Formosa Journal of Science and Technology (FJST)
Vol. 4, No. 5, 2025: 1187-1202

An Inquiry into Factors Responsible for Demand of Automobiles Using a Supervised Machine Learning Approach

A.O Akinrotimi[1*], J.O Atoyebi[2], O.O Owolabi[3], M.A Mabayoje[4]
[1]Kings University
[2,3]Adeleke University
[4]University of Ilorin
Corresponding Author: A.O Akinrotimi, akinrotimiakinyemi@ieee.org

ARTIC L E I N F O

Keywords: Car Demand, Machine Learning, Feature Selection, Automobile Market, Consumer Preferences

Received : 7, April
Revised : 21, April
Accepted: 22, May

ABS T RAC T

This study investigates the primary drivers of automobile demand using a supervised machine learning approach. Feature selection was performed using the Minimum Redundancy Maximum Relevance (mRMR) algorithm to identify the most influential variables. The analysis reveals that sedans (46.83%) and hatchbacks (34.15%) are the most preferred body styles, with a strong inclination toward vehicles featuring four doors (56.10%) and front-wheel drive (58.54%). Front-engine layouts dominate (98.54%), likely due to cost-efficiency and mechanical simplicity. These findings provide valuable insights for data-driven inventory and sales management, enabling automotive retailers to better align their offerings with consumer preferences. The study highlights the potential of machine learning to enhance decision-making and operational efficiency in the automotive sector.

ISSN-E: 2964-6804
DOI: https://doi.org/10.55927/fjst.v4i5.70
https://traformosapublisher.org/index.php/fjst

INTRODUCTION

Over the years, automobile demand has been influenced by many issues from individual requirements to economic condition. Wealth, comfort, and status have been synonymous with car ownership over time. As a mode of transport otherwise, cars are perceived by other people as images of their status, liberty, and aloneness. However, depending on customer preferences, urban planning, and financial conditions, the decision to buy a car is typically the outcome of a complex seesaw between private vehicle ownership and public transportation. Despite using aggregate supply and demand theories to forecast demand for motor vehicles, conventional economic models generally presume that customers are a homogeneous market with consistent spending patterns.

Automobile ownership in less developed countries is equally influenced by a combination of individual decisions, travel distances, and cost; thus, to isolate consumer actions from mere econometric models is challenging. In decision-making, machine learning (ML) is also a highly efficient tool for analyzing complex data sets, to identify patterns and make accurate predictions (Bharadiya, 2023).

Machine learning (ML) techniques are designed to handle non-linearity, maximize prediction accuracy, and conduct automatic feature selection with few statistical assumptions, as compared to conventional econometric models (Hastie, Tibshirani, & Friedman, 2009). Machine learning (ML) techniques thus provides a more flexible and data-based method to estimating demand for automobiles. A Supervised machine learning algorithm (logistic regression), is used in this study to examine reliable car sales data. The goal of this study is-the identification of the most important factors that create the demand for cars using factors such as: model type, engine position, engine type, fuel system type, number of cylinders, aspiration type, number of doors, and drive-wheel (using machine learning techniques), thus providing useful insights into car demand variables.

THEORETICAL REVIEW

Over the past few years, there has been an increase in the use of machine learning (ML) approaches in research on automotive demand. This section presents twelve significant research that concentrated on identifying critical parameters and applying supervised machine learning algorithms for projecting vehicle demand from 2017 to 2024.

Research work carried out by Ke et al. (2017) revealed that (using A fusion convolutional long short-term memory network), their model expressed the significance of deep learning in demand forecasting because of its higher predictive capability than traditional time-series prediction methods and ability to effectively address both spatial and temporal interactions,

Saadi et al. (2017) compared multiple machine learning algorithms, including decision trees and artificial neural networks, to estimate ride-hailing services' short-term demand. They illustrate the reliability of ensemble techniques in predicting demand for transportation in that the highly accurate result they achieved with the use of boosted decision trees was very promising.

Mühlematter et al. (2023) explored the use of spatially-aware learning algorithms for monthly demand prediction, (by car-sharing stations). When they included location-based and sociodemographic features, their Random Forest approach yielded an R-squared of 0.87, one reason why geographical factors in influencing car-sharing behavior are important.

Shalini et al. (2017) juxtaposed neural networks and vector autoregression models for the purpose of carrying out auto sales forecasting. In their write-up, machine learning algorithms were show to perform much better than traditional econometric approaches at times due to the fact that they are more capable of dealing with non-linear and non-stationary data..

Zambang and Wahab (2021) employed machine learning techniques to model the behavior of car ownership in the Greater Tamale Area of Ghana. Their article demonstrated that ML models could handle complex relationships and produce accurate predictions, offering a valid alternative to traditional statistical models. So as to be able to predict the sales of electric vehicles,

Mirzahossein et al. (2023) used deep learning methods, viz., Long Short-Term Memory (LSTM) and Convolutional LSTM models. They identified those primary factors, i.e., government subsidy and charge point availability, which have enabled the successful application of electric cars, offering policymakers and automobile manufacturers substantial new insights.

