Unveiling Customer Dynamics: Data-Driven Strategies for Ecommerce Enhancement

by Christina Gasparinatos

1. Introduction

In the fast-paced world of modern commerce, businesses are continually seeking innovative ways to enhance their customer experiences and drive growth. This project centers around an Ecommerce company that specializes in clothing sales through both online platforms and in-store style and clothing advice sessions. The company's unique approach combines personalized in-store consultations with the convenience of online shopping. Customers have the opportunity to engage with personal stylists during in-store sessions and later make their purchase decisions through either a mobile app or a website.

The overarching challenge faced by the company revolves around optimizing their digital platforms to better align with customer preferences and behaviors. Specifically, the company aims to determine whether their focus should be directed towards refining the mobile app experience or enhancing the usability of their website.

This project delves into an in-depth data analysis and machine learning techniques to provide actionable insights for the company's decision-making process. By leveraging available historical customer data, this analysis aims to uncover patterns, preferences, and trends exhibited by customers using both the mobile app and the website. Ultimately, the goal is to provide evidence-based recommendations that will guide the company towards effectively allocating resources and prioritizing efforts.

2. Data Description

The dataset provides comprehensive insights into the customers of the Ecommerce company, encompassing both personal and numerical attributes. Alongside customers' personal identifiers, such as Email, Address, and Color Avatar, the dataset presents a set of crucial numerical columns that contribute to a comprehensive understanding of customer behavior and engagement. The numerical attributes enable quantitative analysis of customer interactions and spending patterns. The dataset includes the following key numerical columns:

- Avg. Session Length: Average session of in-store style advice sessions.
- Time on App: Average time spent on App in minutes.
- Time on Website: Average time spent on Website in minutes.
- Length of Membership: How many years the customer has been a member.
- Yearly Amount Spent: The amount of money the customer spends on the company annually.

The dataset contains a total of 500 entries, each representing a distinct customer. It is through the analysis of these numerical attributes that the company can gain valuable insights into customer preferences, engagement trends, and spending behaviors.

3. Exploratory Analysis

Initially, a comprehensive data exploration will be conducted to establish familiarity with the dataset's attributes. Notably, the average annual spending by customers is approximately $500, highlighting their financial contribution to the company's revenue stream. Furthermore, the customers' average membership duration stands at 3.5 years, indicating a level of loyalty and sustained engagement with the brand.

An intriguing insight emerges from the dataset's attributes, specifically related to the customers' engagement with the company's platforms. Figure 1 showcases that, on average, customers spend nearly three times more time on the company's mobile app compared to the time spent on the website. This disparity underscores the significance of the mobile app as a preferred platform for customer interaction and product exploration.

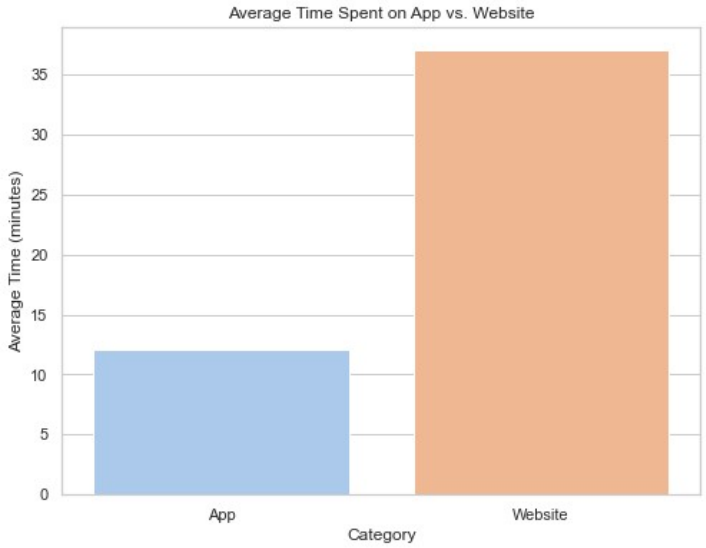

Figure 1: Average Time Spent on App vs. Website

Figure 2Figure 3 illustrate the relationship of yearly amount spent with the time on website and time on app, respectively. The points in the first graph are more spread, indicating varying spending behaviors among customers. Contrarily, the points in the second graph appear to be more centralized, indicating a potential correlation between increased app usage and higher annual spending. This positive correlation underscores the pivotal role of the mobile app in driving customer engagement and spending behaviors.

