Analyzing Consumer Behavior for Marketing Insights

Author: Georgia Iferi Itam

School: Ivy Tech Community College

Course: DBMS 260 (Advanced Data Analytics)

Instructor: Jimmie Flores

Date: 9/2/25

Introduction

This project analyzes consumer behavior to identify patterns that inform marketing strategies. The dataset was retrieved from Kaggle containing demographic, financial, and behavioral variables, including age, annual income, estimated savings, credit score, loyalty years, and preferred shopping categories. The purpose of this project is to observe the consumer behavior patterns and identify meaningful customer segments that can inform targeted advertising and retention strategies. The analysis will include exploring the data and creating visualizations in Tableau to show trends and groups of customers. While the dataset is already clean, one challenge may be making sure the customer groups are meaningful and not just random patterns. To address this, I will use analysis and visualization techniques to confirm results. The project is expected to show how businesses can use customer data to improve advertising, personalization, and long-term relationships.

Materials and Methods

The dataset originates from Kaggle and includes demographic (Gender, Age, Age Group), financial (Annual Income in thousands, Estimated Savings in thousands, Credit Score), and behavioral fields (Loyalty Years, Preferred Category, Spending Score). Data preparation focused on consistency and transparency. Duplicates were removed; four missing values in Age Group were labeled “Unknown” to retain records; numeric fields (Age, Annual Income, Spending Score, Credit Score, Loyalty Years) were stored as whole numbers; Estimated Savings was formatted to one decimal place; categorical fields (Gender, Age Group, Preferred Category) were standardized to proper case. Annual Income is kept in thousands (e.g., 15 = $15,000). Analysis used Excel for cleaning, StatCrunch for descriptive statistics and boxplots, and Tableau for visual exploration and hypothesis testing. Visuals were exported as images and assembled alongside text to support interpretation.

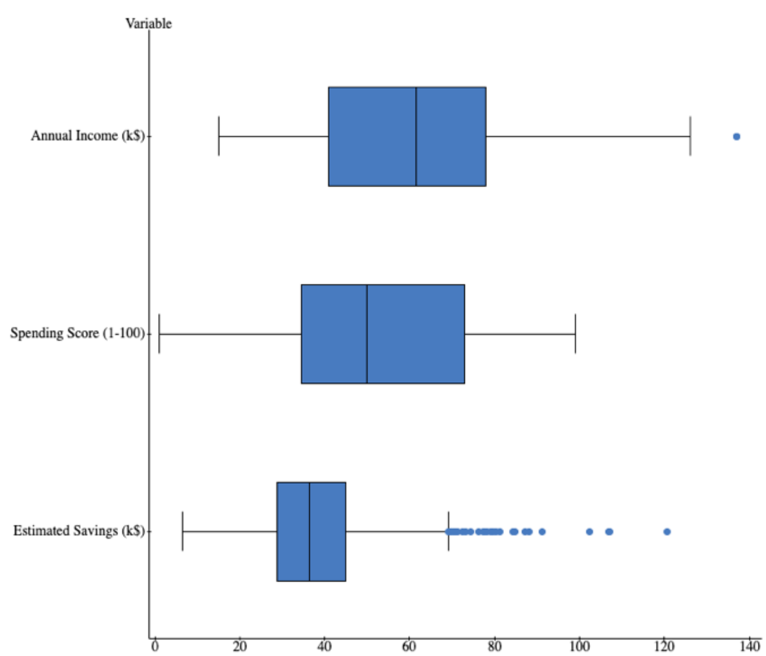
Results

Descriptive statistics summarize central tendencies and ranges for the numeric variables and confirm that values lie within realistic bounds (see Summary Statistics table). Boxplots highlight distribution shapes and flag a small number of high savings values at the upper end; these appear as outliers but remain plausible.

