Analysis of Factors Influencing Sports Shoes Purchase Preferences: An Exploration Based on Logistic Regression Model

Qingrui Liu*
Guangdong University of Finance & Economics, Guangdong, China
* Corresponding Author Email: wozhen@ldy.edu.rs

Abstract. Sports shoes were originally designed as protective footwear for the feet, made from leather and plant materials. Over time, they evolved into various forms to meet different activity requirements during ancient times, including in ancient Egypt, Greece, and Rome. With the advent of the Industrial Revolution in the 19th century, sports shoes began to be mass-produced. Today, sports shoes have transcended their basic function and have become symbols of fashion and lifestyle, characterized by diverse designs and materials. Consumers around the world, faced with a plethora of sports shoe styles, are influenced by various factors in their purchasing decisions. This research obtained sales data for 30 Nike and Off-White collaborative sports shoe models from "Kaggle", spanning from September 1, 2017, to February 13, 2019, across all fifty states of the United States. By constructing a logistic regression model, this research aims to investigate the factors influencing the purchasing preferences for sports shoes. The goal is to deepen our understanding of the characteristics and design preferences of global consumers regarding sports shoes, enhance brand positioning, and promote the prosperity of the sports shoe industry.

Keywords: Sports Shoes; Purchase Preferences; Logistic Regression.

1. Introduction

1.1. Research Background

The origins of sports shoes can be traced back to ancient civilizations, where the earliest sports footwear or foot protection devices were rudimentary, crafted from simple leather or plant materials. During the times of ancient Egypt and Greece, people used basic shoes made from leather or palm leaves to protect their feet during physical activities and sports. In the Roman era, Roman soldiers wore open-toed sandals known as "caligae," which were suitable for their marching and combat activities while also providing ventilation for their feet. By the 19th century, with the advent of the Industrial Revolution, the manufacturing of sports shoes began to industrialize, and the earliest sports shoes, such as basketball and soccer shoes, emerged.

Nowadays, with the rise of fashion and health consciousness, sports shoes have not only become essential gear for athletes but also a symbol of popular fashion in everyday life. Their functionality is no longer confined solely to protecting the feet during sports; they also offer people a multitude of stylish and trendy choices in daily attire and the fashion domain.

From the earliest simple functional sports shoes, which focused solely on providing basic foot protection and comfort, to the diverse designs and materials available today, the sports shoe market continues to evolve. Designers have transcended mere functional requirements, shaping sports shoes into works of fashion art. They are no longer just a type of footwear but a cultural symbol, representing individual taste and lifestyle choices [1].

1.2. Research Objectives and Significance

With the increasing diversity in the sports shoe market, global consumers face numerous choices when purchasing sports shoes. What are their preferences and decision-making factors when it comes to sports footwear? This research aims to delve into consumer inclinations when buying sports shoes through the application of a logistic regression model.
Through this research, we hope to gain a deeper understanding of the characteristics and designs that global consumers favor when purchasing sports shoes. This insight will provide related brands with a more accurate market positioning, enabling them to better meet consumer demands and drive the continuous growth of the sports shoe market. Additionally, these research findings hold significant reference value for companies involved in sports shoe design, production, and sales.

By optimizing product design and marketing strategies and enhancing product competitiveness and market share, this research will further contribute to the prosperity and expansion of the sports shoe industry.

1.3. Research Methodology

The logistic regression model is a classical statistical learning method commonly used to address binary classification problems. It finds extensive applications in various fields, including social sciences, natural sciences, medicine, business, and more. In this research, we have chosen to employ the logistic regression model to investigate the factors influencing the purchase preferences of global soccer enthusiasts for soccer shoes.

The core idea of the logistic regression model is to map the output of linear regression through a logistic function (such as the sigmoid function) into a probability space, resulting in a probability prediction value. For binary classification problems, the objective of the logistic regression model is to fit a boundary that separates the two classes of samples in feature space. The model employs maximum likelihood estimation to determine parameters, thereby identifying the optimal classification boundary.

This research will construct a logistic regression model, formulate the logistic regression equation, and assess the impact of various variables on consumer preferences for purchasing sports shoes based on the feature coefficients within the equation.

