WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Study on Predict Sales of Two Wheelers Using Arima Model

Mr. A. Stephan Akash [1], Dr. A. Sindhiya Rebecca [2]

[1]II MBA Student, Panimalar Engineering College
[2]Assistant Professor, Department of Master of Business Administration, Panimalar Engineering College

ABSTRACT

This study investigates the sales dynamics of two wheelers-, focusing on identifying top-selling models, forecasting sales trends, and analyzing user demographics. Utilizing predictive analytics, the research aims to provide actionable insights for inventory management, marketing strategies, and customer engagement initiatives. However, limitations stemming from data availability, external influences, sample representativeness, and model assumptions may impact the accuracy and generalizability of findings. Despite these challenges, the study contributes to enhancing business decision-making processes in the competitive two-wheeler market landscape.

INTRODUCTION

Predicting sales trends is vital for businesses to maintain competitiveness and optimize operations. In the context of the automotive industry, particularly the two-wheeler segment, accurate sales forecasting holds immense importance. This study delves into the realm of predicting sales for two-wheelers, aiming to provide insights that can inform strategic decision-making processes. As Company continues to shape the industry with its innovative products and customer-centric approach, understanding sales dynamics becomes crucial for dealerships. By leveraging historical sales data, advanced analytical techniques, and predictive modeling, this research endeavors to unravel the complexities underlying sales fluctuations, identify key influencing factors, and develop robust forecasting models tailored to the dealership's context. Ultimately, the findings of this study seek to enhance inventory management, marketing strategies, and overall business performance in the dynamic landscape of the two-wheeler market."

NEED OF THE STUDY

- Understanding sales trends of two-wheelers for informed business decisions.
- Identifying and analyzing the best-selling bike models to optimize inventory and marketing strategies.
- Forecasting sales of top-selling models to anticipate demand and streamline production.
- Collecting user demographics to tailor marketing campaigns and improve customer engagement strategies.

OBJECTIVES OF THE STUDY

PRIMARY OBJECTIVES:

- To study on predict sales of two wheelers using Arima model

SECONDARY OBJECTIVES:

- To Identify the top selling bike Models.
- To Predict sales of Top selling Models.
- To Report User Demographics
- To suggest Measures based on prediction.

SCOPE OF THE STUDY

The scope of this study encompasses an in-depth analysis of sales patterns of two-wheelers aiming to provide actionable insights for strategic decision-making. The study will focus on identifying the top-selling bike models, forecasting their sales trends, and examining user demographics to understand the target market better. By leveraging predictive analytics techniques, the research intends to offer valuable recommendations for inventory management, marketing initiatives, and customer engagement strategies. Additionally, the scope extends to suggesting measures based on predictive models to optimize sales performance and enhance overall business outcomes. Through comprehensive data analysis and interpretation, this study aims to contribute to the sustained growth and competitiveness in the two-wheeler market.

REVIEW OF LITERATURE

Robert H. Shumway (2023)

Classical regression is often insufficient for explaining all of the interesting dynamics of a time series. For example, the ACF of the residuals of the simple linear regression fit to the price of chicken data (see Example 2.4) reveals additional structure in the data that regression did not capture. Instead, the introduction of correlation that may be generated through lagged linear relations leads to proposing the autoregressive (AR) and autoregressive moving average (ARMA) models that were presented in Whittle [209]. Adding nonstationary models to the mix leads to the autoregressive integrated moving average (ARIMA) model popularized in the landmark work by Box and Jenkins [30]. The Box–Jenkins method for identifying ARIMA models is given in this chapter along with techniques for parameter estimation and forecasting for these models. A partial theoretical justification of the use of ARMA models is discussed in Sect.

David S. Stoffer (2023)

This paper investigates the approach to repairable system reliability forecasting based on the Autoregressive Integrated Moving Average (ARIMA) models. This time series technique makes very few assumptions and is very flexible. It is theoretically and statistically sound in its foundation and no a priori postulation of models is required when analysing failure data. An illustrative example on a mechanical system failures is presented. Comparison is also made with the traditional Duane model. It is concluded that ARIMA model is a viable alternative that gives satisfactory results in terms of its predictive performance.