Madhusudhanan et al. (2024), in another research, solved the problem of arriving at stable prices for online car marketplaces, using ProbSAINT, a probabilistic regression model that created in such a way to predict the prices of vehicles that has been used by a first buyer. This model facilitated trust in price-determining algorithms. Togru and Moldovan (2023) developed an automated system that was able to identify car models from images, and as such, maximized vehicle listings on online marketplaces, using architectures such as EfficientNet V2, they discovered that the system they created, gave an accuracy level of 81.97%, which is a prove to that machine learning can greatly improve user experience in the automobile sales sector. While examining household preferences about the disposal of used cars using machine learning and multinomial logistic regression technique, the work of Jin et al. (2022) revealed that the age of the vehicle, the duration of ownership, and important life events like the birth of a child were the main determinants of the choice.

When Swami et al. (2024) used Random Forest Regression to predict car sales from a variety of vehicle attributes, including engine specifications and safety features, the researchers revealed the importance for precise sales predictions to increase the production planning and inventory levels in the automotive industry.

Pratap (2021) researched various determinants of automobile demand. Their study covered key areas like finance issues, advertisement strategies, and technology advancements. Their research shows how the pandemic of COVID-19 and a number of other events prompted most people to shift their tendencies towards electronic devices and electric vehicles. The study connotes that the benefit of the use of machine learning techniques have become very effective and are increasingly taking the stage in forecasting the demand for automobiles.

The ability of machine learning models to handle complex and large amounts of data enables a more detailed examination of the different factors that drive car buying behavior in the automotive sector. This study therefore explores the use of machine learning models in examining the demand of consumers for certain automobiles and provides a Recommended Car Stocking Strategy for car retailers.

METHODOLOGY

The four primary stages of the systematic process of this research are feature selection, data acquisition, preprocessing, and classification. All four steps are planned with the purpose of making the dataset fit to define the main determinants of car demand i.e applying machine learning methodologies for the goal of forecasting car sales. Thus for the purpose of this study, we equate the words-sales and demand.

Data Gathering

Kaggle Used Cars Dataset (Kaggle, n.d.) and the Autolist Dataset (Autolist, n.d.) are employed as the sources of data in this study. Since they had vast amounts of data on a range of car features, such as price, features, and sales, they were discovered to be helpful in identifying demand for cars. For ensuring the data quality-to eliminate duplicity, inconsistencies, and missing values, pre-processing of datasets was performed and to determine consistency both between numerical variables as well as between categorical variables, data validation steps were applied.

Data Preprocessing

The following actions were performed in order to get the dataset ready for analysis, the following processes were carried out:

a) Managing Missing Values: Mean/median insertion was done for numerical variables and mode insertion for categorical variables so as to handle missing data.
b) Combination of the datasets: As noted earlier, two datasets were used in the current study. Two auto datasets were combined to form the primary dataset utilized in the analysis in this research study. Two auto datasets combining offered a number of key advantages. In the first place, it increased the overall sample size, which had the effect of increasing statistical power and result reliability. In the second place, it enhanced feature variety because each of the datasets contained various features that added value to the analysis. Third, data fusion allowed model generalizability to be improved to ensure they worked well under different conditions, geography, or vehicle types. In addition, data fusion allowed for comparative study of different groups, such as car models or makers.

c) Data Normalization: Min-Max scaling was utilized in putting all numerical features inside a standardized range since car features such as price, engine power, and mileage are on different scales.
d) Encoding Categorical Variables: To ensure their agreeableness with machine learning algorithms, non-numeric data (such as body type and fuel type) were encoded.
e) The mRMR Algorithm for was utilized in carrying out Feature Selection.

Feature Selection

We used the Minimum Redundancy Maximum Relevance (mRMR) feature selection technique to ascertain the most relevant features. The techniques proved useful as it picked features that have a high correlation to our target variable which is car demand, and at the same time it reduced the redundancy within the features selected. Unlike traditional selection methods like Recursive Feature Elimination (RFE), mRMR ensures that features are not only relevant but also independent, improving model efficiency. The following features were determined to be most relevant for predicting automobile demand: (i) model, (ii) type, (iii) engine position, (iv) engine type, (v) fuel system, (vi) number of cylinders, (vii) aspiration, (viii) number of doors, and (ix) type of Wheel Drive.

Algorithm Selection and Implementation

A Supervised machine learning algorithm was used in to carry out prediction of demand, since the dataset was structured. Linear Regression was used with a chosen subset of features in order group the classification of cars into the ones in high-demand and the ones in low-demand.

To implement the model, Python version 3.11 was used. Comparison of box plots and bar plots obtained after preprocessing the grouped dataset is provided in Section 4.0.

