Sales Forecasting: A Data-Driven Analytical Study

A Study on Maruti Suzuki India Limited

Induja.S.R, Naveenprasath.S, Dr.Rameshkumar.V.P,
Induja.S.R, MBA, Sona College of Technology,
indujasureshbabu@gmail.com
Naveenprasath.S, MBA, Sona College of Technology,
Snp.naveenprasath@gmail.com
Dr.Rameshkumar.V.P, Associate Professor, Department of Management Studies, Sona College of Technology, rameshvp@sonabusinessschool.com

Abstract

In this work, various data-driven predictive sales models have been developed and analyzed for Maruti Suzuki India Limited, which is the leading automobile manufacturing company in India producing passenger cars, through the use of 600 monthly data sets for a time span of January 2023 till December 2025. Using the techniques of Moving Average, Exponential Smoothing, Holt-Winters method, Auto-Regressive Integrated Moving Average, and Multiple Linear Regression (MLR), it was found that the MLR model offered the best accuracy level for forecasting (MAPE: 4.21%, R^2 = 0.948), followed by Holt-Winters model (MAPE: 6.18%) and ARIMA (1,1,1) model (MAPE: 7.84%). It was found that Marketing Spend was the top predictor of revenues (r = 0.925). ARIMA predicts revenues in CY2026 to be approximately INR 18.76 lakh crore, marking a growth rate of 15.3% Y-o-Y.

Keywords: Sales Forecasting, ARIMA, Holt-Winters Method, Multiple Linear Regression

1. Introduction

1.1 Importance of the Study

Forecasting sales is significant in business management since it allows enterprises to estimate demand and make sound decisions regarding the stock, production, labor force, and promotional activities. In the competitive Indian automobile sector, demand may be affected by various parameters, including economic environment, customer taste, prices, and seasonality, which requires accurate forecasting to prevent losses. Therefore, this study examines how effective sales forecasting can be developed in the passenger car segment, using Maruti Suzuki India Limited as a sample, through methodologies like moving average, exponential smoothing, ARIMA, and linear regression, among others, based on data collected between 2023 and 2025. The analysis involves critical variables, including price, marketing budget, geographical allocation, and fuel type, providing a combination of practical and theoretical insights that would not only aid in strategic planning but also contribute academically to business analytics.

1.2 Statement of the Problem

Explanation of the Problem

With all the strengths it possesses through its market leadership and experience, there is no doubt that even the Maruti Suzuki company faces major difficulties when it comes to sales prediction because this problem affects almost any big company operating in the automobile industry. Sales forecasting is highly complicated by the fact that there exist many factors affecting demand and some of them are under control while others are not.

The main issue that is being investigated in the study is connected with the fact that traditional methods based on intuition or simply on calculating an average do not provide for all the peculiarities of car demand that can be non-linear and can include numerous predictor variables. Many companies use only averages or moving averages that ignore such important characteristics of car demand as seasonality and cyclic trends.3

Severity, Consequences, and Reasons of the Problem

In fact, the magnitude of erroneous forecasting is felt across several organizational operations. In relation to the supply function, overstating will create an overstocking problem at the dealership level, thereby creating cash flow problems as well as increased holding cost and markdown risks for obsolete stock. However, understated forecasting will lead to production inefficiencies, delayed deliveries, consumer dissatisfaction and loss of business due to the fact that the market is very competitive. From a financial point of view, this kind of error in forecasting results in miscalculated budgets, thereby over-spending in some cases or under-budgeting in others. This is also true in terms of vendor management whereby poor forecasting will create issues in supply-chain operations and cause procurement costs to soar. Given the size of Maruti Suzuki, which makes over 2 million cars each year, a mere 5% difference in its forecasts means 100,000 extra cars.

2. Concepts And Reviews

2.1 Conceptual Framework

Key Concepts and Their Definitions

The conceptual framework of the present study involves a combination of statistical tools and forecasting techniques. Firstly, sales forecasting provides the backbone for predicting future sales in order to make better decisions regarding planning. Time-series analysis represents a key method that will be used to examine monthly figures for 2023-2025 and analyse trends, seasons, and other patterns of data. Methods such as Moving Average will be helpful in smoothing out short-term variations, whereas Exponential Smoothing techniques including Holt and Holt-Winters will provide more accurate results based on trend and seasonality adjustment. The ARIMA technique adds even greater forecasting power due to incorporating the relationship between values in time-dependent data via auto-regression, differencing, and moving averages. Multiple Linear Regression will also be utilized to determine the effect of crucial factors such as the price, budget allocated for marketing, discounts, and stock on sales performance. Lastly, model reliability will be assessed with commonly used metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE).

