

Data analysis of customer retention of Multinational chain of Coffee House

Hussain Ghoniem

Cairo, Egypt

+201121332665

glowingdesert@gmail.com

Executive Summary

This data analysis aims to explore customer retention patterns for a multinational chain of coffeehouses. Key insights include factors influencing customer loyalty, trends in customer behavior, and recommendations for improving retention.

Introduction

Coffee House is a multinational chain of coffeehouses and roastery reserves headquartered in Seattle, Washington. It is the world's largest coffeehouse chain. We have a large customer base and have consistently experienced high retention rates in the past. Over the past year, Coffee House noticed a decline in the number of repeat customers and a corresponding decline in revenue. After reviewing customer feedback and conducting surveys, they have determined that the decline in retention rates is likely due to a combination of factors, including a lack of variety in our product offerings, poor customer

service, and increased competition from other retailers. They have decided to conduct a data analysis project to better understand the factors contributing to the decline in retention rates and identify potential solutions to improve retention and increase revenue.

Data Overview

The dataset comprises 122 rows and 20 questions with timestamps to verify that there will be no duplicates, capturing feedback from individuals regarding their experiences with the coffeehouse. The questions cover various aspects, including demographics, visits frequency, preferences, and satisfaction ratings.

Initial Observations:

- Size of the Dataset:** 122 rows
- Variables:** 20 questions capturing different aspects of customer feedback. Questions include information about gender, age, current status, annual income, visit frequency,

enjoyment preferences, spending habits, outlet preferences, membership card status, purchase preferences, and ration on various aspects of the coffee house.

Unique Values

Gender:

Female

Male

Age:

Below 20

From 20 to 29

From 30 to 39

Above 40

Current:

Student

Employed

Self-Employed

House-wife

Income:

Less than RM25,000

RM25,000 till RM50,000

RM50,000 till RM100,000

RM100,000 till RM150,000

More than RM150,000

Visits:

Rarely

Daily

Weekly

Monthly	Billboards	Less than RM20 were 49.6%
Never	Through friends and word of mouth	of the whole interviewed group
Enjoyment:		
Dine in	Application offer	From RM20 to RM40 were
Take away	Continue Buying	39.8% of the whole interviewed group
Drive-thru	Yes	While the more than RM40
Never	No	were 6.2% and the Zero were
Never Buy	Data Handling	4.4% of the whole interviewed group
I don't like coffee	1. Null Values: Identified and removed two null values that were one one row related to a feed backer who haven't visited the coffeehouse.	Distribution Analysis
Duration:	2. Filtered Feed Backers: Excluded feed backers who have never visited the coffeehouse, resulting in 8 filtered records.	Quality Ratings:
Below 30 minutes		Mean: 3.74 (Slightly above average)
Between 30 minutes to 1 hour		Median: 4.0 (Median indicates a relatively balanced distribution)
Between 1 hour to 2 hours		Standard Deviation: 0.89 (Moderate variability)
Between 2 hours and 3 hours		Interpretation: On average, customers rate the quality of the coffeehouse above the midpoint, with a fair amount of consistency.
More than 3 hours		Price Range Ratings:
Distance	Exploratory Data Analysis (EDA)	Mean: 2.93 (Below average)
Within 1km	Exploratory Data Analysis involves gaining insights into the dataset through various statistical and visual exploration techniques.	Median: 3.0 (Median indicates a relatively balanced distribution)
1km - 3km	Descriptive statistics	Standard Deviation: 1.08 (High variability)
More than 3km	Let's begin by examining key summary statistics for numerical features in the dataset	Interpretation: Customers, on average, rate the price range below the midpoint, and there is a significant variability in opinions.
Membership card	1. Distribution of age groups	Promotion Importance Ratings:
Yes	From 20 to 29 were 69.9% of the whole interviewed group.	Mean: 3.88 (Above average)
No	From 30 to 39 were 15.0% of the whole interviewed group.	Median: 4.0 (Median indicates a relatively balanced distribution)
Favorite Product	While the below 20 and 40 and above were the rest 8.8% for the below 20 and 6.2% were for the 40 and above.	Standard Deviation: 1.04 (Moderate variability)
Coffee		Interpretation: On average, customers find promotions
Cold Drinks		
Pastries		
Sandwiches		
Juices		
Jaws chip		
Cake		
Spending Limit		
Less than RM20		
Around RM20 - RM40		
More than RM40		
Zero		
Advertisers		
Starbucks Website/Apps		
Social Media		
Emails		
Deal Sites		
In Store displays		

important, with a fair amount of variability in opinions.