Simon Stevenson (2024)

ARIMA models have been extensively examined in the context of the real estate market. The purpose of this paper is to examine issues relating to their application in a forecasting context. Specifically, the paper seeks to examine whether in-sample measures of best-fit and also past forecasting accuracy bear any relation to future forecasting performance.

Mohamed Reda Abonazel* and Ahmed Ibrahim Abd-Elftah (2022)

The Gross Domestic Product (GDP) is that the value of all product and services made at intervals the borders of a nation in an exceedingly year. In this paper, the Box-Jenkins approach has been used to build the appropriate Autoregressive-Integrated Moving-Average (ARIMA) model for the Egyptian GDP data. Egypt's annual GDP data obtained from the World-Bank for the years 1965 to 2022. We find that the appropriate statistical model for Egyptian GDP is ARIMA (1, 2, 1). Finally, we used the fitted ARIMA model to forecast the GDP of Egypt for the next ten years.

Meyler, Aidan and Kenny, Geoff and Quinn, Terry (2023)

This paper outlines the practical steps which need to be undertaken to use autoregressive integrated moving average (ARIMA) time series models for forecasting Irish inflation. A framework for ARIMA forecasting is drawn up. It considers two alternative approaches to the issue of identifying ARIMA models - the Box Jenkins approach and the objective penalty function methods. The emphasis is on forecast performance which suggests more focus on minimising out-of-sample forecast errors than on maximising in-sample 'goodness of fit'. Thus, the approach followed is unashamedly one of 'model mining' with the aim of optimising forecast performance. Practical issues in ARIMA time series forecasting are illustrated with reference to the harmonised index of consumer prices (HICP) and some of its major sub-components.

Jamal Fattah, Latifa Ezzine (2021)

The work presented in this article constitutes a contribution to modeling and forecasting the demand in a food company, by using time series approach. Our work demonstrates how the historical demand data could be utilized to forecast future demand and how these forecasts affect the supply chain. The historical demand information was used to develop several autoregressive integrated moving average (ARIMA) models by using Box–Jenkins time series procedure and the adequate model was selected according to four performance criteria: Akaike criterion, Schwarz Bayesian criterion, maximum likelihood, and standard error. The selected model corresponded to the ARIMA (1, 0, 1) and it was validated by another historical demand information under the same conditions. The results obtained prove that the model could be utilized to model and forecast the future demand in this food manufacturing. These results

A D Indriyanti *(*2021)

The purpose of this research is to use the linear regression method to predict cycle sales results, the variable used is the period as an independent variable (X) and sales as the dependent variable (Y). The data used in the calculation of linear regression is the last four years data, from January 2014 to December 2019. The implementation of the cycle sales forecasting system is to predict sales in the coming months. To find out the level of accuracy of the prediction error calculation is needed so that it is known how many error levels are obtained. Calculation of forecasting errors using Mean Absolute Deviation (MAD) and Mean Absolute Percentage Error (MAPE). The results of this study are web-based cycle sales prediction systems using linear regression method. From this system, cycle sales forecasting is obtained the following month. In January 2015 with forecasting results of 12.63. To find out how accurate the forecasting level is, the error calculation result using Mean Absolute Deviation (MAD) is 3.40 and Mean Absolute Percentage Error (MAPE) is 44.33%. The results show that the error rate is small and the forecasting results are close to accurate.

RESEARCH METHODOLOGY

ANALYTICS RESEARCH DESIGN

Analytics research design is a comprehensive framework used to structure and guide analytical research endeavors. It involves several key steps, including formulating research questions, selecting relevant data sources, defining variables of interest, and choosing appropriate analytical techniques. Additionally, it encompasses the development of a methodology for data collection, processing, and analysis. A well-designed research plan ensures the accuracy, reliability, and validity of findings derived from data analysis. It also facilitates the interpretation and reporting of results in a clear and meaningful manner. Ultimately, analytics research design plays a crucial role in generating actionable insights that inform decision-making processes and drive business outcomes.