Table 1: Summary Statistics for Customer Variables

Variable	Mean	Median	Std. Dev.	Min	Max
Age	38.9	36	14.0	18	70
Annual Income (k$)	60.6	61.5	26.3	15	137
Spending Score (1–100)	50.2	50	25.8	1	99
Estimated Savings (k$)	40.2	36.4	21.3	6.5	120.6
Credit Score	743.7	833	155.2	300	850
Loyalty Years	5.9	6	1.6	2	9

Boxplot



Brief Interpretation of Boxplot

Looking at the numbers and the boxplots, the dataset is largely clean and realistic. It is key to note that there are a small number of very high savings values (around 120k+) are present; I treated these as outliers to monitor rather than errors. Credit scores spanned the expected range (300–850), which supports data validity, and loyalty years average around 6 years, which is a really positive sign. It shows that many customers stick with the brand long-term.

Discussion

The analysis was guided by three key hypotheses, each aimed at uncovering relationships between customer demographics, financial factors, and shopping behavior. The goal was not only to test whether patterns exist, but also to understand how businesses might act on these insights for more effective marketing.

The first hypothesis asked whether gender influences credit score rates. A bar chart of average credit scores by gender clearly showed that female customers maintain higher credit scores on average compared to male customers (see Figure 1). This suggests that, within this dataset, women may generally manage credit more successfully. From a marketing perspective, this insight could be useful for shaping targeted financial programs, such as premium credit cards, financing plans, or loyalty offers, toward female shoppers who appear to have stronger credit profiles.

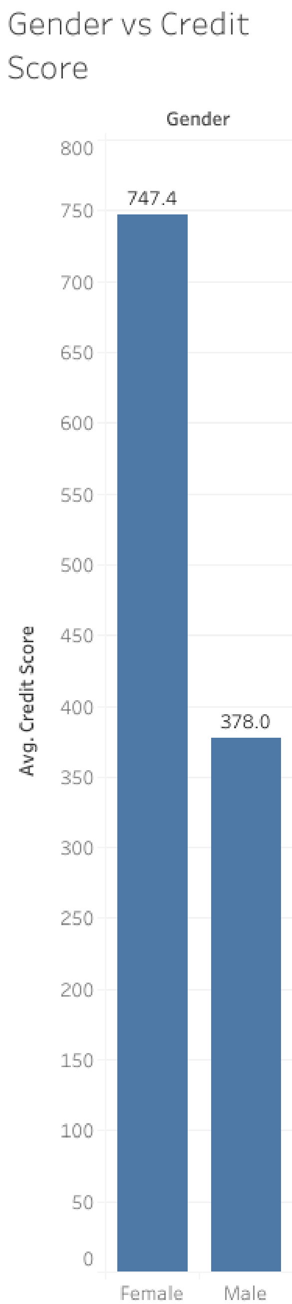


Figure 1: Average Credit Score by Gender

The second hypothesis examined whether gender affects luxury purchases. Filtering for the luxury category revealed that female customers are more likely to prefer luxury products compared to male customers (see Figure 2). This result provides a clear direction for targeted advertising: luxury campaigns could be tailored more toward women, while marketing for other categories such as budget or general items should remain more balanced across genders. By recognizing these differences, businesses can allocate advertising resources more effectively and design messaging that resonates with the right audience.