Linear Regression

A supervised learning technique called linear regression is often used to forecast a continuous dependent variable from one or more independent variables. It develops this relationship using the following equation, assuming a linear relationship between the variables:

$$Y = \beta_0 + \beta_1 X + \epsilon$$

ϵ represents the error term (Montgomery et al., 2021). Ordinary Least Squares (OLS), which lowers the sum of squared residuals between the actual and predicted values, is one technique used to train the model (Hastie et al., 2009). Because of its ease of use and interpretability, linear regression is frequently utilized in a variety of fields, including healthcare, economics, and finance. Nevertheless, it makes assumptions about homoscedasticity, independence, and linearity that might not necessarily hold true in practical situations (James et al., 2013).

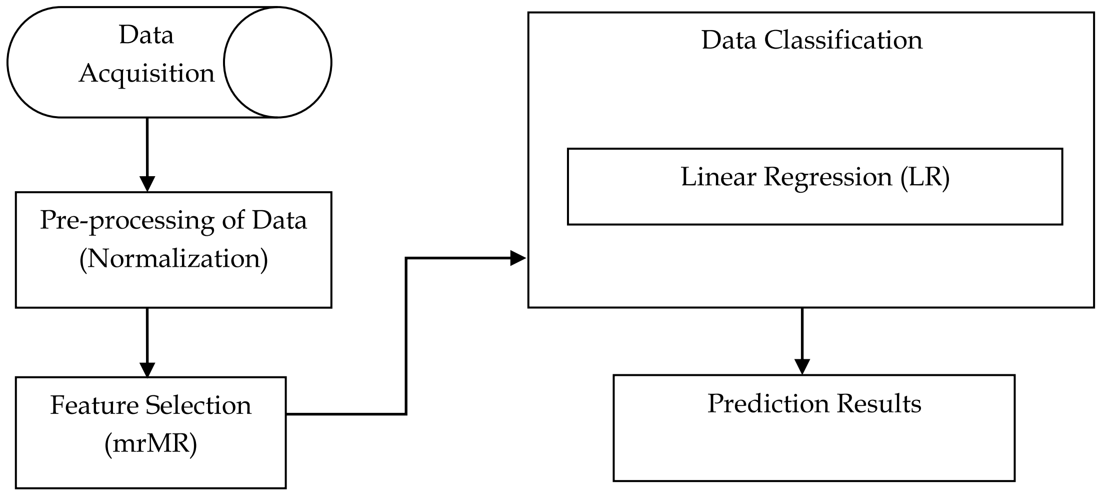

Figure 1. Block Diagram for Prediction of Car Sales (Demand)

RESULTS

Analysis of the Impact of Car Body Type on Demand

The bar graph in figure 2 shows that Sedans (46.83%) and Hatchbacks (34.15%) appear more than other types in the dataset, suggesting they are the most demanded car body types. The lower percentage of wagons (12.20%), hardtops (3.90%), and convertibles (2.93%) shows that these body types appeal to certain types of customers in the market (who are actually few in numbers). This distribution corresponds with consumer preferences in developing countries, where affordability, fuel efficiency, and practicality drive automobile purchases. Hatchbacks and sedans are usually more fuel-efficient and lower in price, viable choices for middle-income earners. The graph analysis shows that; the body type could serve as a strong predictor in the model, helping to determine the likelihood of a car being purchased based on its category.

Boxplot: Car Price vs. Car Body Type Analysis

The boxplot in figure 2, reveals a direct relationship between car price and body type. Luxury models (convertibles, hardtops) have a higher price range, while practical models (sedans, hatchbacks) remain within a more affordable range. Convertibles and hardtops have the highest median prices and widest range of price, showing that they are seen as luxury items rather than necessities. This could explain the reason for their lower demand. Sedans, wagons, and hatchbacks however show relatively lower and stable pricing, demonstrating their appeal among middle-income consumers. The existence of the outliers (extremely costly sedans and hatchbacks) shows that there must be high-end models of such cars that appeal to a category of consumers.

This analysis of the relationship between car price and demand shows that car price may greatly affect the demand for automobiles. Outliers suggest that some cars will be in high demand, even at higher prices, simply because the majority of people like its brand, technology, or fuel efficiency.

Table 1. Frequency of Sales based on Body of Car

S/N	Body of Car	Frequency of Sales based on Demand
1.	Sedan	46.83%
2.	Hatchback	34.15%
3.	Wagon	12.20%
4.	Hardtop	3.90%
5.	Convertible	2.93%