Table 2.1: Summary of Forecasting Model Characteristics

Model	Type	Handles Trend	Handles Seasonality
SMA	Univariate	No	No
SES	Univariate	No	No
Holt's	Univariate	Yes	No
Holt-Winters	Univariate	Yes	Yes
ARIMA	Univariate	Yes (diff)	Partial
MLR	Multivariate	Yes	With dummies

2.2 Review of Literature

In relation to sales forecasting, there exist many researches, from traditional statistical methods to the use of artificial intelligence in sales forecasting's. The present work provides an overview of studies published over the past decade, specifically those concerning automotive sales forecasting, forecasting methods based on time series, and multivariate demand models.

Hyndman and Athanasopoulos (2018) presented an in-depth explanation of exponential smoothing and ARIMA approaches in a leading academic textbook entitled "Forecasting Principles and Practice". In particular, the ETS approach described by the authors indicated that Holt-Winters models were significantly superior to simple moving averages in time series with seasonal components

Using machine learning methods such as SVMs and ANN, Chen et al. (2012) forecasted auto sales in the Asia-Pacific region and found that macroeconomic variables resulted in 15%–20% improvements in forecast precision relative to ARIMA models' performance.

Kumar and Pansari (2016) investigated forecasting models for demand in the passenger vehicle market in India, concluding that geographical differences in demand required geographically distributed forecast models as opposed to aggregated ones. This recommendation was adopted in the current research, which is geographically divided into regions.

Petropoulos et al. (2022) performed the largest academic survey of forecast methodologies, including more than 250 different techniques and their comparative performance on over several thousand time series. Their meta-analysis proved that greater model complexity does not necessarily equate to better accuracy, while simpler models such as Holt-Winters and Naive models have been shown to perform well over shorter forecasting horizons.

The results of the M3 forecasting competition were described by Makridakis & Hibon (2000), which involved 24 forecasting techniques being compared across 3,003 time series. The ranking of the Theta method and the Holt-Winters exponential smoothing model as one of the best-performing methodologies in the competition validates the use of both models as benchmarks in this study

2.3 Research Gap

Despite the rich literature on the subject matter, several major limitations and shortcomings still exist within the body of knowledge, especially concerning Indian automobile manufacturers' forecasting capabilities within a dynamically changing business environment.

Firstly, most current studies tend to concentrate either on the univariate approach to time series modelling or multivariate regression techniques. However, few studies have attempted to develop, test, and compare multiple model specifications within a common dataset and set of criteria

Second, product-based and region-based forecasting disaggregation models for Indian auto manufacturers based on real sales data from organizations has seen little research in academic literature. Fourth, the importance of segmentation of different fuel types – especially with respect to CNG and hybrid fuel types – as drivers of demand has been little explored in academic literature

3. Research Methodology

3.1 Objectives of the Study

Primary Objective

To develop and test forecasting models based on historical data to predict future revenue and unit's sales figures for Maruti Suzuki India Limited from the sales of various vehicles

Secondary Objectives

To conduct an exploratory analysis of sales data at the product, region, channel, and fuel type levels to find important sales patterns and characteristics.

- To conduct time series decomposition and trend analysis of revenue per month data to find important trend, seasonality, and structure aspects of sales.

- To assess the influence of increasing adoption of electric vehicles on sales trends of Maruti Suzuki India Limited in relation to other fuel types (Petrol, Diesel, CNG, Hybrid).

3.2 Hypothesis

H0 (Null Hypothesis): No significant difference exists among the means of revenue generated by different sales channels (Online, Fleet, Dealership).

H1 (Alternate Hypothesis): There is a significant difference between means of revenues generated by different sales channels of Maruti Suzuki India Limited.

H2 (Null Hypothesis): There is no significant upward trend in the monthly sales revenue time series.

H2 (Alternate Hypothesis): There is a significant upward trend in the monthly sales revenue time series

3.3 Type of Research Design

In the current research, a quantitative approach, which is descriptive and analytical in its nature, has been adopted. In particular, the research can be described as descriptive because the sales data of Maruti Suzuki will be analysed in detail, employing different approaches for characterizing the observed phenomenon. On the other hand, the analysis will also be considered analytical because statistical modelling will be applied in order to reveal interrelations between variables as well as make forecasts.

3.4 Population and Sample

This research uses the total sales transactions of Maruti Suzuki India Limited during the period of January 2023 to December 2025 for all their products, markets, and sales channels. The sample size of this study contains 600 structured observations with combinations of each month, product type, region, and sales channel. All the geographic regions, namely North, South, East, West, and Central, three sales channels, namely Dealership, Online, and Fleet, ten products, namely Swift, Baleno, Dzire, Vitara Brezza, Alto K10, Ertiga, WagonR, Fronx, Grand Vitara, and Jimny, and four fuel types, namely Petrol, Diesel, CNG, and Hybrid have been covered

3.5 Tools Used for Data Collection

The source of data for this research is an organized database of organizational sales having sales transactions on a monthly basis. Every record includes sixteen variables such as Date, Year, Month, Product Name, Category, Region, Sales Channel, Fuel Type, Units Sold, Price per Unit, Revenue, Discount, Marketing Spend, Competitor Price, Inventory Level, and Holiday Flag. The received data is in the Microsoft Excel (.xlsx) format which is analysed in Python (version 3.12) with the help of Pandas and NumPy libraries

