subsequently, this article collects the actual values of the company's stocks on the corresponding dates and calculates the error between the predicted values and the actual values, as shown in Table 4. It can be seen that the MAPE value is 8.29%, which is less than 10%. Therefore, it indicates that the predictions are relatively accurate [10]. At the same time, according to the model's predictions, the stock price of Bubble Code Company will show an upward trend in the future, and the actual results also show an upward trend. This shows that the results obtained by the model have high reference value.

Table 4. Analysis of forecast results

Date	Ture value	Forecast value	Absolute error	Relative error	MAPE
2025.06.03	231.00	221.70	9.30	0.04	8.29%
2025.06.04	234.00	223.00	11.00	0.05	
2025.06.05	249.00	224.30	24.70	0.10	
2025.06.06	243.2	225.60	17.60	0.07	
2025.06.09	251.20	226.91	24.29	0.10	
2025.06.10	253.00	228.21	24.79	0.10	
2025.06.11	261.60	229.51	32.09	0.12	

5. Conclusion

The stock market is crucial to many aspects of economic activity. It serves as an efficient mechanism for the allocation of capital to sectors and enterprises that demonstrate superior performance and growth prospects, thereby facilitating optimal resource distribution. Furthermore, the equity market aids businesses in expanding their operational scale, which in turn generates additional employment opportunities. This expansion can lead to the development of vertical and horizontal industry linkages, fostering the formation of industrial clusters and contributing to the overall economic expansion. The stock market also functions as a barometer of macroeconomic health, with stock price movements providing insights into broader economic trends. For both corporate entities and investors, the stock market is of significant relevance, underscoring the importance of accurate stock market forecasting. This study employs a time series analysis to construct an ARIMA (p, d, q) model for predictive purposes. Focusing on the stock performance of Pop Mart from December 2024 to May 2025, an ARIMA (0,2,1) model was developed to forecast the opening prices for the

subsequent five days. The model achieved a MAPE of 8.29%, which is considerably below the 10% threshold, indicating a high level of forecasting accuracy.

This research represents a foundational exploration, yet it acknowledges the complexity and volatility inherent in the securities market, suggesting that future studies should address additional considerations. Firstly, the study's scope is limited to the opening prices, neglecting other pertinent variables such as closing prices, high and low prices, and trading volume. Moreover, external factors that can substantially influence stock prices, especially in relation to industry trends, were not taken into account. The omission of these elements may hinder the precise capture of price dynamics driven by trading volume fluctuations. To enhance the precision of predictions, it is recommended to integrate this model with other methodologies. For example, Long Short-Term Memory (LSTM) networks. The LSTM's robust capacity for non-linear modeling, when combined with the ARIMA model's strength in linear trend fitting, can provide a more comprehensive depiction of stock price patterns, thereby improving the accuracy of forecasts.

References

- [1] Al-Musaylh, M.S., Deo, R.C., Adamowski, J.F. and Li, Y. (2018) Short-Term Electricity Demand Forecasting with MARS, SVR and ARIMA Models using Aggregated Demand Data in Queensland, Australia. *Advanced Engineering Informatics*, 35, 1-16.
- [2] Liapis, C.M., Karanikola, A. and Kotsiantis, S. (2021) A Multi-Method Survey on the use of Sentiment Analysis in Multivariate Financial Time Series Forecasting. *Entropy (Basel, Switzerland)*, 23, 1603.
- [3] Wang, X., Kang, Y., Hyndman, R.J. and Li, F. (2023) Distributed ARIMA Models for Ultra-long Time Series. *International Journal of Forecasting*, 39, 1163-1184.
- [4] Pokou, F., Sadefo Kamdem, J. and Benhmad, F. (2024) Hybridization of ARIMA with Learning Models for Forecasting of Stock Market Time Series. *Computational Economics*, 63, 1349-1399.
- [5] Funde, Y. and Damani, A. (2023) Comparison of ARIMA and Exponential Smoothing Models in Prediction of Stock Prices. *Journal of Prediction Markets*, 17, 21–38.
- [6] Khan, K. (2022) Stock Price Forecasting of Maruti Suzuki using ARIMA Model. *Parikalpana: KIIT Journal of Management*, 18, 90–98.
- [7] Erker, E. (2024) Forecasting Medical inflation in the European Union using the ARIMA model. *Public Sector Economics*, 48, 39–56.
- [8] Singh, S., Parmar, K. S. and Kumar, J. (2021) Soft Computing Model Coupled with Statistical Models to Estimate Future of Stock Market. *Neural Computing & Applications*, 33, 7629–7647.
- [9] Ning, L., Pei, L., Li, F. and Xie, L. (2021) Forecast of China's Carbon Emissions Based on ARIMA Method. *Discrete Dynamics in Nature and Society*, 2021, 1–12.
- [10] Dun, V. (2023) Analysis and Forecasting the Price of the S&P 500 Index Using the Arima Model. *Financial Markets, Institutions and Risks*, 7, 113–134.