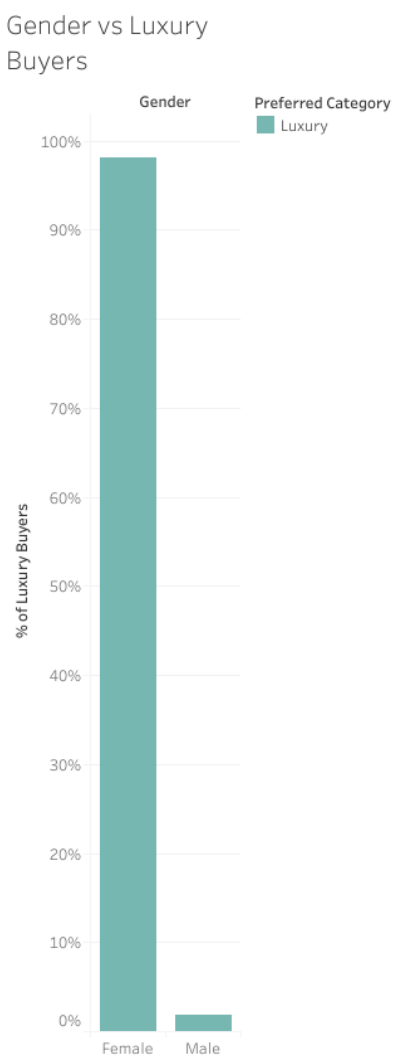


Figure 2: Luxury Purchases by Gender

The third hypothesis focused on whether annual income influences spending scores and preferred shopping categories. The scatterplot of income versus spending scores revealed no strong linear relationship between the two (see Figure 3). High income does not automatically mean high spending, as customers across all income levels showed diverse spending patterns. However, when income was grouped into brackets, higher-income customers were more often associated with luxury purchases, while lower-income groups tended toward budget or general categories. These results suggest that spending behavior is not purely financial but shaped by personal preferences and lifestyle. For marketers, this means that income-based segmentation can be helpful for product positioning (luxury versus budget), but personalized, behavior-driven campaigns are essential to accurately target spending potential.

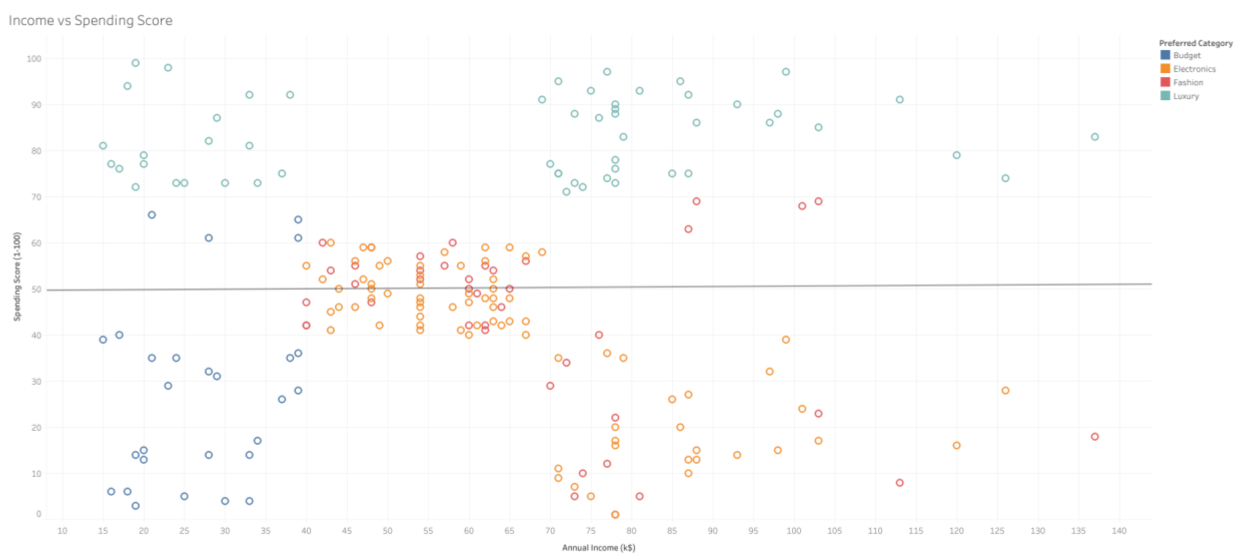


Figure 3: Annual Income vs Spending Score

Conclusion

This project explored how customer demographics and financial factors shape behavior, with a focus on credit strength, luxury preference, and spending patterns. The results showed that women tended to have stronger credit scores, females leaned more toward luxury purchases, and income levels influenced category preference but not spending scores. These findings highlight the importance of combining demographic insights with behavioral data. Businesses that take this dual approach can design more relevant campaigns, deliver personalized offers, and strengthen long-term customer loyalty.

References

Gupta, V. J. (n.d.). Customer analytics practice dataset. Kaggle. https://www.kaggle.com/datasets/vikasjigupta786/customer-analytics-practice-dataset/data