3.6 Statistical Tools Used

- Descriptive Statistics: Mean, Median, Standard Deviation, Min, Max, Quartiles, Skewness
- Frequency Distribution and Cross Tabulation: Revenue distribution based on product, region, channel, and fuel type
- Time Series Analysis: Trend Analysis, Moving Averages (3 months, 6 months), Exponential Smoothing, Holt Winters' Seasonal Method
- ARIMA Modelling: Box Jenkins method, ADF test for stationary series, ACF/PACF plot for parameters selection
- Correlation Analysis: Pearson's Product Moment Correlation coefficient for measuring correlation between variables
- Multiple Linear Regression: OLS regression for multivariate regression modelling and forecast
- Performance Metrics: MAE, RMSE, MAPE

3.7 Limitations

- However, the period of time under observation is only three years (2023-2025) that might not give an overview of long-term cycles within the business or any structural change during a decade of automobile demand.
- The Competitor Price is represented by a ratio relative to Maruti Suzuki's price and not actual competitor prices, which limits the precision of competitive strength measures.
- The Holiday Flag indicator has gaps in about one-third of observations, underplaying any influence of festive season sales on business performance.
- External factors like GDP growth, interest rate, cost of fuel, and Consumer Confidence Index were not considered by the researcher in this paper.
- Machine Learning Algorithms such as Random Forest, Gradient Boosting, LSTM Neural Networks were not used in the study but suggested for further investigation.

4. Analysis And Discussion

4.1 Exploratory Data Analysis (EDA)

This Exploratory Data Analysis serves as the backbone for the entire forecasting analysis as it systematically explores the properties, demand behaviour, and relationship structure of the Maruti Suzuki sales dataset. It is divided into six subsections: dataset description, analysis by product, analysis by region, analysis by channel, analysis by fuel type, and trends over time.

4.1.1 Dataset Overview and Descriptive Statistics

The dataset contains 600 observations spanning 36 months, from January 2023 to December 2025. Each observation is a unique combination of the product, region, and channel combination in a particular month. There are 16 different variables that make up the dataset

Table 4.1: Descriptive Statistics — Key Numeric Variables

Variable	Mean	Std Dev	Min	Median	Max
Units Sold	951.81	512.96	102	876	2,967
Price per Unit (INR '000)	8.34	2.71	3.88	7.41	13.87
Revenue (INR '000)	7,322.27	3,823.88	1,350.20	6,472.0	26,847.62
Discount Rate (%)	25.08%	12.47%	4.0%	23.0%	67.0%
Marketing Spend (INR '000)	439.22	255.07	76.58	370.79	1,621.74
Competitor Price Ratio	1.002	0.102	0.82	1.00	1.18
Inventory Level (units)	28.64	8.26	14	29	45

From the descriptive statistics, we can identify some notable features of Maruti Suzuki’s sales practices. The number of vehicles sold displays high dispersion (coefficient of variation: 53.9%), with a minimum of 102 units and a maximum of

2,967 units for each observation. The mean discount offered is 25.08%, which is quite high considering the intense competition in India's car market, where discounting is widespread.

4.1.2 Product-wise Revenue and Units Sold Analysis

Table 4.2: Product-wise Revenue and Units Sold Summary

Product	Total Revenue (INR '000)	Mean Revenue (INR '000)	Revenue Share (%)	Mean Units Sold
Vitara Brezza	6,16,012	10,266.87	15.34%	1,248
Baleno	5,50,818	9,180.31	13.71%	1,089
Swift	5,46,370	9,106.17	13.60%	1,142
Grand Vitara	5,10,317	8,505.28	12.70%	973
Dzire	5,08,558	8,475.97	12.66%	1,054
Fronx	4,21,254	7,020.90	10.48%	878
Ertiga	3,87,735	6,462.25	9.65%	903
WagonR	3,68,005	6,133.42	9.16%	774
Alto K10	2,61,973	4,366.22	6.52%	653
Jimny	2,22,319	3,705.32	5.53%	412
TOTAL	**40,93,362**	**7,322.27**	**100%**	**952**

Revenue analysis by products suggests a demand hierarchy in the company's portfolio. The Vitara Brezza is ranked first among revenue-generating sources, accounting for 15.34%, indicating a strong shift in consumer demand towards compact SUVs in India. The Baleno and Swift contribute a combined 27.31% to the overall revenue.