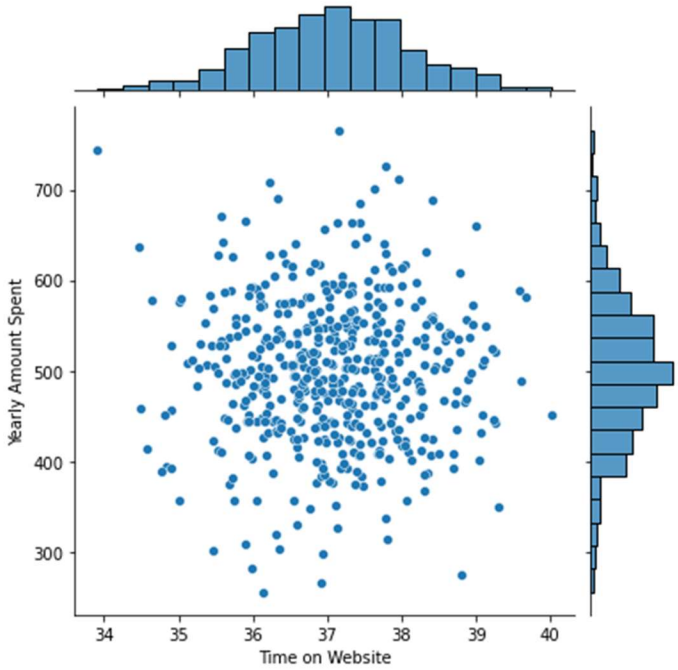

Figure 2: Jointplot of Yearly Amount Spent and Time on Website

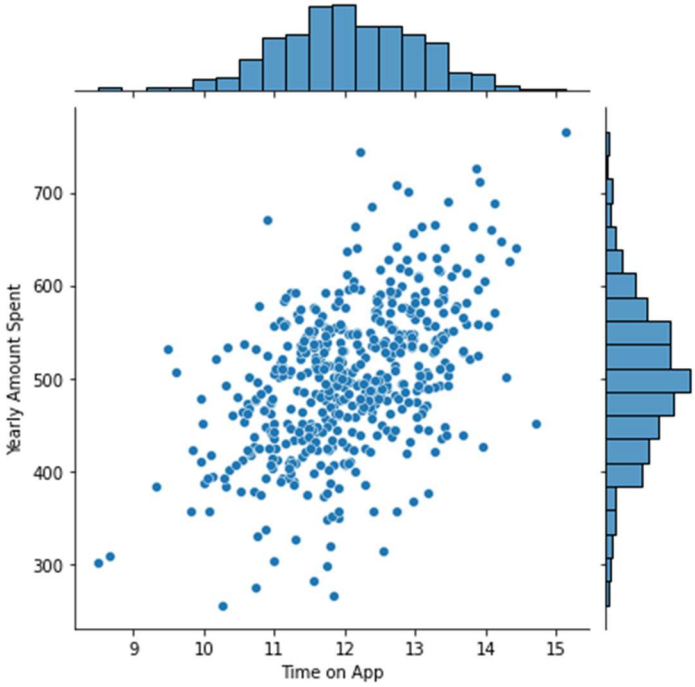

Figure 3: Jointplot of Yearly Amount Spent and Time on App

The correlation heatmap presented in Figure 4 serves as a valuable tool for comprehending the relationships between various variables and the yearly amount spent by customers. Notably, the variable that exhibits the strongest correlation with the yearly amount spent is the length of

membership. This suggests that customers who have been members for a longer duration tend to spend more on an annual basis.