1.4. Research Structural Framework

This research aims to conduct an in-depth analysis of global consumers' preferences and decision factors when purchasing sports shoes. To achieve this objective, we have developed a comprehensive research framework, encompassing key elements such as research hypotheses and model construction, to better understand consumer purchasing preferences and market trends. The specific components of the research are as follows:

1) Research Hypotheses
   As the sports shoe market continues to diversify, consumers are often influenced by various factors when choosing sports footwear. This research will initially formulate hypotheses regarding the factors that affect consumer preferences when purchasing sports shoes.

2) Variable Design
   After formulating the hypotheses, appropriate adjustments will be made to design the dependent and independent variables for the study. To accommodate the construction requirements of the logistic regression model, all variables will be transformed into binary (0-1) variables.

3) Data Collection and Processing
   Obtained the complete sales data of 30 Nike and Off-White collaborative sports shoe models from Kaggle, covering all fifty states in the United States for the period between September 1, 2017, and February 13, 2019. The collected data underwent preprocessing to ensure its effectiveness for model construction.

4) Construction of the Influencing Factors Model for Sports Shoe Purchase Preferences Based on Logistic Regression
   With the processed data, multicollinearity among independent variables will be assessed using the Variance Inflation Factor (VIF) test. Subsequently, the dataset will be divided into training and testing sets to build the logistic regression model. Model performance will be evaluated based on accuracy and the model's AUC value.

5) Results Analysis
The logistic regression equation will be generated, and the influence of each independent variable on consumer sports shoe purchase preferences will be determined based on their respective feature coefficients. The results will be analyzed from various perspectives, including shoe design and consumer preference psychology.

This structured framework will guide the research in systematically exploring and analyzing the factors that influence consumers’ purchasing preferences for sports shoes.

2. Literature Review

2.1. Literature Review

With the flourishing diversity of the sports shoe market, numerous scholars have conducted research from various perspectives on sports shoe design and consumer preferences for purchasing sports shoes.

In 2013, Zijin Mai conducted a study on the impact of sports shoe brand image on consumer purchase intentions. This research targeted the sports shoe consumer group and established a research framework. It included three aspects of brand image (corporate image, product image, user image) as independent variables, customer perceived value as an intermediary variable, and consumer purchase intention as the dependent variable. An empirical survey was conducted based on this framework. The study provided substantial recommendations for brand image development for sports shoe companies [2].

In 2018, Yuyan Li explored and examined the decorative elements of sports shoe appearance. The study investigated aspects such as toe shape, color combinations, and decorative craftsmanship, concluding that the decorative elements of sports shoes had a positive impact on satisfying consumer psychological needs. This research provided new insights for sports shoe manufacturers in product design [3].

Additionally, the logistic regression model, as a classical statistical learning method, has been employed by several scholars to analyze problems in various domains.

In 2021, Hui Miao conducted an analysis of the determining factors in the e-sports game "League of Legends." By constructing logistic regression models and others, the study assessed the impact of factors such as tower defense structures, neutral jungle resources, map vision conditions, and combat information on game outcomes. This data-driven approach offered innovative insights for the development of e-sports in "League of Legends" [4].

In 2023, Xin Zhao conducted research on consumer purchase intentions for prefabricated food and their influencing factors. Through surveys and empirical analysis using binary logistic regression models, the study examined consumer intentions to purchase prefabricated food and the significance of influencing factors. This research provided theoretical support and recommendations for the future development of the prefabricated food industry [5].

2.2. Logistic Regression Model

Logistic regression is a generalized linear regression model widely used in fields such as data mining, automatic disease diagnosis, and economic forecasting. It is employed to explore underlying risk factors that may lead to diseases and predict the probability of disease occurrence based on these factors. In practical applications, binary classification logistic regression is the most common. In certain regression problems, the response variable is binary, meaning there are only two possible outcomes, such as success or failure. The normal linear model is not suitable for this scenario as it cannot handle binary response variables appropriately. In such cases, logistic regression is a highly suitable method [6].