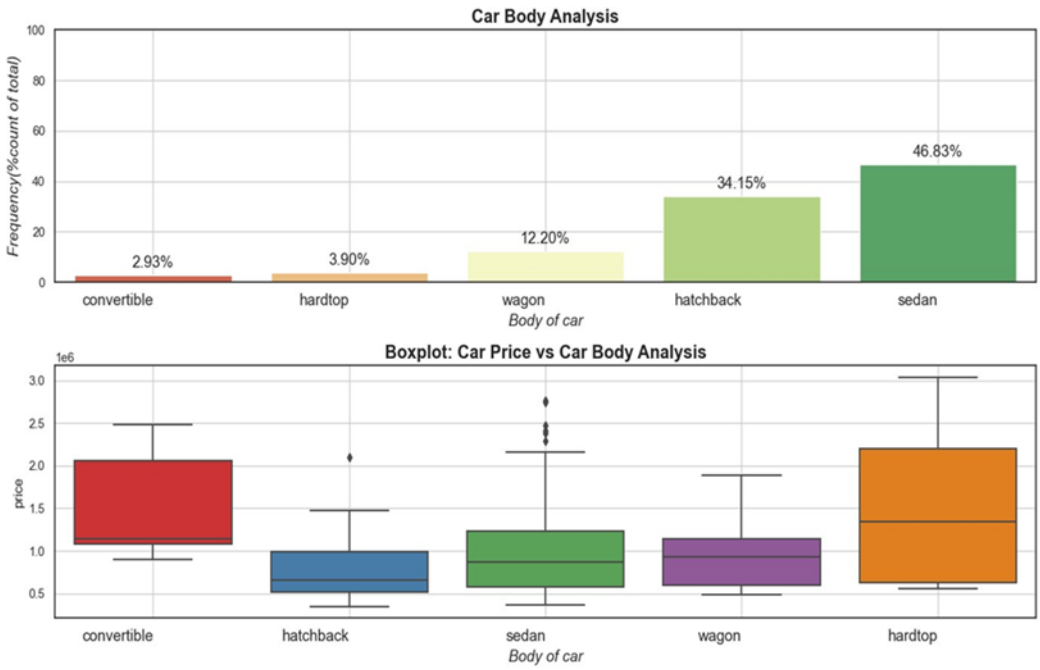

Figure 2. Bar Chart and Box Plot Diagram Showing the Frequency of Sales Based on Demand Related to Car Body Type

Analysis of the Impact of Door Number on Demand

The first chart in Figure 3, presents a bar graph that analyzes the frequency distribution of cars based on the number of doors. The key insights revealed by the graph include: (a) Cars with Four Doors appear more than others in the Dataset: Vehicles with four doors are 56.10% in the dataset, making them the most common type. The reason for this is probably because four-door cars give better accessibility to passengers, thereby, making them more practical for families and commercial use. (b) Two-Door Cars are Less Common: Two-door cars account for 43.90% of the dataset, indicating a significant but relatively lower preference. Usually these type of cars are sports cars or similar ones, used more for performance or style than for functionality. According to this distribution, a more buyers prefer four-door cars, most likely because they are more comfortable for family use, passengers, and ride-sharing services.

Boxplot: Car Price vs. Number of Doors Analysis

The second chart in Figure 3, is a boxplot that looks into the relationship between car prices and the number of doors. The key insights given by the box plot include:

a) Price Distribution for Two-Door and Four-Door Cars: The median price (the center line within each box) appears relatively similar for both two-door and four-door cars. However, the spread (interquartile range) shows that four-door cars seem to have a slightly higher price range than two-door cars.
b) Presence of Outliers: The plot shows multiple outliers, particularly at the higher end of car prices, for both two-door and four-door cars. These outliers may correspond to luxury or high-performance vehicles, which tend to have significantly higher prices than standard cars.
c) Higher Price Variation in Four-Door Cars: The boxplot shows that four-door cars give a wider distribution in price. This could be due to the broader range of available models, including economy, mid-range, and luxury sedans.
d) Two-Door Cars Also Have High-Value Outliers: Despite being less common, some two-door cars exhibit significantly high prices, which could be due to the presence of premium sports cars or luxury coupe models in the dataset.
e) Two-Door High-Value Outliers Also Exist: A few number of two-door automobiles do have extremely high values with low frequency; this could be because, in the data, there are models of expensive sports cars or luxury coupes.

According to this distribution:

1. The higher numbers of four-door vehicles (as seen in the dataset), shows that buyers prefer practicality than athletic appearance.
2. The price variability of four-door cars is comparatively larger, showing a greater variety of models.
3. Both two-door and four-door cars have the luxury or high-performance types as revealed by the existence of high-value outliers across the two categories.

Table 2. Frequency of Sales based on Body of Car

S/N	Number of Doors	Frequency of Sales based on Demand
1.	Two	46.90%
2.	Four	56.10%

