Evaluating the Influence of In-Store Product Placement on Sales: A Statistical Analysis

D'Andrea Donnelly

University of Wisconsin - Madison

PSYCH755

Tim Rogers

August, 11 2024

Abstract

This study examines the effect of product placement within stores on sales volume, focusing on various item types such as clothing, electronics, and food. It explores whether the location of products—aisle, end-cap, or front of store—affects their sales performance. The analysis also incorporates seasonal factors, ongoing promotions, and demographic information. Using a dataset obtained from Kaggle with ten variables and 1,000 observations, the study employs exploratory data analysis (EDA), logistic regressions, random forests, and ANOVA tests. The results indicate limited predictive power of product placement on sales volume. This study found no significant predictors of sales with the given variables.

Introduction

The impact of product placement on sales volume is a critical area of interest in retail management. Marketing research has demonstrated that effective visual merchandising-displays play a vital role in influencing consumer behavior, including the time spent in stores and overall shopping experience. These visual elements can significantly affect customer satisfaction, purchasing decisions, and the duration of store visits (Rituparna et al., 2022). Additionally, in the field of fashion, the rapid evolution of trends necessitates that stores continually update their visual merchandising strategies to meet fast-paced sales goals (Akhilendra and Aravendan, 2023). Given these factors, a key question for companies is how well their current product placement strategies address these evolving requirements.

This project explores how product placements in stores effect profits with a dataset titled "Impact of Product Positioning on Sales". More specifically, it examines if items such as clothing, electronics, and food are affected by their location in stores. These locations include aisle, end-cap, and front of store. The data also include information on what season it is, if there are current promotions taking place, and demographic information.

The dataset used in this analysis was sourced from a public repository gathering site (Kulkarni, 2024) and was downloaded as a .CSV containing ten variables and 1,000 observations. The full list of variables include: "Product ID", "Product Position", "Price", " Competitor's Price", "Promotion", "Foot Traffic", "Consumer Demographic", "Category", "Seasonal", and "Sales Volume". A Repository was made in GitHub to house both the data and the notebook. A notebook was created with VSCode and Jupyter Notebooks, and this final write up is through the use of Overleaf LaTex.

Method

To begin, the data was loaded into the notebook and EDA took place. Variables were explored and adjusted as needed. NAs were searched for, though none existed. Following this, univariate and bivariate statistics were conducted to gain insights into the data and relationships. Though the univariate and bivariate statitistics showed little relationship between variables, linear regressions

and random forests were run to examine predicting variables.

Categorical variables were encoded using one-hot encoding to convert them into numerical data. The dataset was then preprocessed using a pipeline that included scaling the features with StandardScaler and applying a regression model. To evaluate the predictive power of the features on Sales Volume, a linear Regression and a random forest regression were utilized. The Pipeline from Scikit-Learn, incorporating preprocessing with ColumnTransformer and scaling with StandardScaler were used for the linear regression. Model performance was assessed using Mean Squared Error (MSE) and R-squared (R^2) metrics.

To determine the importance of each feature, a random forest regressor was fit to the data. The model's feature importances were extracted and compared to identify which features have the most significant impact on predicting sales volume.

Finally, ANOVA tests were completed to evaluate whether categorical variables such as Product Position, Promotion, Foot Traffic, Consumer Demographics, Product Category, and Seasonal significantly affect sales volume. The ANOVA tests provided F-statistics and p-values to assess the statistical significance of each categorical feature.

Results

The univariate and bivariate exploration revealed several key insights into the factors affecting sales volume. Figure 1 provides a comprehensive overview of various factors affecting sales volume. The box plots demonstrate that promotion significantly increases sales volume, while foot traffic and product positioning show varied impacts. Consumer demographics reveal that young adults and college students contribute to higher sales volumes, whereas seniors and families have a lower impact. Moreover, the product category analysis highlights that food products generally have higher sales volumes compared to electronics and clothing. Finally, seasonality shows a notable influence, with seasonal products achieving higher sales volumes compared to non-seasonal items. Figure 2 examines the impact of product positioning on sales volume, showing that products placed in aisles and at the front of the store exhibit higher sales volumes compared to those on end-caps, suggesting that visibility and accessibility play crucial

roles in driving sales. Additionally, Figure 3 illustrates the relationship between foot traffic and sales volume, indicating that high foot traffic areas tend to have the highest sales volume, followed closely by low and medium traffic areas, though the differences are marginal. Overall, these findings underscore the importance of strategic product placement, targeted promotions, and understanding consumer demographics in optimizing sales performance.