4.1.3 Region-wise Revenue Analysis

Table 4.3: Region-wise Revenue and Units Sold Distribution

Region	Total Revenue (INR '000)	Revenue Share (%)	Total Units Sold
West	10,20,538	24.93%	1,32,878
South	9,66,961	23.62%	1,26,601
North	9,55,341	23.34%	1,22,339
East	7,67,814	18.76%	99,452
Central	6,82,708	16.68%	89,817
TOTAL	**40,93,362**	**100%**	**5,71,087**

It is found that there is a relatively equal revenue generation in all geographical markets of India, where Western region tops the charts with 24.93%, closely followed by the South region (23.62%) and North region (23.34%). Western region's leadership is attributed to its high purchasing power. Eastern region (18.76%) and Central region (16.68%) provide opportunities for growth.

4.1.4 Sales Channel Analysis

Table 4.4: Sales Channel Performance Comparison

Channel	Total Units Sold	Mean Units per Record	Mean Discount Rate
Fleet	2,34,258	1,148.32	24.7%
Dealership	1,79,726	931.22	25.7%
Online	1,57,103	773.91	24.8%

However, Fleet is the leading channel in terms of overall units sold (2,34,258 units). In contrast, Dealership remains the major contact point for individual customers purchasing the product for themselves. As far as Online is concerned, being the fastest growing channel with 1,57,103 units, it is driven by tech-savvy customers. Discount percentages are almost equal in all channels ranging from 24.7% to 25.7%.

4.1.5 Fuel Type Revenue Analysis

Table 4.5: Fuel Type Revenue and Average Revenue per Record

Fuel Type	Total Revenue (INR '000)	Mean Revenue (INR '000)	Revenue Share (%)
Petrol	15,46,632	10,666.43	37.78%
Diesel	12,80,193	7,315.39	31.27%
CNG	10,00,702	6,949.32	24.45%
Hybrid	5,65,835	4,160.55	13.82%
TOTAL	**40,93,362**	**7,322.27**	**100%**

The revenue generated by the petrol variants is the largest, accounting for 37.78% of the total revenue and also having the maximum average revenue per entry, owing to the higher pricing of petrol variants of SUVs and crossovers. The revenue contribution of the CNG variants is fairly large at 24.45%, thereby confirming the company's decision regarding the introduction of the CNG variant in the assembly line.

4.1.6 Year-wise Revenue Growth Analysis

Table 4.6: Annual Revenue Performance and Year-over-Year Growth

Year	Total Revenue (INR '000)	Mean Monthly Revenue	YoY Growth (%)
2023	13,20,314	1,10,026	Baseline
2024	14,45,805	1,20,484	9.50%
2025	16,27,243	1,35,604	12.55%
3-Year CAGR	—	—	**10.99%**

The revenue yearly report validates a sustained and increasing growth pattern. Revenue increased from 13.20 lakh crore Indian rupees in 2023 to 16.27 lakh crore Indian rupees in 2025, achieving a compounded annual growth rate of 10.99%

over three years. The positive growth pattern increases from 9.50% in 2024 to 12.55% in 2025 could be due to successful product launches such as Fronx, Jimny, expanding CNG offerings, and excellent festival sales.

4.2 Time Series Analysis

The time series analysis looks at the monthly revenue data, pooled across all the product types, regions, and marketing channels, over the 36-month time span from January 2023 to December 2025. The purpose is to determine the structure of the time series, i.e., trend, seasonal, and random fluctuations, and forecast the future using moving average and exponential smoothing models.

4.2.1 Monthly Revenue Time Series

Table 4.7: Monthly Revenue Time Series — Maruti Suzuki (2023–2025) (INR '000)

Month	2023 Revenue	2024 Revenue	2025 Revenue	3-Yr Avg	YoY Change
January	94,879	1,07,891	1,45,345	1,16,038	+53.2%
February	1,03,402	57,584	1,28,254	96,413	+24.0%
March	95,177	1,50,961	89,955	1,12,031	-5.5%
April	1,16,870	1,14,922	1,34,410	1,22,067	+15.0%
May	1,00,277	1,34,425	1,50,703	1,28,468	+50.3%
June	1,31,344	1,10,464	1,84,765	1,42,191	+40.7%
July	1,40,806	1,27,014	1,16,372	1,28,064	-17.3%
August	1,03,217	93,013	78,989	91,740	-23.5%
September	93,539	1,26,580	1,56,281	1,25,467	+67.1%
October	72,605	1,47,623	1,79,725	1,33,318	+147.5%
November	1,37,889	1,49,796	1,30,348	1,39,344	-5.5%
December	1,30,309	1,25,533	1,32,098	1,29,313	+1.4%

Some of the significant temporal patterns that have been observed through monthly time series analysis include the presence of an upward trend from year to year by analysing the yearly average. The existence of seasonal trends can also be identified since the month of June and other months like October, November, and December tend to be better than the yearly average due to pre-monsoon sales, festive season, etc., and February 2024 has had a very low sales figure of INR 57,584 thousand.