TIME SERIES

Time series refers to a sequence of data points collected or recorded over successive and equally spaced intervals of time. This data is typically ordered chronologically, with each observation corresponding to a specific time period. Time series data can encompass various domains, including finance, economics, weather forecasting, stock market analysis, and more. Analyzing time series data involves identifying patterns, trends, and fluctuations over time, which can aid in forecasting future values or understanding underlying relationships within the data. Common techniques for analyzing time series data include statistical methods such as moving averages, exponential smoothing, autoregression, and machine learning algorithms like ARIMA (AutoRegressive Integrated Moving Average) and LSTM (Long Short-Term Memory). Time series analysis plays a vital role in decision-making processes across industries, facilitating insights into historical behavior and enabling predictions for future outcomes.

ARIMA MODEL:

ARIMA, or Autoregressive Integrated Moving Average, is a powerful time series forecasting model widely utilized across disciplines such as finance, economics, and meteorology. This model combines three key components: autoregression (AR), which captures the relationship between an observation and its past values; differencing (I), used to achieve stationarity by removing trends or seasonal patterns from the data; and moving average (MA), which accounts for the relationship between an observation and the residual errors from a moving average model. Represented as ARIMA(p, d, q), where p, d, and q denote the number of autoregressive terms, the degree of differencing, and the number of moving average terms, respectively, ARIMA models are adept at handling various time series patterns, including trends, seasonality, and irregular fluctuations. Widely employed for tasks such as sales forecasting, stock price prediction, and demand forecasting, ARIMA models offer a versatile and robust framework for analyzing and predicting sequential data with valuable applications across industries.

ARIMA (P,D,Q):

In ARIMA models, "pdq" refers to the parameters used to specify the model structure. "p" represents the autoregressive order, indicating the number of lagged observations considered in the model to capture past dependencies. "d" signifies the differencing order, representing the number of differencing steps required to achieve stationarity in the time series data. Lastly, "q" denotes the moving average order, indicating the number of lagged forecast errors incorporated into the model to account for random fluctuations. Together, these parameters define the ARIMA model's configuration and are essential for effectively modeling and forecasting time series data, allowing analysts to capture and predict underlying patterns with accuracy.

Data Analysis

Comparision of ARIMA Model

Comparision Of (p,d,q)	ARIMA(1,1,1)	ARIMA(0,1,1)	ARIMA(0,1,0)
RMSE	1.5	2.36	8.5
MAE	1.8	4.55	7.1
MAPE	1.235	5.9	9.8

Inference

The Result of RMSE, MAE, and MAPE. RMSE calculates the average difference between predicted and actual temperatures, while MAE just averages those differences without squaring them. MAPE expresses the average difference as a percentage of the actual temperature. Lower values for these metrics indicate better predictions. Across all metrics, the "111" model performs the best, meaning its predictions are closest to the actual temperatures.

Prediction of XL100HD I-TOUCH START Using ARIMA (1,1,1):

Dep. Variable:	Quantity Sold	No. Observations	12
Model:	ARIMA(1, 1, 1)	Log Likelihood	-47.656
Date:	Wed, 01 May 2024	AIC	101.312
Time:	10:52:26	BIC	102.505
Sample:	04-30-2023	HQIC	100.559

Covariance Type: opg

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	0.5046	0.332	1.520	0.128	1.155	0.146
ma.L1	0.2553	0.455	0.561	0.575	1.148	0.637
sigma2	321.2424	156.047	2.059	0.040	15.396	627.088

Inference:

This data presents the findings of a statistical model applied to sales data over time, indicating a positive relationship between current sales and past sales. While the relationship is modest, it suggests a potential influence of past sales on current performance. Additionally, the model identifies a positive relationship between prediction errors and current sales .This implies some predictability in sales behavior. Although there is variability in sales that the model doesn't fully capture, it provides valuable insights for sales forecasting.With further refinement, the model can be enhanced to better predict future sales trends, aiding decision- making processes for improved business outcomes.