Ambiance Ratings:

Mean: 3.84 (Above average)

Median: 4.0 (Median indicates a relatively balanced distribution)

Standard Deviation: 0.86 (Low to moderate variability)

Interpretation: Customers, on average, rate the ambiance positively with relatively consistent opinions.

WiFi Quality Ratings:

Mean: 3.28 (Slightly above average)

Median: 3.0 (Median indicates a relatively balanced distribution)

Standard Deviation: 0.97

(Moderate variability)

Interpretation: On average, customers rate WiFi quality slightly above the midpoint, with moderate variability in opinions.

Service Quality Ratings:

Mean: 3.81 (Above average)

Median: 4.0 (Median indicates a relatively balanced distribution)

Standard Deviation: 0.80 (Low to moderate variability)

Interpretation: Customers, on average, rate the service quality positively with relatively consistent opinions.

Business Meetings Preference:

Mean: 3.54 (Slightly above average)

Median: 4.0 (Median indicates a relatively balanced distribution)

Standard Deviation: 1.03

(Moderate variability)

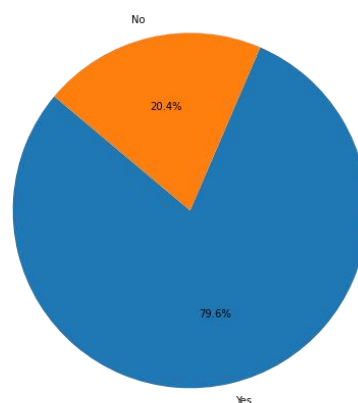
Interpretation: On average, customers express a preference for the coffeehouse for business meetings, with moderate variability in opinions.

Visualizations

Customer Retention Analysis

Upon evaluating the dataset comprising feedback from 113 customers, a notable observation emerges. Approximately 20.4% of respondents explicitly conveyed their decision to discontinue patronage, expressing a choice not to continue purchasing products from the coffeehouse.

Figure 1 Percentage of the Continue Buying



Comparative Analysis of Ratings

I created a data frame for the customers that said no in continue buying and made a box-plot for the rating in both the data frames the whole interviewed group and the customers that said no in continue buying

Figure 2 The ranges in the rating

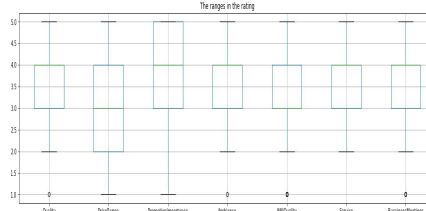
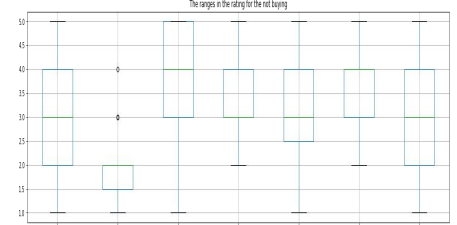


Figure 3 The ranges in the rating for the not buying



To correlate the features, the data should be numeric, so I used the mapping method to encode the unique values of each column into numeric digits, excluding the timestamp column since it's just an index.

Figure 4 The correlation of the column

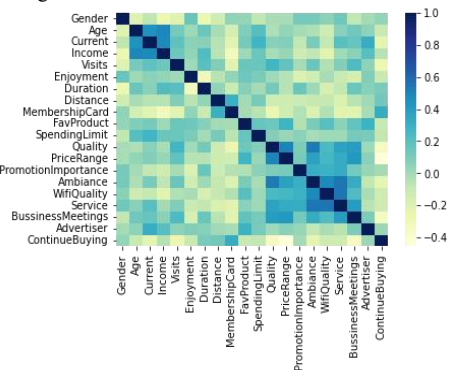
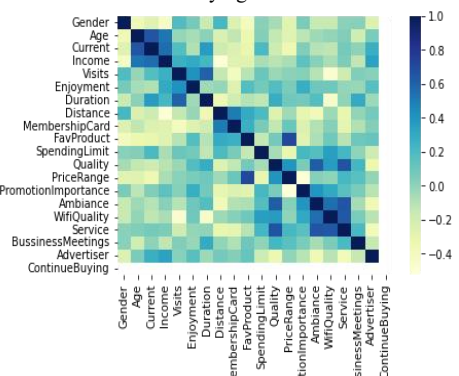


Figure 5 The correlation of the columns for the not buying



Utilizing the correlation analysis, key features have been identified as pivotal in determining customer retention or discontinuation of purchasing behavior. The following features exhibit significant influence:

Membership Card:

The presence or absence of a membership card has demonstrated a notable impact on customer decisions regarding continued engagement.