The linear regression yielded an MSE of 508,094.49 and an R^2 value of 0.0145. These results suggest that the linear regression model explains only a small portion of the variance in sales volume, indicating limited predictive power with the given features. The random forest provided insights into feature importance. The most influential features were found to be Competitor's Price (importance: 0.183), Price (importance: 0.153), and Product ID (importance: 0.134). Other significant features included Foot 'Traffic Low' (importance: 0.074) and 'Seasonal Yes' (importance: 0.064). Features like Product Position and Consumer Demographics had relatively lower importances, with Product 'Position End-cap' having an importance of 0.027.

Finally, ANOVA tests revealed that the categorical features generally did not significantly impact sales volume. Specifically, the Product Position feature had an F-statistic of 0.124 and a p-value of 0.883, indicating no significant effect on sales volume. Similarly, other categorical features like Promotion (F-statistic: 0.142, p-value: 0.707), Foot Traffic (F-statistic: 0.573, p-value: 0.564), Consumer Demographics (F-statistic: 1.752, p-value: 0.155), Product Category (F-statistic: 1.903, p-value: 0.150), and Seasonal (F-statistic: 0.759, p-value: 0.384) also showed no significant effect on sales volume.

Conclusion

This study explored the impact of product placement on sales volume, aiming to determine whether factors such as product positioning, promotion, foot traffic, consumer demographics, product category, and seasonality significantly influence retail sales. Despite the theoretical expectations and insights from prior research, the results from both linear regression and random forest models indicate that the selected features only modestly predict sales volume.

The univariate and bivariate analyses provided preliminary insights, revealing that promotions

and specific consumer demographics (e.g., young adults and college students) are associated with higher sales volumes. Moreover, products positioned in aisles and at the front of the store generally showed better performance than those on end-caps, suggesting that visibility and accessibility are important, although these findings were not statistically significant.

However, the predictive power of the linear regression model was limited, as evidenced by a low R^2 value, suggesting that the model explains only a small portion of the variance in sales volume. The random forest model did provide some insights into feature importance, highlighting the significance of factors such as Competitor's Price, Price, and Product ID, however, the overall impact of product positioning and other categorical features was minimal. The ANOVA tests further supported these findings by showing that the categorical variables, including Product Position, Promotion, Foot Traffic, Consumer Demographics, Product Category, and Seasonality, did not have statistically significant effects on sales volume.

In conclusion, while product placement, consumer demographics, and other factors undoubtedly play roles in influencing sales, this study suggests that the specific dataset and features analyzed may not capture the full complexity of the retail environment. Further research, perhaps incorporating additional variables or different modeling approaches, is needed to more comprehensively understand the factors driving sales volume in retail settings. This study highlights the importance of continuous data exploration and the need for retailers to consider a broader range of factors when devising strategies for product placement and overall sales optimization.

References

Akhilendra, S. P., & Aravendan, M. (2023). A review on apparel fashion trends, visual merchandising and fashion branding. *Intelligent Information Management*, *15*(33), 120–159.

Kulkarni, A. (2024). *Impact of product positioning on sales*.

Rituparna, B., Paul, J., & Singh, K. (2022). Visual merchandising and store atmospherics: An integrated review and future research directions. *Journal of Business Research*, *151*, 397–408.