For the response variable $Y$, there are $p$ independent variables (also known as explanatory variables), denoted as $X_1, X_2, \ldots, X_p$, and the conditional probability of success under the action of $p$
independent variables is denoted as \( P = P(Y = 1|X_1, X_2, \ldots, X_p) \). Therefore, the logistic regression model is:

\[
P = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p)}
\]  

(1)

Among them, \( \beta_0 \) is the constant term, \( \beta_1, \beta_2, \ldots, \beta_p \) is the regression coefficient of the logistic model. It can be seen that the logistic regression model is a typical nonlinear regression model. We perform a logistic transformation on the above equation, which can be transformed into the following linear form:

\[
\logit(P) = \ln \left( \frac{P}{1-P} \right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p
\]  

(2)

Parameter estimation can be done using linear regression models here, so generalized linear models include logistic regression models.

3. Research Hypothesis

As the diversity in the sports shoe market continues to grow, consumers worldwide face a plethora of choices when purchasing sports shoes. Each individual's needs and preferences vary, leading to the emergence of a wide range of brands, styles, and price ranges in the market. Consumers can select suitable sports shoes based on factors such as their athletic activities, personal style, budget constraints, and health considerations [7]. In this research, we propose several hypotheses regarding the factors influencing consumer preferences for purchasing sports shoes.

Hypothesis 1: For any product, price is a significant factor that consumers consider when making a purchase. This research anticipates that lower prices will encourage more consumers to buy sports shoes.

Hypothesis 2: The dominant color scheme of sports shoes may be an aesthetic factor that consumers consider when making a purchase. This research speculates that brightly colored sports shoes may attract a larger number of consumers.

Hypothesis 3: Different types of outsoles in sports shoes possess distinct advantages. This research predicts that sports shoes equipped with air-cushioned outsoles will be more popular among consumers.

4. Research Methods

4.1. Design of Variables

In this research, to delve into the preferences and decision factors of global consumers when purchasing sports shoes, we have selected a series of key variables. These variables will be used to construct a logistic regression model to study their impact on sports shoe sales [8]. These variables include both dependent and independent variables, and here are their detailed descriptions:

1) Sales (Dependent Variable)

Sales are the primary dependent variable in this research. Since the logistic regression model is primarily used for binary classification problems, this research selects shoes with actual sales above the average as 1 and shoes with actual sales below the average as 0.

2) Price (Independent Variable \( X_1 \))

The price of a product plays a significant and undeniable role in shaping consumers' willingness to purchase. Consumers often evaluate the material value of a product based on its pricing and weigh price against factors like product quality, level of demand satisfaction, and more during their purchase decisions. Therefore, the price of an item directly relates to whether consumers are willing to buy it and their ultimate purchasing choices. In a fiercely competitive market, businesses need to consider pricing strategies that enhance purchasing intent while meeting consumer expectations of value, thus
promoting sales and market share growth. In this research, retail prices higher than the average are selected as 1, and retail prices lower than the average are selected as 0.

3) Dominant Color Scheme (Independent Variable $X_2$)

The predominant color scheme of sports shoes plays a crucial role in shaping consumers’ purchasing intent. Consumers often pay high attention to the color of shoes because color is not only a visual perception but is often intertwined with emotions, tastes, and personal image factors. The predominant color scheme of shoes can convey product characteristics, style, and target audience positioning, thereby resonating with consumers. In this research, shoes are categorized based on color scheme, with predominantly bright colors selected as 1 and predominantly dark colors selected as 0.

4) Outsole Type (Independent Variable $X_3$)

The outsole, as a critical component of sports shoes, not only affects the appearance of the shoes but also directly influences comfort, functionality, and usage scenarios. When purchasing sports shoes, consumers often carefully consider the material, design, and features of the outsole. Currently, the sports shoe market mainly consists of two types of outsoles: innovative air-cushioned outsoles and traditional rubber outsoles. This research selects shoes with air-cushioned outsoles as 1 and shoes with traditional rubber outsoles as 0.