Distance:

Proximity to the Coffee House outlet plays a crucial role, indicating that customers may be influenced by the convenience of the location.

Duration:

The amount of time customers spend during their visits emerges as a contributing factor, suggesting that a positive or negative experience duration may affect their likelihood to continue purchasing.

Promotion Importance:

The perceived importance of promotions in the decision making process significantly influences customer behavior, indicating the effectiveness of promotional strategies.

Advertiser:

The source or channel through which customers become aware of promotions, such as online platforms, social media, or word of mouth, contributes substantially to their decision-making process.

Gender:

Gender-based preferences have shown correlation, underscoring that certain aspects of the customer experience may

resonate differently with male and female customers.

Income:

Annual income levels have demonstrated a discernible impact on customer retention, highlighting the role of financial considerations in purchasing decisions.

Enjoyment:

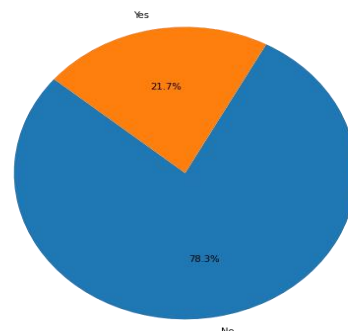
Customer enjoyment preferences, encompassing choices such as dine-in, take-away, or drive-thru, are indicative of factors that significantly shape the overall customer experience.

These identified features serve as strategic focal points for businesses seeking to enhance customer retention and tailor their approaches to meet the diverse needs and preferences of their clientele. Further exploration and targeted strategies in these areas may yield valuable insights for optimizing customer satisfaction and loyalty.

So I analyzed the data of these features.

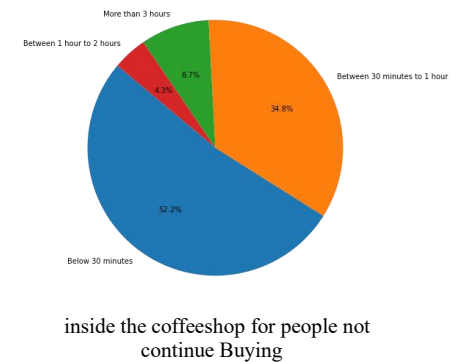
Out of the customers not continue buying 21.7% only has a membership card.

Figure 6 Percentage of the people not continue Buying with Membership Cards
Percentage of the people not continue Buying with Membership Cards



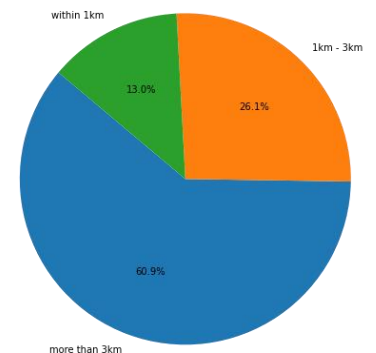
Out of the customers not continue buying 87% of them spend below than one hour in the coffeshop.

Figure 7 Percentage of the duration spent
Percentage of the duration spent inside the coffeshop for people not continue Buying



Out of the customers not continue buying 60.9% needs to travel more than 3km to the coffeshop out of the interviewed customers that needs to travel more than 3km to the coffeshop 25% said they are not continue buying.

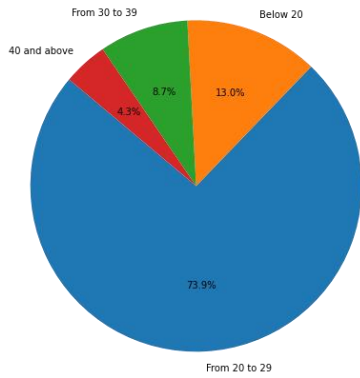
Figure 8 Percentage of the distance away from the coffeshop for people not continue Buying
Percentage of the distance away from the coffeshop for people not continue Buying



Out of the customers not continue buying 86.9% of them were from two age groups the below 20 and the from 20 to 29, which means that young people are not favoring the multinational chain of coffeshops.

