International Journal of Scientific Research & Engineering Trends
Volume 11, Issue 5, Sep-Oct-2025, ISSN (Online): 2395-566X

Predicting Stock Market Trends with ARIMA: A Data-Centric Approach to the BSE Index

[1]Kalyan Kumar Bethu, [2]Dr.M.Sravani
[1]Scholar Krishna University
[2]Assistant Professor Krishna University

Abstract - Stock market volatility makes accurate forecasting vital for informed trading decisions and profit maximization. Over the years, various models have been introduced to enhance the reliability of time series predictions. This study applies the ARIMA model to evaluate data stability and forecast movements in the BSE Index. Model selection was guided by statistical measures including SIGMASQ, Adjusted R^2, AIC, and BIC, with ARIMA (2,1,2) emerging as the most suitable specification. Using monthly data from January 2021 to January 2025 (49 observations), the model generated forecasts for February 2025 through December 2025, yielding 11 projected values. The results highlight ARIMA's effectiveness as a short-term forecasting tool, offering actionable insights for informed investment decisions.

Keywords - BSE_SENSEX_CLOSE, ARIMA model, AIC, BIC Forecasting,

INTRODUCTION

A vital aspect of stock market forecasting is the prediction of future index values, such as SENSEX and NIFTY, which play a key role in developing profitable investment strategies and guiding business decisions. Accurate predictions enable investors, market regulators, and policymakers to better manage market fluctuations and maximize returns. Forecasting approaches are broadly classified into two categories: linear methods, including time series regression and moving averages, and nonlinear methods that capture more complex market dynamics. Among these, the ARIMA model—introduced by Box and Jenkins in 1976— remains one of the most influential techniques for stock market forecasting. This study aims to determine the optimal ARIMA model through rigorous statistical evaluation, thereby improving the accuracy of market trend predictions and strengthening strategic financial decision-making.

II. REVIEW OF LITERATURE

In time series forecasting, ARIMA models have been widely applied to predict economic variables such as exchange rates, power consumption, inflation, oil palm prices, and gold prices. However, research focused on forecasting stock prices and market indices remains relatively limited. This study seeks to address this gap by identifying the optimal ARIMA model for predicting future stock prices and indices. By concentrating on these crucial financial measures, the research aims to improve forecasting accuracy and reliability while providing fresh insights to assist investors and financial analysts in making strategic decisions.

C.V. (2019): analysed the BSE and NSE stock indices over a 22-week period (January–June 2018) using weekly data from January 2017 to December 2018. The study concluded that ARIMA (0,1,1) and ARIMA (1,0,0) provided the best fit, highlighting ARIMA's effectiveness for short-term forecasting and its value in portfolio decision-making.

Suprabha, K.R. & Arjun, R. (2020): conducted an extensive study of stock index forecasting on the NSE and BSE between 2004 and 2016, incorporating both technical and operational factors. Their findings revealed that while linear models performed better in predicting BSE trends, Artificial Neural Networks (ANNs) captured NSE variability more effectively, thereby strengthening prediction models and enhancing business intelligence.

Dadhich, Manvinder Singh, Pahwa, Ruchi, Doshi, M.Sc., Dadhich & Vipin Jain (2021): employed econometric forecasting techniques to evaluate the predictability of major Indian stock indices. Using ARIMA models, the best fits identified were ARIMA (0,1,4) for DLOG_NSE and ARIMA (3,1,2) for DLOG_BSE. Stationarity was confirmed using the Augmented Dickey-Fuller (ADF) test. Based on 474 variables, the study demonstrated that ARIMA models could accurately forecast daily stock prices between March 1 and March 28, 2018, underscoring the advantage of daily data over weekly or monthly patterns for optimal portfolio strategies.

Murali, K., Omar, S.N., Naidu, D.C., Reddy, M.R., Mani, C.K., and Murthy, B.R. (2021): examined ARIMA applications in predicting stock buy-and-sell signals. Their results validated the use of ARIMA for short-term forecasting and emphasized its practical utility for investors and financial strategists.

H.R., T.S., et al., and V., J.B. (2023): compared ARIMA with the Generalized Autoregressive Conditional Heteroskedasticity (GARCH (1,1)) model to forecast the BSE 30 (Sensex) and NSE 50 (Nifty) indices. Using PP, ADF, and KPSS tests to ensure stationarity, their study revealed that the GARCH (1,1) model outperformed ARIMA, achieving nearly twice the accuracy. These findings highlight the importance of selecting advanced models for reliable forecasting, thereby aiding investors, analysts, and financial institutions in making more effective market predictions.