5) Upper Type (Independent Variable $X_4$)

Different upper types in sports shoes provide varying degrees of support, snugness, and stability for wearers to cater to various sports activities and personal preferences. Additionally, the appearance of the upper is closely linked to personal fashion taste. Therefore, thoughtfully selecting the appropriate upper type can guide purchasers towards forming positive purchase intentions while meeting their comfort and performance requirements. In this research, high-upper shoes are categorized as 1, while low-upper shoes are categorized as 0.

6) Presence of Strong Contrast Color Elements or not (Independent Variable $X_5$)

The design of strong contrast color elements, as seen in some of the Nike and Off-White’s collaborative sports shoes, can enhance visual appeal and personalization. For example, some predominantly white shoes may feature a contrasting black Nike logo, while some predominantly dark-colored shoes may have vibrant stripes. Contrast colors can highlight the design and details of the shoes, arousing consumer interest. However, some consumers may prefer shoes without strong color conflicts. This research selects shoes with strong contrast color elements as 1, and shoes without strong contrast color elements as 0.

4.2. Data Collection and selection

Kaggle (www.kaggle.com) is an online platform catering to data scientists, machine learning engineers, and data analysts, providing users with a rich repository of datasets from various domains. Users can explore, download, and analyze these datasets for academic research, practical projects, and exercises.

For this research, sales data for 30 Nike and Off-White collaborative sports shoe models, sold in all 50 states of the United States from September 1, 2017, to February 13, 2019, were obtained from Kaggle. Variable selection for the dependent and independent variables was made by comparing the actual sales to the average sales for each shoe model. Additionally, criteria were applied to determine the variables related to price, dominant color scheme, outsole type, upper type, and the presence of strong contrast color elements.

In the end, this research obtained a dataset comprising 30 rows, encompassing one dependent variable and five independent variables. Below (Table 1), a partial display of the data is presented:
4.3. Model Setting and Calculating

To begin with, this research examines the existence of multicollinearity among the five independent variables using the Variance Inflation Factor (VIF).

Multicollinearity refers to the presence of linear relationships among independent variables, where one independent variable can be a linear combination of one or more other independent variables. If multicollinearity exists, the matrix used for calculating the partial regression coefficients of the independent variables becomes non-invertible. This can lead to inconsistencies between the results of the variance analysis of the entire model and the tests of the regression coefficients of individual independent variables, ultimately undermining the effectiveness of the model.

The Variance Inflation Factor (VIF) measures the severity of multicollinearity by comparing the variance of the regression coefficient estimates to the variance when the independent variables are assumed to be linearly uncorrelated. If multicollinearity is indeed present, VIF is used to gauge its severity. The value of VIF should be greater than 1 and approach 1; the smaller the VIF, the lower the degree of multicollinearity, and conversely, the larger the VIF, the higher the degree of multicollinearity. In academia, the typical threshold for identifying multicollinearity is set at 10. When VIF < 10, it is generally considered as an absence of multicollinearity. When 10 ≤ VIF < 100, the degree of multicollinearity is considered high, and when VIF ≥ 100, multicollinearity is deemed very severe.

Table 1. Partial Data Display.

<table>
<thead>
<tr>
<th>Sports Shoes</th>
<th>Sale</th>
<th>Price</th>
<th>Color</th>
<th>Outsole</th>
<th>Upper</th>
<th>Contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nike-Zoom-Fly-Off-White-Pink</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nike-Zoom-Fly-Off-White-Black-Silver</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Nike-Zoom-Fly-Off-White</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nike-Zoom-Fly-Mercurial-Off-White-Total-Orange</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Nike-Zoom-Fly-Mercurial-Off-White-Black</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Nike-React-Hyperdunk-2017-Flyknit-Off-White</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Nike-Blazer-Mid-Off-White-Wolf-Grey</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Nike-Blazer-Mid-Off-White-Grim-Reaper</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Nike-Blazer-Mid-Off-White-All-Hallows-Eve</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2. VIF values of each Variable.

<table>
<thead>
<tr>
<th>Variables</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>X₁</td>
<td>1.206886</td>
</tr>
<tr>
<td>X₂</td>
<td>1.625227</td>
</tr>
<tr>
<td>X₃</td>
<td>1.418348</td>
</tr>
<tr>
<td>X₄</td>
<td>1.247043</td>
</tr>
<tr>
<td>X₅</td>
<td>1.484090</td>
</tr>
</tbody>
</table>

From Table 2, it is evident that the VIF values for the five independent variables are all close to 1. This indicates that there is no multicollinearity among the selected variables, and it will not negatively impact the effectiveness of the model. Therefore, further operations can be carried out.