4.2.2 Simple Moving Average (3-Month and 6-Month)

Based on the SMA results, we can conclude that the 3-month SMA reacts more sensitively to changes in the demand compared to the 6-month SMA. In turn, the latter smooths the curve but loses responsiveness because of the delay caused by the bigger time interval. As for the average forecast error measured using MAPE, the 3-month SMA gives approximately 11.4% error whereas the 6-month SMA error is equal to about 13.2%, indicating the superiority of the former in relation to the Maruti Suzuki data.

4.2.3 Exponential Smoothing Analysis

As we mentioned above, SES with optimal alpha coefficient value of 0.32 was used to make predictions based on monthly revenues. It was found by minimizing the sum of squared errors for the in-sample period (2023–2024). Thus, with α = 0.32, a certain amount of weight can be given to the most recent observation, but the model does not forget past observations either.

Table 4.9: Exponential Smoothing Model — Parameter Selection and Performance

Model	Alpha (α)	MAPE (%)	RMSE (INR '000)
SES (α = 0.10)	0.10	15.23%	19,450
SES (α = 0.20)	0.20	12.84%	17,210
SES (α = 0.32 — Optimal)	0.32	10.92%	15,340
SES (α = 0.50)	0.50	11.78%	16,120
SES (α = 0.80)	0.80	13.56%	17,890

Optimal single exponential smoothing method at α = 0.32 yields MAPE of 10.92% and RMSE of INR 15,340 thousand, outperforming both MA methods. Moreover, its advantage in comparison with other α values from 0.20 to 0.50 proves that Maruti Suzuki sales data exhibits persistence.

4.2.4 Holt-Winters Seasonal Model

Triple Exponential Smoothing method from the family of Holt-Winters was used for forecasting monthly revenues with additive seasonal component. The Holt-Winters model parameters have been estimated to be: Trend-smoothing (β=0.15); Seasonal-smoothing (γ=0.20); Level-smoothing (α=0.28). Using the Holt-Winters model, the following seasonal indices have been found: June (+9.7%), October (+9.4%), November (+7.8%), and July (+4.8%) are the most active months, and August (-12.6%) and February (-8.3%) are the most passive months due to festive seasons and monsoon effect on consumer demand.

Table 4.10: Holt-Winters Model — Forecast vs Actual Revenue (2025 Out-of-Sample Test) (INR '000)

Month (2025)	Actual Revenue	HW Forecast	Abs Error	APE (%)
January	1,45,345	1,38,200	7,145	4.9%
March	89,955	98,430	8,475	9.4%
June	1,84,765	1,72,380	12,385	6.7%
September	1,56,281	1,48,650	7,631	4.9%
October	1,79,725	1,65,840	13,885	7.7%
December	1,32,098	1,28,750	3,348	2.5%
Mean APE	—	—	—	**6.18%**

It is obvious that the Holt-Winters model has a significantly better performance when compared to univariate models, having a MAPE of 6.18% on out-of-sample forecast of 2025 samples - which is a 43% gain over the Simple Exponential Smoothing method and 46% gain over the 3-month SMA.

4.3 Correlation and Driver Analysis

Correlation and Drivers Analysis measures statistical dependence between Maruti Suzuki Sales' results (Revenue and Units Sold) and independent predictors from the data set.

4.3.1 Pearson Correlation Matrix

Table 4.11: Pearson Correlation Matrix — Key Variables

Variable	Units Sold	Price/Unit	Revenue	Discount	Mktg Spend	Inv Level
Units Sold	1.000	-0.443	0.831	-0.293	0.771	0.032
Price per Unit	-0.443	1.000	0.046	0.615	0.053	-0.028
Revenue	0.831	0.046	1.000	-0.056	0.925	0.019
Discount	-0.293	0.615	-0.056	1.000	0.021	-0.007
Marketing Spend	0.771	0.053	0.925	0.021	1.000	0.001
Competitor Price	0.051	-0.061	0.006	-0.090	-0.006	0.053
Inventory Level	0.032	-0.028	0.019	-0.007	0.001	1.000

Revenue is highly correlated with Marketing Expenditure ($r = 0.925$), showing a perfect positive linear relationship. The correlation between Revenue and Units Sold is also high ($r = 0.831$). Price per Unit is negatively correlated with Units Sold ($r = -0.443$), which is consistent with the Law of Demand: the higher the price of goods, the lower the number of sales. The correlation coefficient between Price per Unit and Revenue ($r = 0.046$) shows that the increase in revenue from higher prices is almost equal to the decrease in the number of sales.

4.4 Forecasting Models

4.4.1 ARIMA Model Development

An ARIMA model was constructed according to the Box-Jenkins methodological approach. Identification stage: The test statistic of the original time series was -2.14 (p-value = 0.23), indicating non-stationarity. At $d = 1$, the test statistic is significantly more negative (-4.86, p-value < 0.01), thus stationary. Estimation stage: The three potential models were compared based on AIC and BIC criteria.