.

III. STATEMENT OF THE PROBLEM

Forecasting in a market economy is inherently challenging due to its unpredictable and volatile nature. Nevertheless, businesses and entrepreneurs must develop strategies to manage these risks effectively. To address this need, researchers have introduced increasingly precise forecasting models. The accuracy of such models—including regression analysis, Box–Jenkins methods, smoothing techniques, and classical classification approaches—is typically evaluated using statistical measures such as Root Mean Square Error (RMSE) and Adjusted R-Squared (R^2). Among these approaches, the Box–Jenkins ARIMA model stands out for its adaptability and predictive accuracy. This study focuses on applying ARIMA to time series data from the BSE in order to identify the most suitable model for short-term forecasting. The findings are expected to provide valuable insights for investors, supporting more informed and strategic investment decisions.

Objective

- To ensure the consistency and reliability of monthly BSE SENSEX closing values and other time series data.
- To identify the most suitable ARIMA model for accurate time series forecasting.
- To apply the selected ARIMA model for predicting future movements of the BSE Index.

Research Methodology

- Research Design: The study used an exploratory research strategy with contingency modelling to reach its goals. Secondary data is mostly used in this type of study to look at and judge current material. We used stochastic modelling to find the best ARIMA model and then used that model to predict the time series.
- Data Source: The data for this study came from BSE INDIA, which collected monthly closing indices for the BSE_SENSEX from January 2021 to February 2025.We will use this dataset, which has 49 observations and 11 projections, to guess prices from February 2025 to December 2025.

Statement of Hypothesis

- For unit root tests: Null hypothesis (H0): $\delta = 1$, which implies a unit root of the series (BSE_SENSEX_CLOSE) and nonstationary. Alternative hypothesis (Here): $\delta < 1$, the series (BSE_SENSEX_CLOSE) does not have a unit root and does not show stationarity (or trend stationarity).

Methods of Data Analysis

When choosing the best ARIMA model based on experimental data, there are a number of important statistics that need to be considered. The Bayesian Information Criterion (BIC) adjusted R-squared (R2), Akaike Information Criterion (AIC), volatility (sigma2), and the significance of the coefficients are all examples of these. Each of these indicators gives us useful information about how complicated the model is, how well it explains the data, and how well it fits the data. By looking at these variables, we can figure out the best ARIMA setup for making accurate predictions.

Augmented Dickey Fuller Unit Root Test Constructing an ARIMA model begins with determining if the target variable is stationary across time. It is implied behavior that is unpredictable might result from non- stationarity, which is frequently brought on by unit roots. Transforming non-stationary data into a stationary form is necessary to guarantee accurate forecasting. A popular technique for determining stationarity and identifying the existence of a unit root in a random walk model is the Augmented Dickey- Fuller (ADF) test. The calculated t-statistic is compared with critical values in the test. Stationarity is confirmed if the p- value is less than 0.05, which rejects the null hypothesis of non- stationarity. If not, other actions might be needed to reach stationarity, like applying the first or second differences.

Table I. Results of The Augmented Dickey-Fuller (Adf) Test

Section	Category	Level Form	First Difference
ADF Test Summary	Test Description	Baseline Level Form	First Difference
	Null Hypothesis (H_0)	The unit root of BSE SENSEX CLOSE is non-stationary	A unit root exists for D(BSE SENSEX CLOSE)
	Exogenous Variable	Constant	Constant
	Lag Length	0 (Automatic - SIC, max=10)	0 (Automatic - SIC, max=10)
	ADF Test Statistic	-1.125673	-6.972752
	Critical Value (1%)	-3.574446	-3.577723
	Critical Value (5%)	-2.92378	-2.925169
	Critical Value (10%)	-2.599925	-2.600658
	p-value	0.8982	0
	Conclusion	Fail to reject H_0 (Non- Stationary)	Reject H_0 (Stationary at 1st Difference)
Equation Coefficients	BSE SENSEX CLOSE(-1)	-0.039741 (0.032839)	-1.031904 (0.147991)
	t-Statistic (BSE)	-1.125673	-6.972752
	p-value (BSE)	0.2861	0
	Constant (C)	2974.976 (2089.708)	625.8900 (342.3324)
	t-Statistic (Constant)	1.423384	1.828311
	p-value (Constant)	0.1618	0.0744
Model Summary	R-squared	0.028608	0.51933
	Modified R- squared	0.006582	0.508648
	Regression Std. Error	2213.58	2243.925