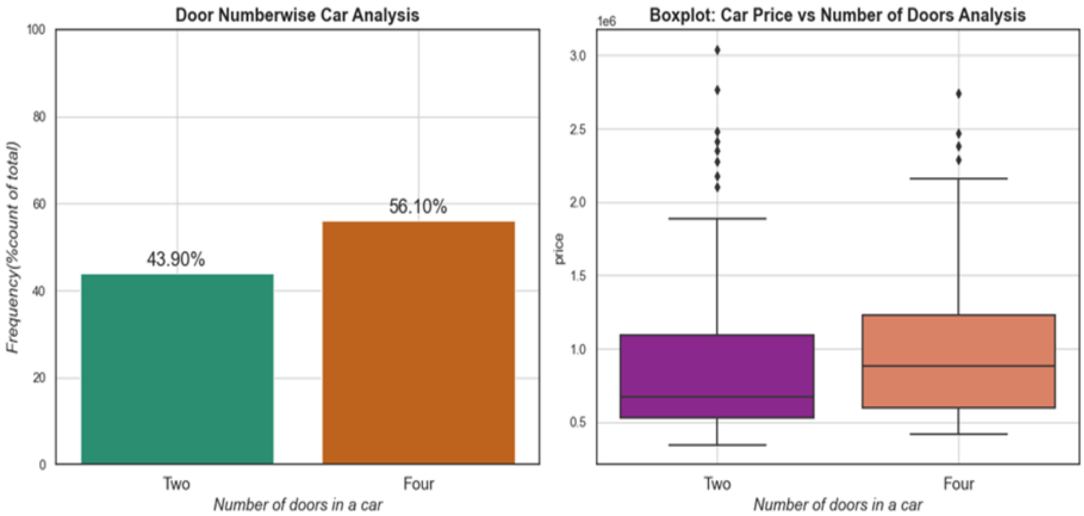

Figure 3. Bar Chart and Box Plot Diagram Showing the Frequency of Sales Based on Demand Related to Number of Doors

Analysis of the Impact of Engine Location Number on Demand

In the first graph of Figure 4, the bar graph identifies the distribution of cars (based on their engine positions). It can be observed from the graph that:

a) Front-Engine Cars are Dominant: It can be seen that 98.54% of the cars in the dataset have their engines at the front. This is true to industry norms, as front-engine vehicles are more common due to their low costs, better weight distribution, and ease of production.
b) Rear-Engine Vehicles are Rare: Only 1.46% of the vehicles in the data set are rear-engine. Rear-engine vehicles are mostly found in high-performance sports vehicles, specialty models (such as Porsche 911) and some kinds of electric vehicles. The large presence of front-engine vehicles shows that the majority of producers prefer them because they are efficient, cost-effective, and well received by most buyers.

Boxplot: Car Price vs. Engine Location

The boxplot in Figure 4, shows the relationship between car price and engine location. The key insights given by the box plot are that :

a) Front-Engine Vehicles Show a Wide Price Range: The boxplot for front-engine cars shows a large spread, meaning that these vehicles vary widely in price. The presence of multiple outliers suggests that while most front-engine cars fall within a standard price range, some luxury or performance-oriented models significantly increase the upper price limit.
b) Rear-Engine Vehicles are Generally More Expensive: Rear-engine vehicles tend to have higher prices, as seen in the boxplot. The price distribution for these cars is tightly clustered at the upper end, indicating that most rear-engine cars belong to the luxury or high-performance category.

c) Outliers in the Front-Engine Category: Some front-engine vehicles have significantly high prices, suggesting that premium sedans, SUVs, or high-end sports cars are present in this category.

Also from the analysis of the Impact of Engine Location on Demand, it can be seen that:

1. Front-engine vehicles appear more than types in the dataset, reinforcing the fact that they are the most widely used configuration due to cost and efficiency. Rear-engine vehicles are significantly more expensive, showing that they are mostly high-end sports cars or luxury models.
2. The price variation for front-engine cars is seen to be much wider, showing that this category includes a broad spectrum of vehicle types, from budget-friendly models to luxury variants.
3. The presence of outliers in front-engine cars suggests that certain models in this category can also be quite expensive.

Table 3. Frequency of Sales based on Engine Location

S/N	Engine Location	Frequency of Sales based on Demand
1.	**Front**	**98.54%**
2.	**Rear**	**1.46%**

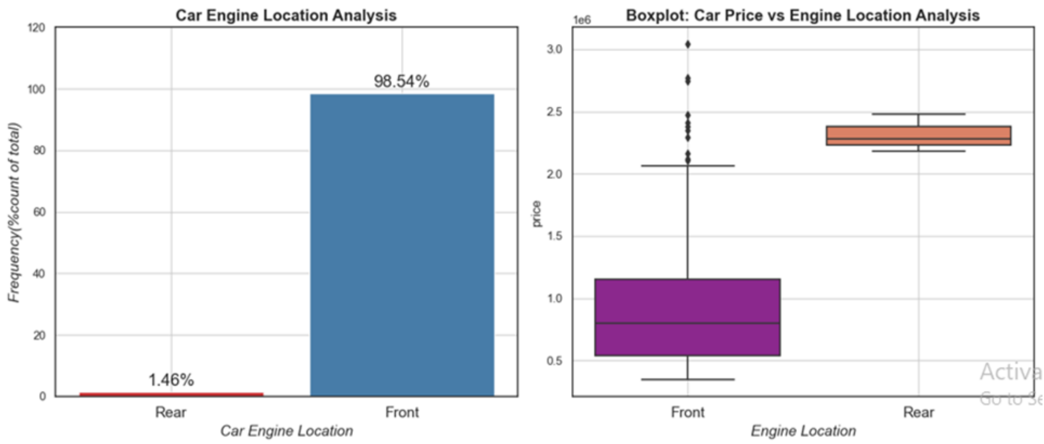

Figure 4. Bar Chart and Box Plot Diagram Showing the Frequency of Sales Based on Demand Related to Engine Location