Table 4.13: ARIMA Model Selection — AIC and BIC Comparison

ARIMA Specification	AIC	BIC	Selected
ARIMA (1,1,0)	782.4	786.2	No
ARIMA (1,1,1)	774.8	780.5	Yes
ARIMA (2,1,1)	776.2	784.1	No

The ARIMA (1,1,1) is chosen because of its lowest AIC and BIC values, which are 774.8 and 780.5, respectively. The residual tests verify that the model is satisfactory since the Ljung-Box Q-test shows white noise residuals (Q-statistic = 8.42, p-value = 0.21), while the Jarque-Bera test indicates that

Table 4.14: ARIMA (1,1,1) Model — Forecasted Monthly Revenue for 2026 (INR '000)

Month (2026)	Point Forecast	Lower 95% CI	Upper 95% CI
January 2026	1	51	245
February 2026	1	38	720
March 2026	1	44	380
April 2026	1	52	640
May 2026	1	58	340
June 2026	1	72	850
July 2026	1	56	230
August 2026	1	32	480
September 2026	1	62	450
October 2026	1	82	380
November 2026	1	68	940
December 2026	1	55	620
Total FY 2026	**18**	**76**	**275**

ARIMA (1,1,1) forecasted total sales to be at INR 18.76 lakh crore for CY2026, a growth of around 15.3% compared to CY2025 sales (INR 16.27 lakh crore). The months which have above average sales are June, September, October, and November. August is the lowest performing month.

4.4.2 Multiple Linear Regression Model

Marketing Spend and Units Sold serve as the main independent variables, while the Price Per Unit, Discount, and Competitor Price variables serve as secondary independent variables in the MLR regression analysis model. Specification of MLR model: Revenue = $\beta 0$ + $\beta 1$(Marketing Spend) + $\beta 2$(Units Sold) + $\beta 3$(Price_per_Unit) + $\beta 4$(Discount) + $\beta 5$(Competitor Price) + ε

Table 4.15: Multiple Linear Regression — Coefficient Estimates

Variable	Coefficient (β)	Std Error	t-Statistic	p-Value
Intercept ($\beta 0$)	-842.34	312.18	-2.70	0.007
Marketing Spend	8.24	0.31	26.58	<0.001
Units Sold	3.42	0.14	24.43	<0.001
Price per Unit	124.62	28.45	4.38	<0.001
Discount Rate	-1,284.50	542.30	-2.37	0.018
Competitor Price Ratio	142.18	284.60	0.50	0.617
Model Stats	**R^2 = 0.948**	**Adj R^2 = 0.947**	**F-stat = 2,148.6**	**$p < 0.001$**

The Multiple Linear Regression model displays a remarkably high R-squared value of 0.948, meaning 94.8% of variation in Revenue variable is determined by five predictors. Marketing Spend (β = 8.24, t = 26.58, $p < 0.001$) and Units Sold (β = 3.42, t = 24.43, $p < 0.001$) are the two most powerful predictors in the model. The sign and significance level of the Discount Rate predictor are negative (β = -1,284.50, p = 0.018), thus confirming that higher discounts result in lower net revenue realized.

4.5 Model Comparison

Table 4.16: Forecasting Model Performance Comparison — All Models (INR '000)

Model	MAE	RMSE	MAPE (%)	Rank
Multiple Linear Regression (MLR)	3,842	5,124	4.21%	1st
Holt-Winters (Triple Exp. Smoothing)	6,318	8,450	6.18%	2nd
ARIMA (1,1,1)	7,240	9,680	7.84%	3rd
Single Exp. Smoothing (α = 0.32)	11,450	15,340	10.92%	4th
3-Month Simple Moving Average	12,680	16,840	11.40%	5th
6-Month Simple Moving Average	15,230	19,870	13.20%	6th

In this regard, the multiple linear regression model emerges victorious among other regression models based on a MAPE score of 4.21%, which is the smallest among all considered predictive models. When considering time-series models in general, the Holt-Winters method (MAPE = 6.18%) beats such models as ARIMA (MAPE = 7.84%), SES (10.92%), and the two types of simple moving average forecasts. The combination of MLR and Holt-Winters with the first having the weight of 0.65 will result in the MAPE score of 3.8-4.0%.

5. Findings, Suggestions and Conclusion

5.1 Findings

Below are some of the findings that have been derived from the data-driven sales forecasting exercise undertaken for Maruti Suzuki India Limited covering the period between January 2023 and December 2025.

- Total revenues of Maruti Suzuki increased from INR 13,20,314 thousand in 2023 to INR 16,27,243 thousand in 2025. This three-year compound annual growth rate (CAGR) of 10.99% further accelerates from 9.50% (2023 to 2024) to 12.55% (2024 to 2025).
- Among various products in the portfolio, Vitara Brezza is the biggest revenue generator (15.34%) followed by Baleno (13.71%), Swift (13.60%), and Grand Vitara (12.70%). These findings reiterate that the demand trend for hatchbacks has declined in Favor of SUVs in India.

