Proximity as the Key Ingredient in Starbucks' Brand Loyalty

Part - III

Section A: Project Description (As per Part 2 and Part 1 revised)

Our objective is to find out if Starbucks' proximity is the main factor affecting customer loyalty, especially when the most convenient or nearest Starbucks location is no longer an option. The purpose of this study is to see whether consumers will stick with Starbucks even if there are more convenient or nearby options. By looking at loyalty in these particular contexts, we hope to identify the underlying factors that influence customers' dedication to Starbucks and determine how much proximity influences their decision. This study will shed important light on what keeps customers loyal even when it's difficult to get to their favorite Starbucks store.

Method: (Not part of the original project description, added to maintain continuity)

We collected survey responses from over 150 individuals for our planned analysis questions and filtered them to include only those relevant to our statistical analysis, resulting in a final sample size of 72 responses. The dataset captures participants' loyalty to Starbucks, focusing on the relationship between proximity to the nearest outlet and self-reported and behavior-inferred loyalty measures. It includes distance to the nearest Starbucks (distance) as an independent variable and two types of loyalty measures, self-reported and inferred, as dependent variables. A detailed explanation of the dependent and independent variables is provided in Appendix -Part 2.

Section B: Analysis and Results

As discussed in Part-II of the project, we decided to proceed with Correlation Analysis for Distance and Self-Reported Loyalty and Logistic Regression Analysis for Distance and Inferred Loyalty. The analysis for this study was conducted using R, utilizing the base packages.

Correlation Analysis (Distance and Self-Reported Loyalty):

Null Hypothesis (H_0): There is no significant correlation between distance and self-reported customer loyalty.

Alternative Hypothesis (H_1): There is a significant correlation between distance and self-reported customer loyalty.

We conducted a Pearson's correlation test to examine the relationship between distance and self-reported loyalty, which showed a weak negative correlation, $r(70) = -0.18$, $t = -1.51$, $p = 0.136$. Since the p-value is greater than the typical significance level of 0.05, we fail to reject the null hypothesis.

Logistic Regression Analysis (Distance and Inferred Loyalty):
Null Hypothesis (H_0): Distance does not significantly affect inferred customer loyalty.
Alternative Hypothesis (H_1): Distance significantly affects inferred customer loyalty.
We fit a logistic regression model to examine the relationship between distance and inferred loyalty. The coefficient for distance was -0.645 (SE = 0.4118), z = -1.567, with a p-value of 0.117, which exceeded the typical 0.05 significance threshold. Therefore, we fail to reject the null hypothesis.

Part C: Evaluation and Error Analysis:
The results of both the correlation analysis and logistic regression did not support the hypothesis, as we failed to reject the null hypothesis in both tests. In the correlation analysis, although we observed a weak negative correlation, $r(70) = -0.18$, the p-value of 0.136 exceeded the standard significance level of 0.05. This indicates insufficient evidence to support a significant correlation between distance and self-reported loyalty. Similarly, the logistic regression analysis yielded a p-value of 0.117, which also exceeded the 0.05 threshold, suggesting that proximity does not significantly influence inferred loyalty.
Given that we failed to reject the null hypothesis, there is a potential for a Type II error, which occurs when an effect exists but is not detected. This could be because of insufficient statistical power, as for both our tests, we have a higher p-value than the standard value of 0.05 due to a smaller sample size. A larger sample size is necessary to have sufficient power to reject the null hypothesis. A power analysis could provide us with more details about the same.

Part D: Power Analysis:
A power analysis was conducted using G*Power, with a Pearson's correlation effect size of $r = -0.18$, a sample size of 72, and an alpha level of 0.05. The power analysis revealed a power of 0.33 [Figure 5], well below the desired threshold of 0.8. This low power, likely due to the small sample size, suggests insufficient sensitivity to detect significant effects. Increasing the sample size could improve the reliability of the results.

Part E: Discussion:
The results indicate that proximity does not significantly affect customer loyalty to Starbucks, as both the correlation and logistic regression analyses failed to reject the null hypothesis. The weak negative correlation between distance and self-reported loyalty (r = -0.18) suggests a slight, but insignificant, relationship. For both the correlation and logistic regression analyses, the p-values (0.13 and 0.11, respectively) are not far from the 0.1 significance level. This suggests that using a larger sample size could potentially reduce the p-value below 0.1, revealing a more significant relationship.