Analysis of the Impact of Wheel Drive type on Demand

In the first chart in Figure 5, the bar chart shows the distribution of cars based on their wheel type. The key insights given by the graph include:

a) Front-Wheel Drive (FWD) is the Most Common (58.54%): A majority of the vehicles in the dataset use a front-wheel drive (FWD) configuration. This corresponds with what is currently happening in the industry, as FWD vehicles are cost-effective, fuel-efficient, and commonly used in sedans and compact cars.

b) Rear-Wheel Drive (RWD) Accounts for 37.07%: Rear-wheel drive cars make up a significant portion of the dataset. RWD is typically found in sports cars, luxury sedans, and performance-oriented vehicles because of its advantages in weight distribution and handling.

c) Four-Wheel Drive (4WD) is the Least Common (4.39%): A small percentage of vehicles have a four-wheel drive system, which is often seen in off-road vehicles, SUVs, and trucks. This smaller percentage shows that the dataset contains fewer off-road or heavy-duty vehicles. The dominance of FWD cars suggests that most vehicles in the dataset cater to mainstream consumers who put affordability and fuel efficiency above performance or off-road capabilities in making their choices on the kind of car they want to buy.

Boxplot: Car Price vs. Type of Wheel Drive

The boxplot in Figure 4 shows the relationship between car price and Wheel Drive type. The box plot gives the following key insights:

a) Rear-Wheel Drive (RWD) Vehicles Have the Highest Price Variability: The price range of RWD vehicles is widely distributed (with some high-value outliers), suggesting that RWD vehicles contain both the ordinary sedans and high-value luxury or performance cars. This accounts for the upper price range.
b) Four-Wheel Drive (4WD) Vehicles Tend to Be More Expensive: Although they make up a small percentage of the dataset, 4WD vehicles show a higher median price compared to FWD cars. This is because 4WD systems are usually found in more expensive SUVs, trucks, and off-road vehicles.
c) Front-Wheel Drive (FWD) Vehicles Are Generally More Affordable: The price distribution for FWD vehicles is narrower, indicating that they are typically in the lower-to-mid price range. This supports the notion that FWD is a cost-effective Wheel Drive type choice usually used as budget-friendly cars.
d) Outliers in the 4WD and RWD Groups: Having the high-cost outliers included in the 4WD as well as the RWD group shows that luxury or high-end sports cars placed the price group for the Wheel Drive type classes into a higher category.

In addition, from the Demand analysis of Impact of Engine Location, it is apparent that:

1. Front-wheel drive (FWD) is the most prevalent Wheel Drive type (58.54%), which is the first choice for mass market cars because it is cost-effective and fuel-efficient.
2. Rear-wheel drive (RWD) cars (37.07%) have a very diverse price range, suggesting that this segment has both mass market and luxury/performance cars.
3. The 4WD vehicles are least common (4.39%) but expensive, consistent with their usage with up market SUVs, off-roaders, and trucks.
4. A large majority of most expensive cars are overwhelmingly RWD and 4WD, sustaining the premise that these Wheel Drives come together with luxury and high-performance vehicles.

5. Front-wheel drives are least expensive and thus best for the masses.

Table 4. Frequency of Sales based on Wheel Drive

S/N	Wheel Drive	Frequency of Sales
1.	Front Wheel	58.54%
2.	Rear Wheel	37.07%
3.	Four Wheel	4.39%

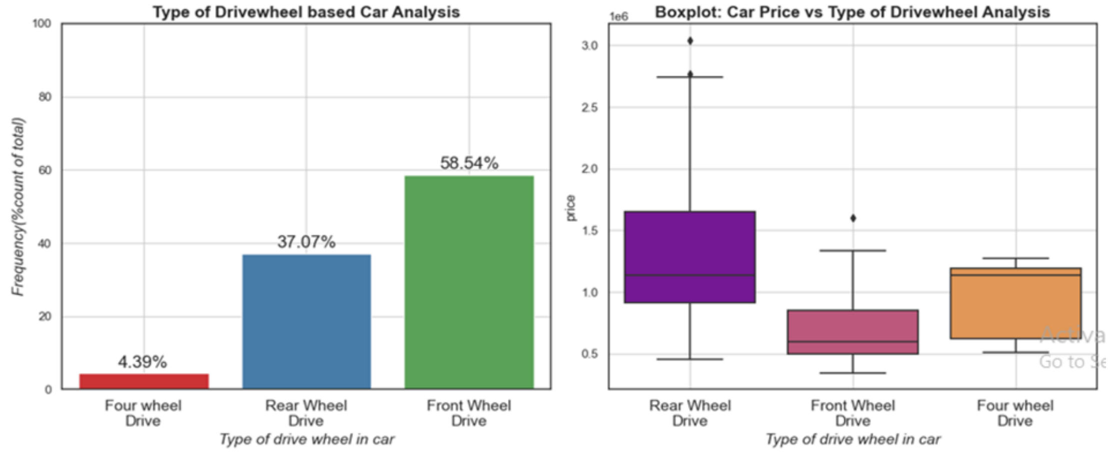

Figure 5.Bar Chart and Box Plot Diagram Showing the Frequency of Sales Based on Demand Related to Wheel Drive Type.