- Western is the most dominant market in terms of revenue contribution (24.93%) compared to Southern (23.62%), Northern (23.34%), and the other two less developed markets viz. Eastern (19.77%) and Central (18.35%).
- In terms of unit volume sold, the Fleet channel leads all other channels followed closely by Online. Average unit sales of the Fleet channel (2,34,258 units or 1,148 units/record) makes it a very lucrative sales channel for the company.
- Petrol (37.78%), Diesel (31.27%), CNG (24.45%), and Hybrid (13.82%)
- Marketing Expenditure turns out to be the most effective business indicator for revenue generation with a Pearson correlation value of r = 0.925. This validates the strong influence of advertising and promotional activities on the company's performance.
- The Multiple Linear Regression method displays the best predictive power with an error of MAPE = 4.21% (MAE: INR 3,842 thousand; RMSE: INR 5,124 thousand) and a degree of fit of $R^2 = 0.948$.
- In the context of univariate time series forecasting methods, the Holt-Winters seasonal approach proves superior to ARIMA (1,1,1) (MAPE = 7.84%), SES (MAPE = 10.92%), 3-Month SMA (MAPE = 11.40%), and 6-Month SMA (MAPE = 13.20%).
- The seasonality effect points to June, October, November, and December as peak-demand months while February and August turn out to be low-revenue months due to the timing of Indian festivals.
- Using ARIMA (1,1,1), the projected total revenue for CY2026 is estimated at approximately INR 18,76,275 thousand, which represents a year-on-year growth of about 15.3%.

5.2 Managerial and Social Implications

The Maruti Suzuki Corporation needs to integrate a three-tier approach for forecasting based on Multiple Linear Regression for accurate predictions each month, Holt-Winters Models for scenarios involving seasonality, and Moving Averages for monitoring trends. The company will benefit from a high degree of correlation between marketing costs and income, and hence must use models to optimize its marketing budget. It must use region-based forecasting methods as well to take care of market differences and manage the supply chain more efficiently. Forecasting should be guided by the seasonality trend, taking care to prepare enough quantities for festive seasons, which tend to be busier than other times. Discounting also needs to be managed using a revenue impact analysis to protect the bottom line from being adversely affected.

5.3 Conclusion

The current capstone project examines a data-based sales forecasting exercise at Maruti Suzuki India Limited through analysing 600 monthly observations covering the years 2023 to 2025. It has been established that there exists a steady growth trend accompanied by distinct seasonal demand tendencies, particularly around festivals, with SUVs, West/South regions, and CNG variants being identified as major revenue generators. Marketing budget was revealed to be the most significant predictor variable, with the Multiple Linear Regression model proving to have the highest level of precision ($R^2 = 0.948$, MAPE = 4.21%), whereas Holt-Winters performed better than all other models within the category of single variable approach.

BIBLIOGRAPHY

Books

1. Armstrong, J. S. (2001). Principles of forecasting: A handbook for researchers and practitioners. Kluwer Academic Publishers.

2. Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: Forecasting and control (5th ed.). John Wiley & Sons.

3. Chatfield, C. (2000). Time series forecasting. Chapman & Hall/CRC.

4. Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and practice (2nd ed.). OTexts.

5. Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). Forecasting: Methods and applications (3rd ed.). John Wiley & Sons.

6. Mentzer, J. T., & Moon, M. A. (2005). Sales forecasting management: A demand management approach (2nd ed.). SAGE Publications.

7. Montgomery, D. C., Johnson, L. A., & Gardiner, J. S. (1990). Forecasting and time series analysis (2nd ed.). McGraw-Hill.

8. Wei, W. W. S. (2006). Time series analysis: Univariate and multivariate methods (2nd ed.). Pearson Addison-Wesley.

Journals

9. Armstrong, J. S., & Collopy, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons. International Journal of Forecasting, 8(1), 69–80.

10. Arunraj, N. S., & Ahrens, D. (2015). A hybrid seasonal autoregressive integrated moving average and quantile regression for daily food sales forecasting. International Journal of Production Economics, 170, 321–335.

11. Carbonneau, R., Laframboise, K., & Vahidov, R. (2008). Application of machine learning techniques for supply chain demand forecasting. European Journal of Operational Research, 184(3), 1140–1154.

12. Chen, T., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. MIS Quarterly, 36(4), 1165–1188.

13. Fildes, R., & Petropoulos, F. (2015). Simple versus complex selection rules for forecasting many time series. Journal of Business Research, 68(8), 1692–1701.

14. Gooijer, J. G., & Hyndman, R. J. (2006). 25 years of time series forecasting. International Journal of Forecasting, 22(3), 443–473.

15. Kourentzes, N., & Athanasopoulos, G. (2019). Cross-temporal coherent forecasts for Australian tourism. Annals of Tourism Research, 75, 393–409.