From the exploratory data analysis, we found that taste emerged as the dominant factor influencing customer loyalty [Figure 1], with 68% of participants citing it as the primary driver, compared to only 17% for proximity. This suggests that factors beyond proximity, such as product quality, play a more significant role in fostering loyalty to Starbucks. The lack of significance between proximity and loyalty can be attributed to the sample's limited variation in distance [Figure 2], as many respondents lived within a 0.5-mile radius of a Starbucks. This reduced variability made it difficult to detect an effect of proximity on loyalty. Another consideration is the relatively high number of participants in our study who reported low loyalty in the self-reported measure. This skew toward non-loyalty may have further reduced the sensitivity of our analysis, making it harder to detect subtle effects of proximity on loyalty [Figure 3, Figure 4].

Given these limitations, future research should aim for a larger, more representative sample to more accurately assess the relationship between proximity and loyalty. Additionally, exploring other factors such as taste could provide a more comprehensive understanding of the drivers of customer retention at Starbucks.

Figures and Tables

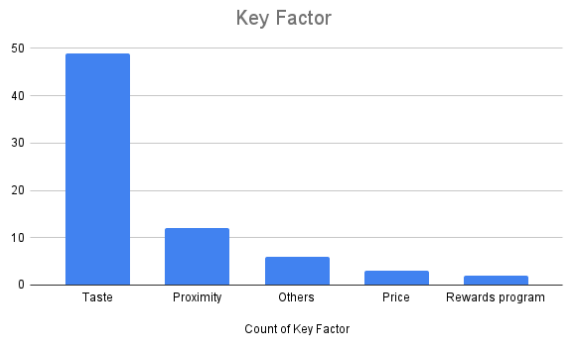

Figure 1: Distribution of Key Factor responses

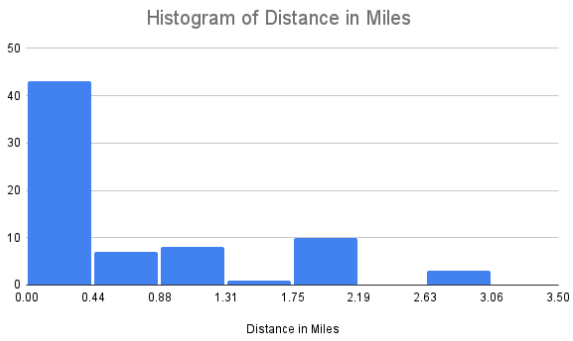

Figure 2: Distribution of Distances

Figure 3: Distribution of Self Reported Loyalty scores

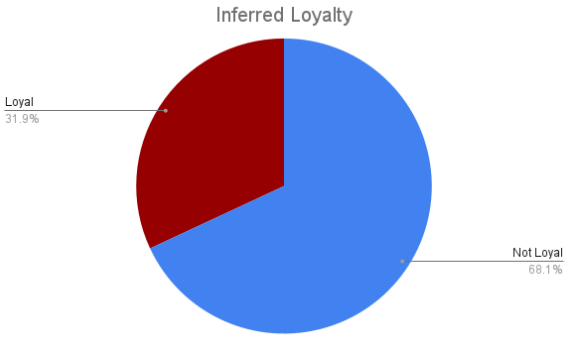

Figure 4: Pie chart of Inferred Loyalty

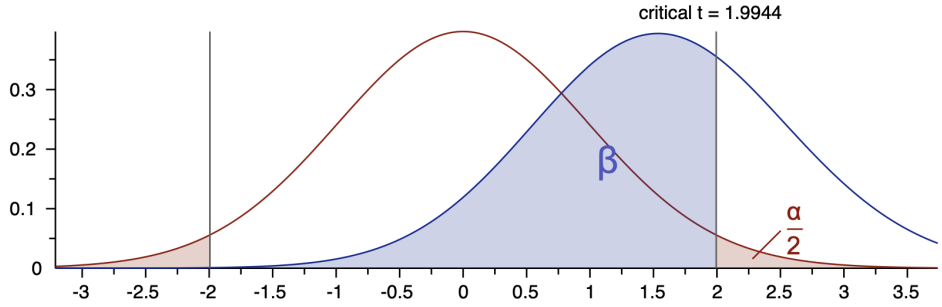

Figure 5: Power distribution plot