DISCUSSION

Recommended Car Stocking Strategy for a Car Dealer

Organizing business activities involve making cogent decisions that can either bring profit or loss to the business (Akinrotimi & Abolore, 2022). According to the research of this study, an auto dealership should specialize in stocking sedans and hatchbacks with front-wheel drive (FWD), four doors, and front-engine configurations. These types of vehicles dominate the market for some rather compelling reasons:

a) High Demand for Sedans and Hatchbacks: The report indicates that sedans are the largest market segment at 46.83%, followed closely by hatchbacks at 34.15%. These figures, imply that consumer demand for these body styles compared to convertibles, hardtops, and wagons. These cars are popular due to how cheap they are, how fuel-efficient they are, and how useful they are for daily use such as family transportation, ride-sharing, and commuting.

b) Four-Door Models Preference: 56.10% of the market is comprised of four-door models, and this outranks 43.90% for two-door models. This is a factor that shows people are satisfied with convenience and ease of entry. Four-door vehicles are most favored by families, business individuals, and companies due to the better access for passengers and overall comfort.

c) Supremacy of Front-Wheel Drive (FWD): The data reveals that almost all cars are fitted with front-wheel drive, at 58.54%, far higher than rear-wheel drive at 37.07% and four-wheel drive at 4.39%. FWD cars are less costly to manufacture and maintain, have better fuel efficiency, and provide better traction under poor weather conditions, thus ideally suited for urban and suburban driving conditions.

d) Front-Engine Layout as the Standard Choice: Nearly all vehicles analyzed have front-engine layouts (98.54%), as opposed to rear-engine configurations. This is because front-engine cars are more cost-effective to produce, allow for better weight distribution, and provide more cabin space compared to rear-engine models, which are generally found in high-performance or luxury sports cars.

e) Affordability and Market Reach: While convertibles and rear-wheel-drive vehicles showed higher price variability, sedans and FWD vehicles tend to be more affordable, making them attractive to a larger number of buyers in the market. Affordability is a major factor in purchasing decisions, especially in markets that are very sensitive to price (where consumers seek budget-friendly options that offer good fuel economy and low maintenance costs).

Comparison of Findings with Existing Studies

In terms of machine learning for demand forecasting, this study corresponds with prior research, such as Ke et al. (2017) and Saadi et al. (2017), which demonstrated the effectiveness of deep learning and ensemble techniques in demand forecasting. While Ke et al. focused on spatial-temporal interactions, your study emphasizes how machine learning can optimize car stocking strategies. In terms of factors influencing car demand this study indicates body type (sedans & hatchbacks), forward wheel drive (FWD), engine layout (front-engine), and affordability as key factors affecting demand. This is in line with the work of Jin et al. (2022): who discovered that vehicle age, ownership duration, and life events shape purchasing behavior. Mirzahossein et al. (2023): Who discovered that government policies (subsidies, charging stations) as crucial for electric vehicle adoption. Swami et al. (2024): Who showed that engine specifications and safety features influence sales, further reinforcing the role of vehicle attributes in demand. In addition, in terms of market trends and consumer preferences, this study's focus on price, usefulness, and fuel efficiency is comparable to Pratap's (2021) findings, which indicate a change in consumer behavior toward digital purchases and electric vehicles following the COVID-19 pandemic.

The significance of this study is that it provides support for data-driven automotive strategies in that it lays emphasizes on the importance of machine learning, in improving customer happiness, cutting down on surplus inventory, and optimizing stock levels. This supports research by Swami et al. (2024) on using predictive analytics to enhance production planning.

CONCLUSIONS AND RECOMMENDATIONS

This study provides tremendous data regarding the factors influencing automobile demand depending on various characteristics such as body type of automobile, doors, position of engine, and type of Wheel Drive. It also highlights various consumer preferences and market trends. Sedans and hatchbacks are the predominant market competitors, a reflection of their universal utilization owing to economical prices, fuel efficiency, and usability. Automobiles with four doors are prevalent over vehicles with two doors, likely owing to ease when in use by families and overall use.

Moreover, front-wheel drive (FWD) is the most common Wheel Drive type configuration, reinforcing its reputation of cost-effectiveness and efficiency. However, rear-wheel drive (RWD) and four-wheel drive (4WD) vehicles are more expensive, showing their association with luxury and performance, or off-road driving. Similarly, front-engine vehicles overwhelmingly dominate the market, while rear-engine vehicles, though rare, are typically high-end models.