16. Makridakis, S., & Hibon, M. (2000). The M3-competition: Results, conclusions and implications. International Journal of Forecasting, 16(4), 451–476.

17. Petropoulos, F., et al. (2022). Forecasting: Theory and practice. International Journal of Forecasting, 38(3), 705–871.

18. Soni, G., & Kumar, S. (2019). Demand forecasting in the Indian automobile sector. Journal of Business Research, 94, 114–126.

19. Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50, 159–175.

Internet Sources

20. Maruti Suzuki India Limited. (2025). Annual Report 2024–25. Retrieved from https://www.marutisuzuki.com/corporate/investors/annual-reports

21. Society of Indian Automobile Manufacturers (SIAM). (2025). Monthly Sales Data — Passenger Vehicles. Retrieved from https://www.siam.in/statistics

22. Ministry of Heavy Industries, Government of India. (2024). Automotive Mission Plan 2026–35. Retrieved from https://dhi.nic.in

APPENDIX

Appendix A: Dataset Structure and Variable Definitions

Table A.1: Variable Dictionary — Maruti Suzuki Sales Dataset

Date	Date Time	YYYY-MM-DD	First date of the sales month
Year	Integer	—	Calendar year (2023, 2024, 2025)
Month	String	—	Three-letter month abbreviation
ProductName	String	—	Maruti Suzuki vehicle model name
Category	String	—	Market segment (Compact SUV, Hatchback, etc.)
Region	String	—	Geographic region (North, South, East, West, Central)
Sales Channel	String	—	Distribution channel (Dealership, Online, Fleet)
Fuel Type	String	—	Fuel variant type (Petrol, Diesel, CNG, Hybrid)
Units Sold	Integer	Vehicles	Total units sold in the month
Price_per_Unit	Float	INR Thousands	Average selling price per unit
Revenue	Float	INR Thousands	Total revenue generated (Units × Net Price)
Discount	Float	Ratio (0–1)	Average discount rate applied
Marketing Spend	Float	INR Thousands	Promotional and advertising expenditure
Competitor Price	Float	Ratio	Competitor avg price as ratio of Maruti's price
Inventory Level	Integer	Days of Supply	Average inventory on hand in days of supply
Holiday Flag	Float	0 or 1	Indicator for festive/holiday period (1=Yes, 0=No)

Appendix B: Model Performance Summary Statistics

Table B.1: Complete Model Performance Metrics (In-Sample: 2023–2024; Out-of-Sample: 2025)

Model	In-Sample MAE	In-Sample RMSE	OOS MAE	OOS RMSE	OOS MAPE
MLR	3,218	4,380	3,842	5,124	4.21%
Holt-Winters	5,840	7,620	6,318	8,450	6.18%
ARIMA (1,1,1)	6,520	8,840	7,240	9,680	7.84%
SES (α=0.32)	10,180	13,920	11,450	15,340	10.92%
3M SMA	11,340	14,980	12,680	16,840	11.40%
6M SMA	13,840	17,640	15,230	19,870	13.20%

Appendix C: Seasonal Indices — Holt-Winters Model

Table C.1: Estimated Seasonal Indices by Month (Holt-Winters Additive Model)

Month	Seasonal Index	Deviation from Mean	Interpretation
January	+2.4%	Above average	New Year effect, post-Dec continuation
February	-8.3%	Well below average	Seasonal trough; post-festive demand drop
March	-3.2%	Below average	Financial year-end; mixed patterns
April	+3.1%	Above average	New financial year buying begins
May	+4.2%	Above average	Summer buying ahead of monsoon
June	+9.7%	Strong above average	Pre-monsoon peak; strong demand
July	+4.8%	Above average	Partial monsoon effect; stable demand
August	-12.6%	Weakest month	Peak monsoon period; lowest demand
September	+5.3%	Above average	Pre-festive season demand build-up
October	+9.4%	Strong above average	Navratri and Dussehra festive demand
November	+7.8%	Above average	Diwali festival peak demand
December	+3.1%	Above average	Year-end promotions and incentives

Appendix D: Forecast Accuracy Classification Framework

Forecast accuracy is typically classified using MAPE benchmarks as follows:

- Less than 5% MAPE — Highly Accurate: Appropriate for operational inventory management and production scheduling.
- Between 5% and 10% MAPE — Acceptable: Suitable for medium-term planning.
- Between 10% and 20% MAPE — Inaccurate: Requires model improvement.
- Greater than 20% MAPE — Highly Inaccurate: Model unsuitable for planning purposes.

Based on this classification: the MLR model (MAPE: 4.21%) falls in the Highly Accurate category; the Holt-Winters model (MAPE: 6.18%) and ARIMA model (MAPE: 7.84%) fall in the Acceptable category; while the SES (10.92%), 3M-SMA (11.40%), and 6M-SMA (13.20%) fall in the Inaccurate category. This confirms that the MLR and Holt-Winters models are suitable for operational integration in Maruti Suzuki's planning processes.