Appendix

Figure 1
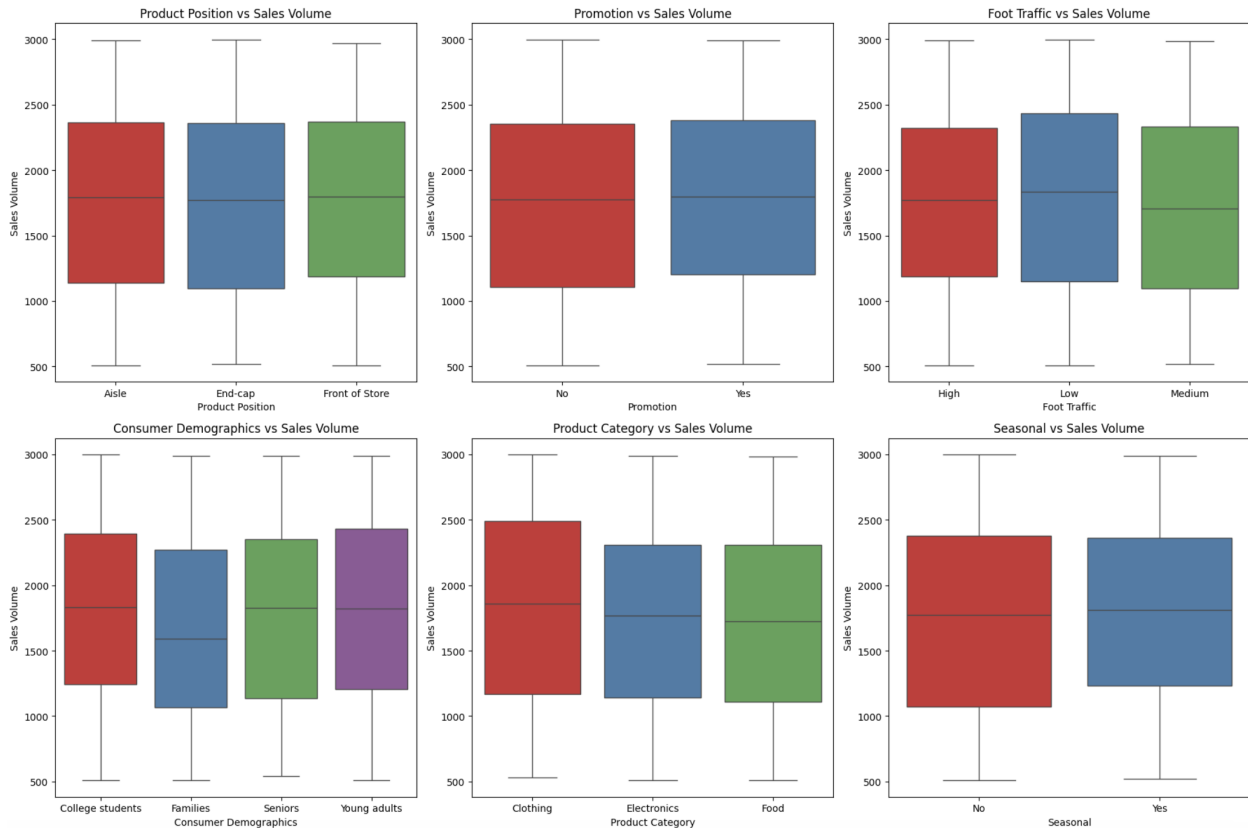
Note. *Box plots showing the relationship between different factors and sales volume. The plots represent the effects of product position, promotion, foot traffic, consumer demographics, product category, and seasonality on sales volume, with each box representing the interquartile range and the whiskers extending to the minimum and maximum values.*

Figure 2
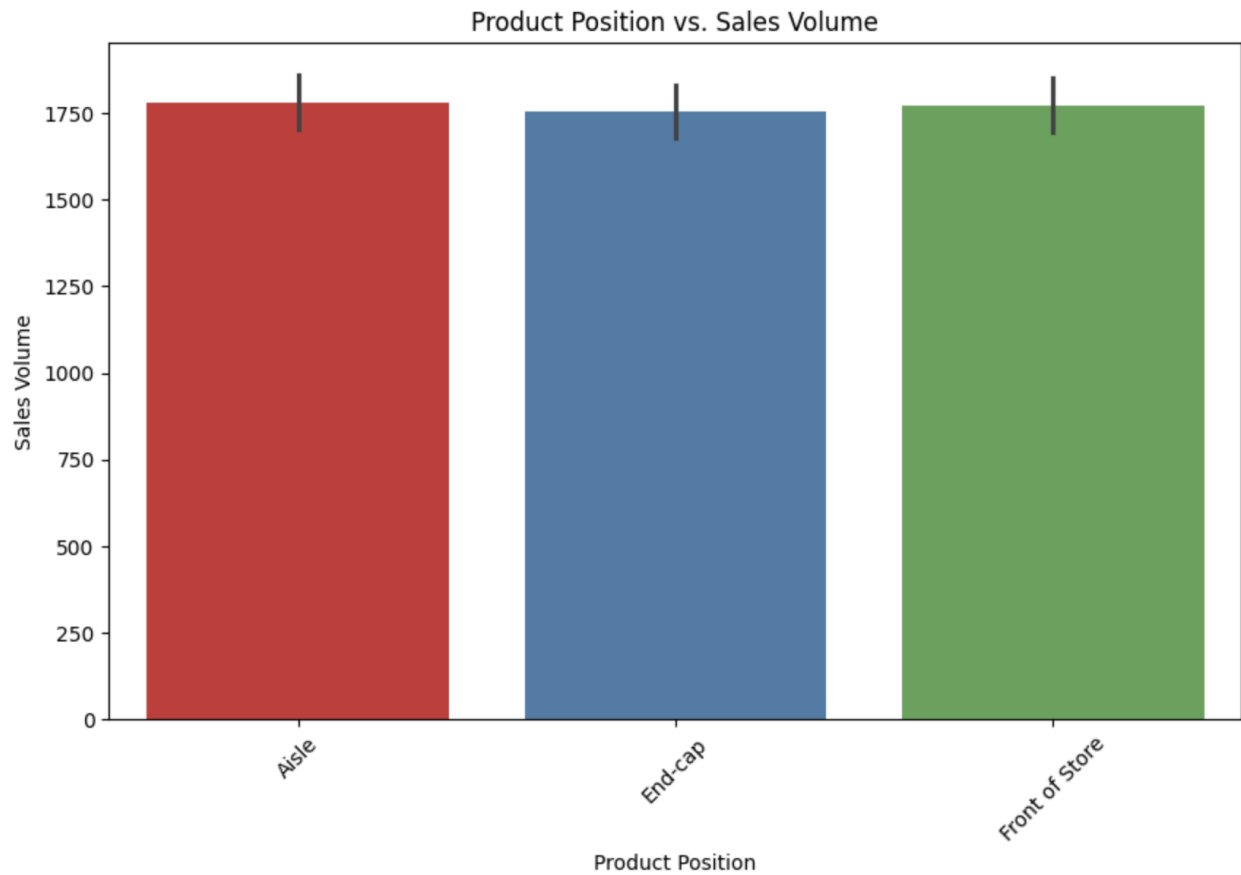
Note. *The impact of product position on sales volume. The bar plot illustrates the average sales volume for products positioned in the aisle, end-cap, and front of the store, with error bars representing the standard error.*