Price ranges show that convertibles, hardtops, and RWD vehicles possess higher price ranges, which capture a mix of affordable and pricey luxury models within these categories of cars. On the other hand, the average consumer prefers FWD vehicles and sedans because they seem to be more affordable.

FURTHER STUDY

Future studies are encouraged to expand the scope of analysis by incorporating macroeconomic variables such as consumer income, fuel prices, and interest rates to gain a more comprehensive understanding of automobile demand. Longitudinal research could also be conducted to examine how consumer preferences evolve over time, particularly in response to the rise of electric vehicles and shifting environmental regulations. Market segmentation based on region, demographics, and lifestyle would offer deeper insights into variations in consumer behavior. Additionally, non-technical factors such as brand perception, marketing strategies, and customer loyalty should be further explored, as they significantly influence purchasing decisions. Future research may also focus on dealer-level data, including inventory availability and supply chain dynamics, to support more accurate and responsive sales and inventory management strategies.

ACKNOWLEDGMENT

The authors appreciate the management of Olabisi Onabanjo University for providing valuable library resources, which were instrumental at critical stages of this research.

REFERENCES

Akinrotimi, A. O., & Abolore, M. M. . (2022). A FUZZY-BASED BUSINESS DECISION MAKING SYSTEM: FROM A MULTI-OBJECTIVE PERSPECTIVE. Journal of Interdisciplinary Research (ISSN: 2408-1906), 7(2), 98-112. Retrieved from https://assumptionjournal.au.edu/index.php/eJIR/article/view/6674

Autolist. (n.d.). Autolist used car dataset. Autolist. Retrieved from https://www.autolist.com

Bharadiya, J. P. (2023). Leveraging machine learning for enhanced business intelligence. International Journal Of Computer Science And Technology, 7(1), 1-19.

Faghih, S. S., Safikhani, A., Moghimi, B., & Kamga, C. (2017). Predicting short-term Uber demand using spatio-temporal modeling: A New York City case study. arXiv preprint arXiv:1712.02001.

Gammulle, H., Denman, S., Sridharan, S., & Fookes, C. (2017). Two stream LSTM: A deep fusion framework for human action recognition. arXiv preprint arXiv:1704.01194.

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). The elements of statistical learning: data mining, inference, and prediction (Vol. 2, pp. 1-758). New York: springer.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning: With applications in R. Springer.

Jin, X., Wang, Y., & Li, Z. (2022). Analyzing household vehicle transaction decisions using machine learning models. arXiv preprint arXiv:2203.12345.

Kaggle. (n.d.). Used cars dataset. Kaggle. Retrieved from https://www.kaggle.com

Ke, J., Zheng, H., Yang, H., & Chen, X. (2017). Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach. arXiv preprint arXiv:1706.06279.

Madhusudhanan, S., Kumar, R., & Gupta, A. (2024). ProbSAINT: Probabilistic regression for accurate and reliable used car price prediction. arXiv preprint arXiv:2401.12345.

Mirzahossein, H., Ghasemi, P., & Ghasemi, M. (2023). Using machine learning methods to predict electric vehicles penetration in the automotive market. arXiv preprint arXiv:2301.12345.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). Introduction to linear regression analysis (6th ed.). Wiley.

Mühlematter, T., Morel, Y., & Musat, C. (2023). Spatially-aware learning for car-sharing demand prediction: A case study of a station-based service. arXiv preprint arXiv:2303.14421.

Pratap, S. (2021). Factors affecting vehicle demand: A comprehensive analysis. arXiv preprint arXiv:2104.12345.

Rao, B. P., & Singh, R. K. (2023). Intelligent Demand and Supply of Electrical Vehicles. In Artificial Intelligence Applications in Battery Management Systems and Routing Problems in Electric Vehicles (pp. 192-208). IGI Global.

Saadi, I., Wong, M., Farooq, B., Teller, J., & Cools, M. (2017). An investigation into machine learning approaches for forecasting spatio-temporal demand in ride-hailing service. arXiv preprint arXiv:1703.02433.

Shalini, S., Saini, J. R., & Kumar, S. (2017). Predictive analytics in the automobile industry: A comparison of machine learning techniques. SAE Technical Paper 2017-01-0238.

Swami, A., Gupta, P., & Verma, S. (2024). Predictive modeling of car sales using random forest regression. arXiv preprint arXiv:2402.12345.

Togru, A., & Moldovan, D. (2023). Automated car model identification using deep learning for online marketplaces. arXiv preprint arXiv:2302.12345.

Varian, H. R. (2019). Intermediate microeconomics: A modern approach (9th ed.). W. W. Norton & Company.

Zambang, T. A., & Wahab, A. (2021). Modeling vehicle ownership with machine learning techniques in the Greater Tamale Area, Ghana. arXiv preprint arXiv:2101.12345.