International Journal of Scientific Research & Engineering Trends
Volume 11, Issue 5, Sep-Oct-2025, ISSN (Online): 2395-566X

	Sum of Squared Residuals	2.25E+08	2.27E+08
	Log Likelihood (Log Prob.)	-418.18	-428.3194

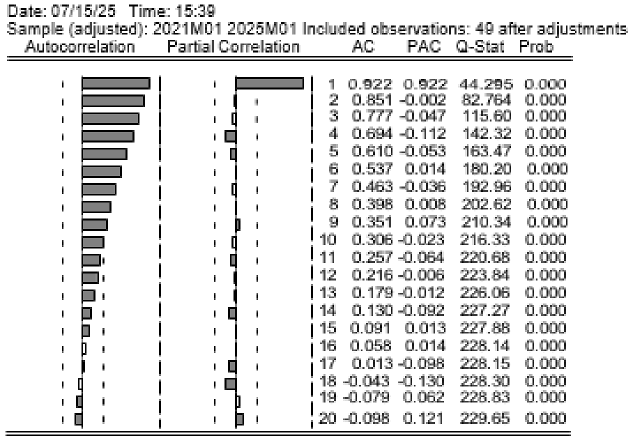
Table 1: The null hypothesis cannot be rejected because the p-value (0.6982) > 0.05 and the ADF test statistic (- 1.125673) are both greater than any critical values (in absolute terms).

This demonstrates that at the baseline level, the BSE_SENSEX_CLOSE series is not stationary. Table 2: With a p-value of 0.0000, the ADF test statistic (-6.972752) is far below the critical values (in absolute terms) following the first difference. As a result, the null hypothesis is rejected, proving by stationarity that statistical parameters like mean and autocovariance don't change over time.
Data the stationary nature of the first-differenced BSE_SENSEX_CLOSE series.

Analysis of Correlogram

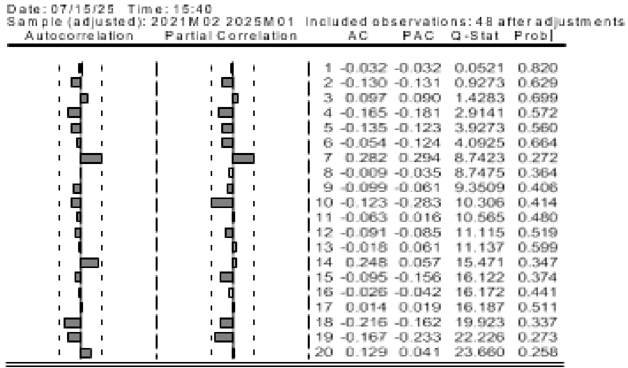
A Correlogram, additionally called an Autocorrelation Function (ACF) plot, visually represents correlation facts that summarize the correlation at extraordinary time intervals, indicating serial correlation in a time series. Serial correlation happens whilst an blunders at one factor in time impacts next points in time, highlighting styles and dependencies inside the data.

The Correlogram generally consists of both the Autocorrelation Function (ACF), which measures how observations in a time collection relate to every different across exclusive lags, and the Partial Autocorrelation Function (PACF), which isolates the direct courting among observations at unique lags after putting off the consequences of intervening observations (Brooks, 2008; Abdullah, 2012).
Date: 07/15/25 Time: 15:39 Sample (adjusted): 2021M01 2025M01 Included observations: 49 after adjustments Autocorrelation Partial Correlation AC PAC Q-Stat Prob



Source :Authors calculation

Fig.1. Correlogram for BSE_SENSEX _CLOSE.