Forecast :

Predicted Month	**Predicted Sales Units**
30APRIL 2024	74.077433
31 MAY 2024	81.010862
30 JUNE 2024	72.558202
31 JULY 2024	60.777440
31 AUGUST 2024	65.171399
30 SEPTEMBER 2024	52.972614
31 OCTOBER 2024	82.072918
30 NOVEMBER 2024	80.022306
31 DECEMBER 2024	72.047844
31 JANUARY 2025	82.034958
28 FEBUARY 2025	68.041460
31 MARCH 2025	70.038179

Inference:

The forecasted sales data for the upcoming months indicates fluctuating trends, with varying projected sales volumes. May 2024 is expected to have the highest forecasted sales at approximately 81 units, while September 2024 is anticipated to have the lowest forecasted sales, around 53 units. This variability underscores the challenge of accurately predicting future sales trends. Continuous monitoring and refinement of forecasting models will be essential to improve the accuracy of future predictions and enable effective resource allocation.

Prediction of ZEST Using ARIMA (1,1,1):

Dep. Variable:	Quantity Sold	No. Observations	12
Model:	ARIMA(1, 1, 1)	Log Likelihood	-39.346
Date:	Wed, 01 May 2024	AIC	84.692
Time:	10:52:26	BIC	85.886
Sample:	04-30-2023	HQIC	83.940

Covariance Type: opg

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	0.4700	0.943	0.498	0.618	2.319	1.379
ma.L1	0.5080	0.456	0.349	0.727	3.363	2.347
sigma2	68.5315	30.921	2.216	0.027	7.927	129.136

Inference:

This data reflects the outcomes of a Predictive analysis applied to sales data trends. While the model indicates statistically significant relationship between current and past sales, or between prediction errors and past observations, it does recognize some variability in sales not fully accounted for. This highlights areas for potential refinement to enhance accuracy, ultimately providing opportunities for improving future forecasting capabilities.

Forecast:

Predicted Month	**Predicted Sales Units**
30APRIL 2024	42.230879
31 MAY 2024	54.592346
30 JUNE 2024	38.422466
31 JULY 2024	48.502305
31 AUGUST 2024	30.464783
30 SEPTEMBER 2024	52.482418
31 OCTOBER 2024	49.474130
30 NOVEMBER 2024	46.478025
31 DECEMBER 2024	40.476194
31 JANUARY 2025	50.477054
28 FEBUARY 2025	32.476650
31 MARCH 2025	27.476840

Inference :

The predicted sales data indicates fluctuating trends over the next year, ranging from around 27 to 54 units per month. May 2024 anticipates the highest sales at approximately 54 units while March 2025 shows the lowest forecasted sales, around 27 units. Such variability highlights the challenge of accurately predicting future sales trends. To effectively manage inventory and resources, continuous monitoring and adaptive strategies are crucial. Refinement of forecasting models is necessary for improved accuracy. By aligning sales forecasts with actual data and leveraging advanced analytical techniques, businesses can optimize inventory levels and stay agile in responding to market dynamics.

Prediction of JUPITER 125 Using ARIMA (1,1,1):

Dep. Variable:	Quantity Sold	No. Observations	12
Model:	ARIMA(1, 1, 1)	Log Likelihood	-45.294
Date:	Wed, 01 May 2024	AIC	96.589
Time:	11:40:30	BIC	97.497
Sample:	05-31-2023	HQIC	95.593

Covariance Type: opg

	coef	std err	Z	P>|z|	[0.025	0.975]
ar.L1	0.1280	0.867	0.148	0.883	1.827	1.571
ma.L1	0.7227	0.693	1.043	0.297	2.081	0.635
sigma2	458.4042	367.197	1.248	0.212	261.289	178.098

Inference:

This data presents insights from a Predictive model analysing sales trends over time. While the model indicates significant relationship between current and past sales or between prediction errors and past observations, it offers opportunities for enhancement. By acknowledging potential variability in sales data, the model encourages further refinement to ensure greater accuracy and effectiveness in predicting future sales trends.