Next, this research randomly selects data from 8 out of the 30 sports shoe models, representing approximately 25% of the data, as the test set, while the remaining data from the remaining 22 shoe models are used as the training set. A logistic regression model is constructed using the training set's data. This trained model is then used to make predictions on the test set, and the predictions are compared to the actual values of the dependent variable in the test set to determine accuracy.

Simultaneously, this research employs ROC (Receiver Operating Characteristic) curve analysis to visualize and assess the model's predictive performance. The ROC curve is a commonly used tool for analyzing classifier performance. It is constructed by plotting the false positive rate (FPR) on the horizontal axis and the true positive rate (TPR) on the vertical axis. Each classifier's performance is mapped as a point on the ROC plane, and adjusting the classifier's threshold results in different points.
forming the ROC curve. Ideally, the curve should lie above the line connecting (0, 0) and (1, 1), representing better performance than a random classifier. If the curve falls below, the classification results can be improved by reversing them. To quantify classifier performance, the Area under the ROC Curve (AUC) is typically used, representing the area covered by the ROC curve. The AUC value ranges from 0.5 to 1.0, with larger values indicating better performance.

Table 3. Indicators of the Model's performance.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>87.5%</td>
</tr>
<tr>
<td>AUC</td>
<td>0.8</td>
</tr>
</tbody>
</table>

From Table 3, it is evident that the model achieved an accuracy of 87.5%, with an AUC value of 0.8. Therefore, this research can conclude that the logistic model has performed well and can be used to explain the relationship between various factors and consumer purchase intentions.

4.4. Research Results

Based on the results of the constructed logistic regression model, this study obtained the logistic linear regression equations for five Feature variables as follows (Table 4):

$$\text{Logit}(P) = 0.9386 - 0.8195 \times X_1 - 0.7653 \times X_2 - 1.3969 \times X_3 + 0.5566 \times X_4 - 0.8807 \times X_5$$

Table 4. Feature coefficients of each variable.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Feature coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>-0.8195</td>
</tr>
<tr>
<td>$X_2$</td>
<td>-0.7653</td>
</tr>
<tr>
<td>$X_3$</td>
<td>-1.3969</td>
</tr>
<tr>
<td>$X_4$</td>
<td>0.5566</td>
</tr>
<tr>
<td>$X_5$</td>
<td>-0.8807</td>
</tr>
</tbody>
</table>

Based on the coefficients of the respective independent variables, this study can derive the following results:

The importance of variables affecting consumers' purchase intentions for sports shoes, ranked from highest to lowest, is as follows: Outsole Type, Presence of Strong Contrast Color Elements, Price, Dominant Color Scheme, and Upper Type.

The coefficient for $X_3$ is -1.3969, indicating that consumers tend to prefer traditional rubber sole sports shoes over air-cushioned sole athletic shoes.

The coefficient for $X_5$ is -0.88077, suggesting that consumers may favor sports shoes with a consistent color scheme that lacks strong contrast color elements.

The coefficient for $X_1$ is -0.8195, demonstrating that sports shoes with lower prices are more likely to elicit purchase intentions from consumers.

The coefficient for $X_2$ is -0.7653, revealing an interesting phenomenon: sports shoes with darker dominant color schemes are more likely to be popular among consumers compared to brighter ones.

The coefficient for $X_4$ is 0.5566. Although the influence of Upper Type on consumer purchase intentions is not as significant as other factors, we can still conclude that high-top athletic shoes are more likely to achieve high sales compared to low-top athletic shoes.

4.5. Research Discussion

Based on the coefficients of the various independent variables, this research will analyze consumers' preferences for sports shoes from the perspective of shoe design and consumer psychological preferences.