Subsequently, the heatmap reveals the following in descending order of correlation magnitude: the time spent on the mobile app, and the average session length during in-store style advice sessions. These variables demonstrate notable but comparatively lesser correlations with the yearly amount spent. Conversely, the time spent on the website appears to exhibit a weak correlation with the yearly amount spent.

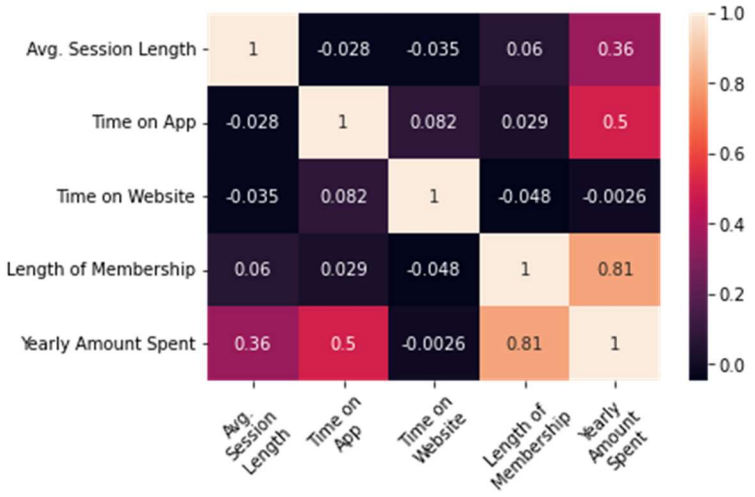

Figure 4: Correlation Heatmap

4. Methodology

Given the objective of understanding the relationship between different variables and the yearly amount spent by customers, linear regression was chosen as the most suitable technique. Linear regression is particularly adept at identifying and quantifying the linear relationship between dependent and independent variables. In this context, the dependent variable is the yearly amount spent by customers, while the independent variables encompass attributes such as the length of membership, time spent on the app, time spent on the website, and the average session length during in-store style advice sessions. Linear regression allows us to quantify this impact by assigning coefficients to each independent variable, indicating how much the dependent variable changes for a unit change in the independent variable, while holding other variables constant. Finally, linear regression offers interpretability, which is paramount for gaining actionable insights.

5. Results

This section presents the outcomes of the linear regression analysis conducted on the dataset. The primary objective was to unravel the relationship between various predictor variables and the yearly amount spent by customers.

The scatter plot (Figure 5) comparing the predicted and actual values of the annual amount spent reveals a strong linear relationship, indicating that the model's predictions closely align with the true values. The plot exhibits a pattern where most points are situated near the line y=x, suggesting that the model's predictions are generally accurate and consistent with the actual spending patterns.

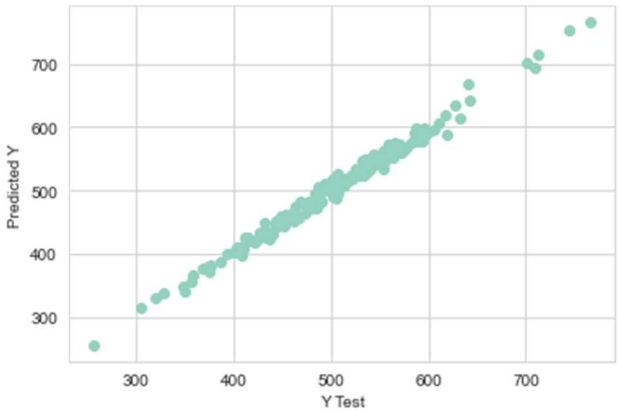

Figure 5: Plot of Real and Predicted Yearly Amount Spent

The evaluation metrics provide further insight into the model's performance:

- Mean Absolute Error (MAE) = 7.23. This means that, on average, the model's predictions for the yearly amount spent are off by $7.23 from the actual values.
- Mean Squared Error (MSE) = 79.81. MSE focuses on larger errors and provides a comprehensive view of the model's overall performance.
- Root Mean Squared Error (RMSE) = 8.93. This indicates that, on average, the model's predictions for the yearly amount spent deviate by approximately $8.93 from the actual values.