Forecast:

Predicted Month	**Predicted Sales Units**
30 APRIL 2024	48.841066
31 MAY 2024	58.221342
30 JUNE 2024	46.300676
31 JULY 2024	39.290520
31 AUGUST 2024	50.291820
30 SEPTEMBER 2024	62.291653
31 OCTOBER 2024	70.291672
30 NOVEMBER 2024	48.291675
31 DECEMBER 2024	70.291675
31 JANUARY 2025	68.291672
28 FEBUARY 2025	48.300676
31 MARCH 2025	39.291672

Inference:

The forecasted sales data indicates a mix of fluctuating and stable trends over the forecast period. While some months, like May 2024 and October 2024, show an increase in projected sales, others, such as July 2024 and March 2025, demonstrate a decrease. Notably, September 2024 has a sinificant uptick in forecasted sales. The variability suggests the importance of closely monitoring market dynamics and implementing adaptive strategies to effectively manage resources.

SUGGESTIONS

- Targeted marketing strategies should address significant sales differences between genders, optimizing product offerings and enhancing customer satisfaction by understanding and accommodating gender-specific buying habits.

- Capitalizing on the popularity of XL100HD I-TOUCH START requires exploring avenues for product development or marketing focus through market research, maintaining its competitive edge while continuously refining forecasting models to adapt to fluctuating sales trends, improving inventory management and resource allocation.
- Expanding product range or marketing efforts tailored to female consumers' strong preference for scooters presents growth opportunities, addressing any purchasing behavior discrepancies between genders to optimize pricing strategies, ultimately enhancing sales performance and market competitiveness for Company through continuous monitoring of market trends and adaptive strategies.

CONCLUSION

In conclusion, the analysis underscores the importance for Company to implement targeted marketing strategies addressing gender-specific preferences, capitalize on the popularity of top-selling models like XL100HD I-TOUCH START through continuous product development, and refine forecasting models to adapt to fluctuating sales trends. Additionally, there's a clear opportunity to expand market share among female consumers through tailored product offerings and marketing efforts. Addressing purchasing behavior discrepancies between genders can further optimize sales performance. By focusing on these areas of improvement, Company can enhance its competitiveness and achieve sustained growth in the dynamic automotive industry landscape.

ANNEXURE

BIBLIOGRAPHY:

BOOKS REFFERED :

- "Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die" by Eric Siegel, Publisher: Wiley, Year: 2013
- "Practical Predictive Analytics: Using Python and Machine Learning" by Vishal Chaurasia, Publisher: Apress, Year: 2018
- "Marketing Analytics: Data-Driven Techniques with Microsoft Excel" by Wayne L. Winston, Publisher: Wiley, Year: 2016

JOURNALS REFFERED :

- Wang, Shixiong; Li, Chongshou; Lim, Andrew (2019). "Why Are the ARIMA and SARIMA not Sufficient".
- Eckerson, Wayne, W (2019). "Predictive Analytics. Extending the Value of Your Data Warehousing Investment"
- Swain, S; et al. (2018). "Development of an ARIMA Model for Monthly Rainfall Forecasting over Khordha District, Odisha, India". *Recent Findings in Intelligent Computing Techniques*. Advances in Intelligent Systems and Computing.
- Finlay, Steven (2020). Predictive Analytics, Data Mining and Big Data. Myths, Misconceptions and Methods Basingstoke: Palgrave Macmillan.

WEBSITES REFFERED :

- https://scholar.google.com
- https://www.mendeley.com/
- https://www.researchgate.net/
- https://www.mcdermott.com/Home