Figure 3
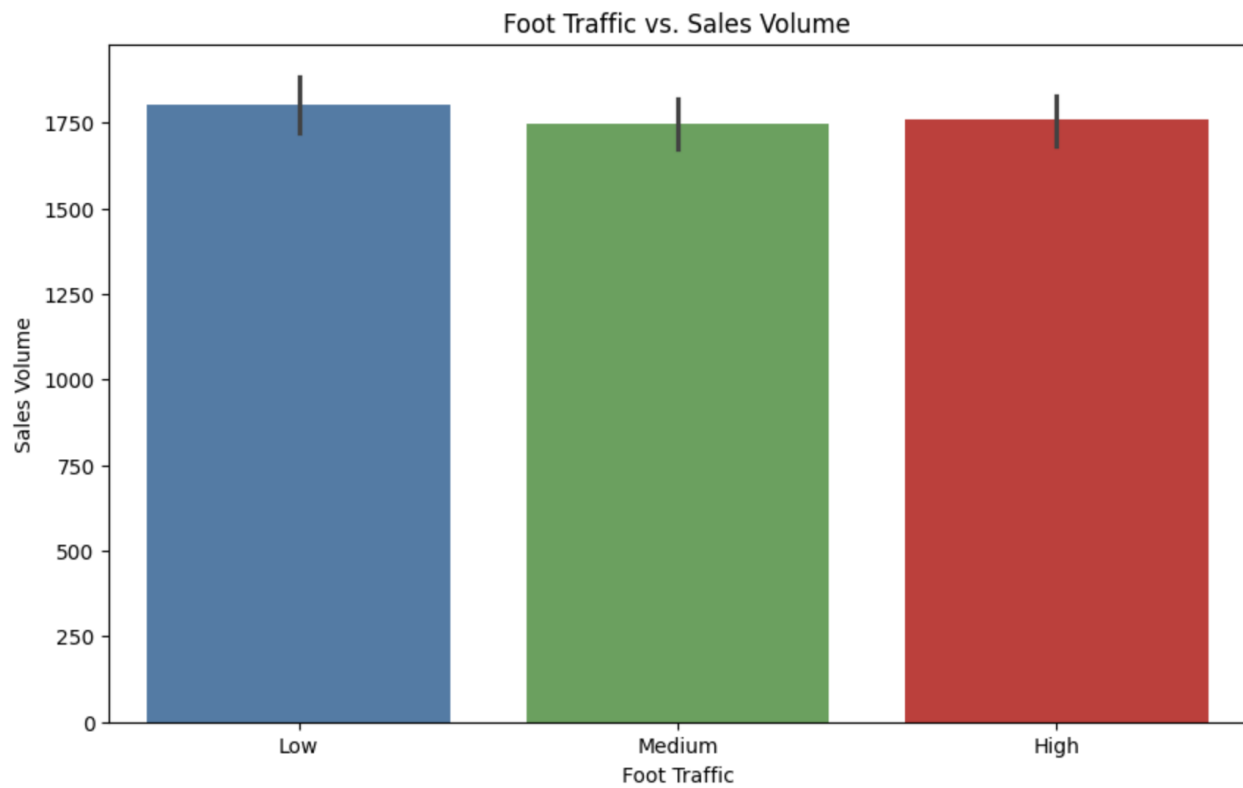
Note. *The relationship between foot traffic and sales volume. The bar plot shows the average sales volume for low, medium, and high foot traffic areas with error bars representing the standard error.*