The histogram of residuals (Figure 6) – a measure of the differences between actual and predicted values – displays a normal distribution, indicative of a well-fitted model. The normal distribution suggests that the model's errors are symmetrically distributed around zero, further confirming the appropriateness of the linear regression approach.

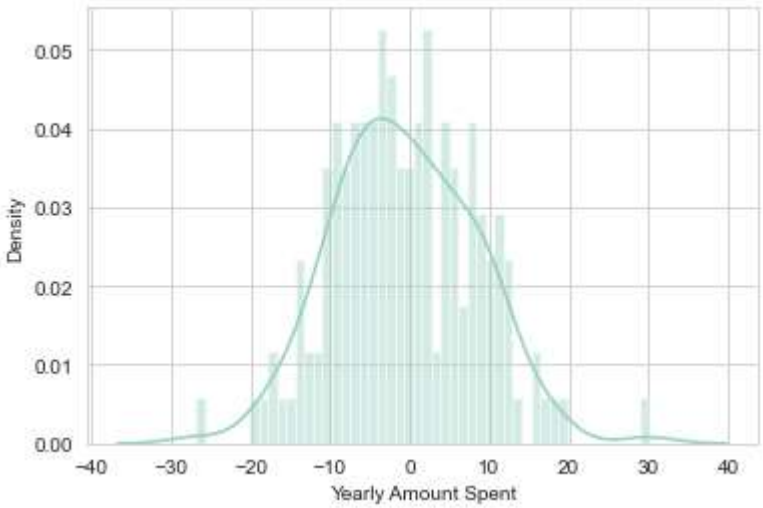

Figure 6:Histogram of Residuals

The estimated coefficients for the independent variables shed light on their respective impacts on the yearly amount spent:

- Avg. Session Length: 25.98. This suggests that, on average, a one-unit increase in the average session length during in-store style advice sessions is associated with an approximate increase of $25.98 in the yearly amount spent by customers.
- Time on App: 38.59. It is evident that spending more time on the mobile app is associated with a more substantial increase in the yearly amount spent compared to other variables.
- Time on Website: 0.19, which implies a minor influence of time spent on the website on the yearly amount spent.
- Length of Membership: 61.28. It is the largest coefficient, which underscores the significance of customer loyalty, indicating that each additional year of membership corresponds to a significant increase in the yearly amount spent.

These coefficients offer actionable insights for the company's decision-making process. The model highlights that the length of membership and time spent on the app have the most pronounced impact on customers' annual spending habits. This information is crucial for the company's strategic planning, allowing them to allocate resources effectively and tailor their efforts to maximize customer engagement and loyalty.

6. Conclusion

This project has navigated through an in-depth analysis of an Ecommerce company's customer data to derive meaningful insights that can steer the company's future directions. By delving into customer behavior, preferences, and engagement patterns, this analysis aims to guide the company's strategic choices regarding its digital platforms and customer interactions.

Based on the insights gained from the analysis, the following recommendations can guide the company's decisions:

- Enhanced Mobile App Experience: Given the strong positive correlation between time spent on the mobile app and yearly spending, the company should prioritize enhancing the mobile app experience. This can include improvements in user interface, personalization features, and seamless navigation, all of which can encourage extended app usage and drive higher spending.
- Leverage Customer Loyalty: The length of membership exhibited the most substantial impact on yearly spending. The company should explore strategies to further nurture and reward customer loyalty. Tailored promotions, exclusive offers, and personalized experiences can foster stronger relationships and bolster spending habits.
- Website Optimization: Although the impact of time spent on the website on yearly spending is relatively minor, optimizing the website has the potential to enhance new customer engagement.
- Holistic Approach: Rather than an exclusive focus on the mobile app or website, the company can adopt a holistic approach. Integrating the strengths of both platforms can create a seamless omnichannel experience, allowing customers to transition seamlessly from in-store consultations to online purchases.

By unraveling customer behaviors and uncovering relationships between variables, the company gains a competitive edge in crafting strategies that resonate with its customer base.