Source: Authors calculation

Fig.2.: First Order Difference Correlogram for BSE SENSEX CLOSE

There is non-stationarity in the BSE_SENSEX_CLOSE time series, as seen by the strong autocorrelations and notable spikes in the ACF and PACF plots prior to differencing (Figure 1). This is also supported by the Q-statistics with p- values of 0.000. The series becomes stationary after first-order differencing is applied (Figure 2), as the ACF and PACF spikes fall within the 95% confidence limits and the majority of p-values are above 0.05. Given that the residuals resemble white noise and show no discernible autocorrelations, the ARIMA model appears to fit the data well. As a result, the model works

International Journal of Scientific Research & Engineering Trends
Volume 11, Issue 5, Sep-Oct-2025, ISSN (Online): 2395-566X

well, and the time series is effectively converted for forecasting.

Arima Model for Forecasting

ARIMA models work well for short-term economic forecasting, providing valuable insights and predictability (Moreh, Saxena, & Pardasani, 2010; Pie & Senglin, 2005). The ARIMA model is denoted as ARIMA (p, d, q), where p is the autoregressive term, d is the differencing order, and q is the moving average component. In our investigation, the BSE_SENSEX_CLOSE time series became stationary after the first difference (d = 1), making it acceptable for ARIMA (p, 1, q).

To calculate p and q, we used the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF). The lack of substantial peaks indicated that high-order AR (p) and MA (q) terms were superfluous.

The best ARIMA model was chosen based on Akaike (AIC), Bayesian (BIC), and Schwarz (SIC) Information Criteria (minimization technique).

The smallest standard error of regression Low SIGMASQ values, High adjusted R^2.

Table 2: Comparative Analysis of ARIMA Parameters for D(BSE_SENSEX_CLOSE)

ARIMA	SIGMASQ	Adjusted R^2	AIC	SBIC	S.E. of Regression
(1,1,0)	48,19,946	-0.04335	18.35117	18.49567	2,232.24
(0,1,1)	48,18,167	-0.04297	18.35082	18.46777	2,109.96
(1,1,1)	45,11,432	0.001228	18.33515	18.49008	2,256.80
(1,1,2)	46,68,718	0.033583	18.36347	18.51941	2,218.46
(2,1,1)	47,36,424	-0.04857	18.37606	18.53199	2,248.99
(2,1,0)	47,39,858	-0.02601	18.33511	18.45206	2,267.44
(0,1,2)	46,71,471	-0.01121	18.32242	18.43937	2,267.47
(2,1,2)	40,80,923	0.095646	18.26308	18.41901	2,232.24

Source: Authors calculation

Table 2 displays the outcomes of various AR(p) and MA(q) parameters in the ARIMA model, which can be used to find the best settings for forecasting D(BSE_SENSEX_CLOSING). The examination considered many statistical metrics, such as the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Sigmund-Sigma Quotation (SQ), modified R^2, and regression standard error. A thorough examination of these characteristics reveals that the ARIMA model (2,1,2) is the best reliable predictor of future values. This model demonstrates that the best fit with little complexity has the lowest AIC (18.26308) and BIC (18.41901) values. •The lowest SIGMASQ (4080923) increases prediction reliability by presenting little residual variance. The highest Adjusted R2 (0.096546) means that, when compared to other examined models, the model accounts for the greatest percentage of the variance in the dependent variable. These results give solid evidence that ARIMA (2,1,2) is the best forecasting approach for predicting future changes in the BSE_SENSEX_CLOSE index.

Prediction of Bse_Sensex_Close The Usage Of The Arima (2,1,2) Version.

Table 3: Prediction of Bse_Sensex_Close

Month	Forecasted BSE_SENSEX_CLOSE
2025M02	79498.91
2025M03	79555.98
2025M04	81020
2025M05	81318.31
2025M06	82461.49
2025M07	82904.67

2025M08	83855.17
2025M09	84385.35
2025M10	85220.16
2025M11	85802.57
2025M12	86567.91

Source: Authors calculation

Table 3 shows a steady upward trend in the forecasted BSE_SENSEX_CLOSE from February 2025 to December 2025. The ARIMA (2,1,2) model predicts continuous growth, with the index consistently rising each month. This suggests a positive market outlook for the forecast period, reflecting optimism and potential stability in the BSE SENSEX

IV. CONCLUSION

This study used the latest monthly index data from BSEINDIA from January 2021 to January 2025 to evaluate the stability of the BSE_Index_Close and its projected value using the ARIMA model. The instability was validated by the first set of ADF index tests, but it became more complicated by accounting for the initial differences. The most dependable of the ten models that were tested was the ARIMA model (2,1,2), which satisfied all of the goodness of fit requirements, including a high R-squared adjustment and minimum AIC, BIC, and SIGMASQ. Therefore, the ARIMA (2,1,2) model was used to predict BSE_CLOSE values for the period February 2025 to December 2025. The results confirm the ARIMA (2,1,2) model's effectiveness in providing sharp short-term forecasts to investors seeking profitable opportunities in financial markets, providing valuable insights.

