



## **CPE 329 Business Intelligence**

### **Technical report**

#### **Submitted to**

Asst. Prof. Dr. Unchalisa Taetragool

#### **By**

#### **Sausage Dog**

- |              |                    |                 |
|--------------|--------------------|-----------------|
| 1. Jiajia    | Bai                | ID: 63070503404 |
| 2. Tamonwan  | Boonpa             | ID: 63070503418 |
| 3. Apirak    | Senarak            | ID: 63070503458 |
| 4. Kunnithi  | Treecharoensomboon | ID: 63070503461 |
| 5. Nunthanut | Kittirattanaviwat  | ID: 63070503471 |

Faculty of Engineering  
Department of Computer Engineering  
King Mongkut's University of Technology Thonburi

## Introduction

The Food and Beverage industry is an ever-evolving sector that serves as an integral part of the global economy. This industry encompasses a wide range of businesses, from fine-dining restaurants to food trucks and grocery stores, and it employs millions of people worldwide. However, with fierce competition in the industry, achieving success is becoming increasingly difficult.

Food and Beverage businesses face a range of challenges, from managing costs to providing excellent customer service to staying ahead of the competition. These challenges are magnified by the changing needs and preferences of customers who are more discerning than ever before. Today's customers are looking for more than just good food and drinks; they want an exceptional experience that they can't find anywhere else.

As a result, customer satisfaction and loyalty have become essential factors in determining the success of any food and beverage business. Achieving customer satisfaction and loyalty requires a deep understanding of customers' needs and preferences. A satisfied customer is more likely to return and bring more business, while a dissatisfied customer will spread negative word of mouth and decrease profitability. [1]

Therefore, building guests' satisfaction and loyalty is not only important but also has become obligatory for Food and Beverage businesses for their survival and growth in such a fierce competitive environment. As a result, the purpose of this study is to support the Food and Beverage industry by analyzing customer data and feedback to find out the various antecedents of customer loyalty and help them implement the necessary strategies.

## Business Question

1. **Predictive Analytics question:** What factors are most likely to influence customer loyalty?
2. **Descriptive Analytics question:** What is the level of customer satisfaction with the pricing of our products based on their characteristics?
3. **Another Analytics question:** What are the differences in spending patterns of each visiting frequency?

## Dataset Details

The dataset utilized in this project is provided by Mahira Hamzah and available on Kaggle, licensed under Data files © Original Authors. Two files are included in the dataset, which provide information on two different subjects related to Starbucks' customer satisfaction in Malaysia. These files contain information on various aspects related to the business, including customer demographics such as age and gender, their satisfaction levels with different aspects of the business such as beverages or facilities, and their buying behavior, including spending patterns and visit frequency.

File 1: “Starbucks satisfactory survey encode cleaned.csv” - This file contains data that has been encoded and cleaned for analysis. It comprises 122 customer records with 33 columns of information. Here is a detailed description of the dataset attributes:

Attribute Name	Meaning
Id	A unique identifier for each customer
Gender	The gender of the customer (0 - Male, 1 - Female)
Age	The age of the customer (0 - Below 20, 1 - From 20 to 29, 2 - From 30 to 39, 3 - 40 and above)
Status	The occupation status of the customer (0 - Student, 1 - Self-Employed, 2 - Employed, 3 - Housewife)
Income	The annual income of the customer (0 - Less than RM25,000, 1 - RM25,000 – RM50,000, 2 - RM50,000 – RM100,000, 3 - RM100,000 – RM150,000, 4 - More than RM150,000)
VisitNo	How often the customer visits Starbucks (0 - Daily 1 - Weekly 3 - Monthly 4 - Never)
Method	The method the customer uses to enjoy Starbucks, such as dining in, drive-thru, or takeaway (0 - Dine In, 1 - Drive-thru, 2 - Take away, 3 - Never, 4 - Others).
TimeSpend	How much time the customer spends at Starbucks (0 - Below 30 mins, 1 - 30 mins to 1h, 2 - 1h to 2h, 3 - 2h to 3h, 4 - More than 3h)
Location	The distance of the Starbucks store from the customer's location (0 - Within 1km, 1- 1km to 3km, 2 - More than 3km)
MembershipCard	Whether the customer is a member of the Starbucks rewards program (0 - Yes, 1 - No)
ItemPurchaseCoffee	Whether or not the customer frequently purchases coffee at Starbucks (0 - Yes, 1 - No)
itempurchaseCold	Whether or not the customer frequently purchase cold drinks at