Regarding the outsole type, air-cushioned soles typically contain air chambers filled with nitrogen or air, which can absorb and distribute impact forces from the ground, reducing foot fatigue and providing a more comfortable feel. However, despite continuous advancements in air-cushioning
technology, air-cushioned sole sports shoes may still have slightly lower durability than traditional sole sports shoes in some cases. Their air chambers can be affected by wear and external conditions, leading to leaks or damage. Additionally, due to their transparent appearance, air-cushioned sole sports shoes may not align with the aesthetic preferences of some consumers. On the other hand, traditional rubber sole sports shoes, while possibly lacking the same level of cushioning and lightweight feel as air-cushioned soles, excel in durability and grip. They offer better stability and support, suitable for various sports activities from everyday walking to indoor and outdoor sports. In terms of aesthetics, traditional rubber sole sports shoes can incorporate a wider range of color design elements. Therefore, most consumers prefer sports shoes equipped with traditional rubber soles [9].

Concerning whether sports shoes have strong contrast color elements, even though the design with strong contrast colors may visually add complexity and impact to the shoes, it appears that consumers lean towards sports shoes with a more balanced and harmonious color scheme. Consumers who favor highly individualistic and attention-grabbing designs, showcasing their unique tastes in footwear, do not represent the majority.

Price often correlates with consumers' perception of a product's value, and lower prices may make consumers feel they are getting greater value for their money, thus increasing their desire to purchase. When faced with multiple options, a lower price can make one pair of shoes appear more attractive when compared to competitors. Additionally, many consumers have budget constraints and need to make purchase decisions within limited budgets. Lower prices make products more accessible within consumers' budget constraints. Therefore, setting a reasonably lower price is crucial for increasing the sales of a pair of sports shoes.

Regarding the dominant color scheme of sports shoes, darker color schemes may convey a sense of stability and understatement, which aligns with some consumers' attitudes towards their purchases. Many consumers may prefer shoes that are not attention-grabbing and flamboyant, making them easier to coordinate with daily outfits. Moreover, many sports shoes with darker color schemes are considered classics in the fashion industry and can withstand the test of trends. This may lead some consumers to favor this type of footwear.

Additionally, from the perspective of the upper type, while there may be differences in comfort, high-top sports shoes often offer better ankle support, which is crucial for certain sports like basketball and soccer. Consumers who are enthusiastic about such sports may lean towards high-top sports shoes to meet their needs for foot stability and protection during physical activities [10].

5. Conclusion

Through the construction of a logistic regression model, this research has drawn several important conclusions. Firstly, the type of outsole significantly influences consumers' purchase intentions, with traditional rubber sole sports shoes being more popular, possibly due to their durability and versatility. Secondly, consumers tend to choose sports shoes with a balanced color scheme, emphasizing overall harmony over designs with strong contrast color elements. Thirdly, price plays a crucial role in consumers' purchase decisions, as lower prices can attract more buyers. Additionally, sports shoes with darker color schemes are more favored, possibly due to their understated appearance and classic appeal. Lastly, the type of upper is also a consideration for some consumers, as high-top sports shoes may be more suitable for those in need of ankle support during physical activities.

However, this research has some limitations. Firstly, the research dataset only includes sales data for 30 Nike and Off-White collaborative sports shoe models, limited to this specific brand and collaboration series. Future research can expand the sample range to encompass more brands and types of sports shoes for a comprehensive understanding. Secondly, while the logistic regression model has provided useful information for this study, it is primarily used for analyzing binary classification problems, and its accuracy in analyzing factors that are commonly continuous variables, such as price and sales, is not very high. Furthermore, this research has not considered the impact of
individual differences and cultural factors on purchase intentions, which could be addressed in future research.

Future studies can also explore more factors, such as brand reputation, social media influence, and personal lifestyles, to provide a more comprehensive analysis of consumers' sports shoe purchase decisions. Additionally, the use of machine learning and artificial intelligence technologies can enable more accurate predictions of consumer purchase intentions, aiding sports shoe brands in better meeting market demands.

In summary, this research provides valuable insights into the sports shoe market and consumer purchasing behavior. Despite some limitations, our study offers useful guidance for future research directions, helping to better understand and cater to consumer needs, and driving the continuous development of the sports shoe market.

References