REFERENCES

1. C, V. (2019). Predicting the stock market index using stochastic time series ARIMA modeling: The sample of BSE and NSE. ERN: Forecasting Techniques (Topic).
2. Arjun, R., & Suprabha, K. R. (2020). Modeling hybrid indicators for stock index prediction. In A. Abraham, A. K. Cherukuri, P. Melin, &
3. N. Gandhi (Eds.), Intelligent Systems Design and Applications. ISDA 2018 2018. Advances in Intelligent Systems and Computing (Vol. 940).
4. Springer, Cham. https://doi.org/10.1007/978-3-030-16657-1_18
5. Dadhich, M., Pahwa, M. S., Jain, V., & Doshi, R. (2021). Predictive models for stock market index using stochastic time series ARIMA modeling in emerging economy. Advances in Mechanical Engineering.
6. Murali, K., Umar, S. N., Naidu, D. C., Reddy, M. R., Mani, C., & Murthy, B. R. (2021). Stock market prediction using auto- regressive integrated moving averages (ARIMA).
7. H. R., T., & V., J. B. (2023). Forecasting the Sensex and Nifty indices using ARIMA and GARCH models. MUDRA: Journal of Finance and Accounting.
8. Jarrett, E. J., & Kyper, E. (2011). ARIMA modeling with intervention to forecast and analyze Chinese stock prices. International Journal of Engineering,
9. Business and Management, 3(3), 53-58.
10. Pai, P. F., & Shenglin, C. (2005). A hybrid ARIMA and support vector machines model in stock price prediction. Omega, 33(2005), 497-505.
11. Merh, N., Saxena, V. P., & Pardasani, K. R. (2010). A comparison between hybrid approaches of ANN and ARIMA for Indian stock trend forecasting. Journal of Business Intelligence, 3(2), 23-43.
12. Adebiyi, A. A., Adewumi, A. O., & Ayo, C. K. (2014). Stock Price Prediction Using the ARIMA Model. Proceedings of the UKSim-AMSS 16th International Conference on Modelling and Simulation. (This study applies ARIMA to NYSE and Nigerian stock data and finds ARIMA useful for short-term forecasting) ResearchGate
13. Mashadihasanli, T. (2022). Stock Market Price Forecasting Using the ARIMA Model: An Application to Istanbul, Türkiye. Journal of Economic Policy Researches, 9(2).
14. (The study finds ARIMA(3,1,5) to be the best fit for Istanbul's stock market index) ResearchGate+2Istanbul University Press+2
15. Bijesh Dhyani, Manish Kumar, Poonam Verma, Abhisheik Jain. Stock Market Forecasting Technique using ARIMA Model.
16. (Authors use ARIMA on NIFTY daily data and discuss model implementation details) ResearchGate
17. Mondal, P., Shit, L., & Goswami, S. (2014). Study of Effectiveness of Time Series Modelling (ARIMA) in Forecasting Stock Prices. International Journal of Computer Science, Engineering and Applications. (Examines ARIMA forecasting performance for Indian stocks) SpringerOpen+1
18. Meher, B. K., & Hawaldar, I. T. (2021). Forecasting Stock Market Prices Using Mixed ARIMA Model: A Case Study

of Indian Pharmaceutical Companies. Investments Management and Financial Innovations, 18

19. (The paper develops a mixed ARIMA model for sector-specific forecasting) PMC+1
20. "S&P BSE Sensex and S&P BSE IT return forecasting using ARIMA." (2020). Journal of Financial Studies and Economics (JFin-SWUFE).
21. (Uses ARIMA to forecast returns of S&P BSE Sensex and BSE IT indices; finds forecasts closely match real values) SpringerOpen
22. "ARIMA vs LSTM on NASDAQ stock exchange data." Procedia / Elsevier (2022).
23. (Compares ARIMA and LSTM performance on stock data; highlights strengths and limitations)