Figure 9 Percentage of the age groups for people not continue Buying

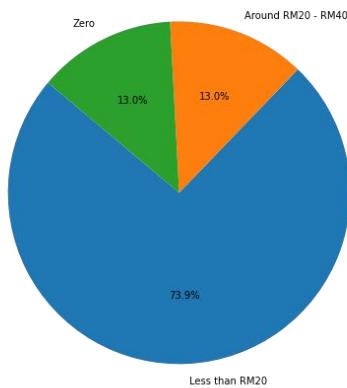
Percentage of the age groups for people not continue Buying



Out of the customer not continue buying 86.9% are spending less than RM20 per visit.

Figure 10 Percentage of the spending limit groups for all the not continuing

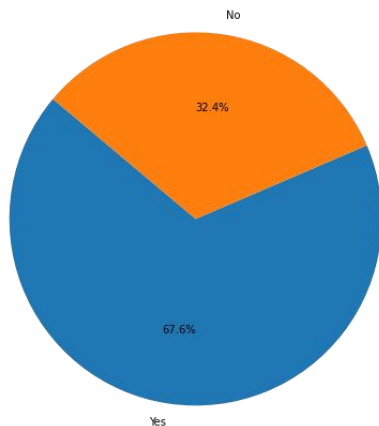
Percentage of the spending limit groups for all the not continuing



The students made up 32.7% of the whole interviewed group but they represented 52.2% of the customers not continue buying, 32.4% of the students interviewed said no in continue buying.

Figure 11 Percentage of the Students group in the continue buying

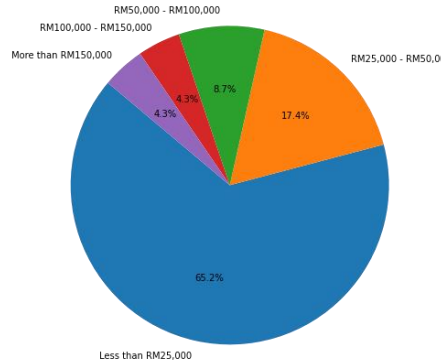
Percentage of the Students group in the continue buying



Out of the customers not buying 82.6% of there were from two income groups the less than RM25,000 and the RM25,000 to RM50,000. 21.6% of those two income groups said no in continue buying.

Figure 12 Percentage of the income group for people not continue Buying

Percentage of the income group for people not continue Buying

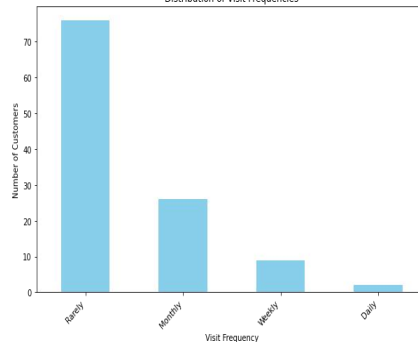


Out of the customers not continue buying 72.0% favored coffee, 20% of them favored cold drinks, and the last 8% of them favored pastries. Which were the top 3 products favored by the whole interviewed group.

Most of the interviewed do rarely visit the coffeehouses of this multinational chain coffeehouse.

Figure 13 Distribution of Visit Frequencies

Distribution of Visit Frequencies



Recommendations

Membership Engagement:

Encourage membership card usage with targeted promotions for cardholders.

Location Optimization:

Consider opening outlets in areas with higher customer concentration.

Evaluate the feasibility of delivery services for customers at a distance.

Experience

Duration

Enhancement:

Introduce initiatives to enhance the overall customer experience within a shorter duration.

Targeted Promotions:

Tailor promotions based on customer segments to maximize impact.

Ad Campaign Optimization:

Focus on the most effective advertising channels for promotions.

Demographic

Centric

Approaches:

Tailor strategies based on age groups and income levels.

Product Diversification:

Consider expanding product offerings or highlighting lesser-known favorites.

Student Engagement:

Develop student centric promotions and experiences.

Limitations

Sample Representation:

The dataset may not fully represent all customer segments.

Temporal Aspects:

External factors influencing customer behavior may not be fully captured.

Survey Bias:

Responses might be influenced by survey wording or framing.

Conclusion

This analysis highlights crucial insights into customer retention for Coffee House. Key factors influencing retention include membership engagement, location optimization, and targeted promotional strategies. Understanding demographic nuances, optimizing the customer experience duration, and diversifying product offerings are vital for sustained success. Acknowledging the limitations, this analysis provides a foundation for strategic decision-making to enhance customer satisfaction and retention, ultimately driving revenue growth. Future research could explore evolving customer preferences and external market dynamics for a comprehensive strategy.