	Starbucks (0 - Yes, 1 - No)
itemPurchasePastries	Whether or not the customer frequently purchase pastries at Starbucks (0 - Yes, 1 - No)
itemPurchaseJuices	Whether or not the customer frequently purchase juices at Starbucks (0 - Yes, 1 - No)
itemPurchaseSandwiches	Whether or not the customer frequently purchase sandwiches at Starbucks (0 - Yes, 1 - No)
itemPurchaseOthers	Whether or not the customer frequently purchase other product at Starbucks (0 - Yes, 1 - No)
SpendPurchase	Amount of money the customer spends on Starbucks purchases. ( 0 - Zero, 1 - Less than RM20, 2 - RM 20 to RM40, 3 - More than RM40 )
ProductRate	The customer's rating of the quality of Starbucks' products, ranging from 1 (Very Bad) to 5 (Excellent).
PriceRate	The customer's rating of the value for money of Starbucks' products, ranging from 1 (Very Bad) to 5 (Excellent).
PromoRate	The customer's rating of Starbucks' promotional activities, ranging from 1 (Very Bad) to 5 (Excellent).
AmbianceRate	The customer's rating of the ambiance of Starbucks stores, ranging from 1 (Very Bad) to 5 (Excellent).
WifiRate	The customer's rating of the quality of Starbucks' Wi-Fi, ranging from 1 (Very Bad) to 5 (Excellent).
ServiceRate	The customer's rating of the quality of service at Starbucks, ranging from 1 (Very Bad) to 5 (Excellent).
ChooseRate	The customer's rating of the variety of products available at Starbucks, ranging from 1 (Very Bad) to 5 (Excellent).
promoMethodApp	This column shows whether customers use the mobile app to stay up to date on Starbucks promotions and offers. (0 - Yes, 1 - No)
promoMethodSoc	This column shows whether customers use social media channels (e.g., Facebook, Instagram, Twitter, etc.) to stay up to date on Starbucks promotions and offers. (0 - Yes, 1 - No)
promoMethodEmail	This column shows whether customers use email to stay up to date on Starbucks promotions and offers. (0 - Yes, 1 - No)
promoMethodFriend	Whether customers utilize deals or discounts for promotional purposes at Starbucks.(0 - Yes, 1 - No)

promoMethodDisplay	This attribute indicates whether the customer was exposed to the Starbucks promotion through a display or signage.(0 - Yes, 1 - No)
promoMethodBillboard	This attribute indicates whether the customer was exposed to the Starbucks promotion through billboard. (0 - Yes, 1 - No)
promoMethodOthers	This attribute indicates whether the customer was exposed to the Starbucks promotion through other means. (0 - Yes, 1 - No)
Loyal	Whether the customer considers themselves loyal to Starbucks or not. (0 - Yes, 1 - No)

File 2: “Starbucks satisfactory survey.csv” - This file contains the raw data collected from the customer satisfaction survey conducted by Starbucks. The dataset consists of responses from 122 customers and includes 21 columns

Attribute Name	Meaning
Timestamp	The date and time when the survey response was recorded.
Your Gender	The gender of the customer with two possible values (male or female).
Your Age	The age of the customer (From 20 to 29, From 30 to 39, Other)
Are you currently...?	The occupation of the customer with four possible values (Student, Self-Employed, Employed, Housewife).
What is your annual income?	The annual income of the customer in Malaysian Ringgits.
How often do you visit Starbucks?	How often the customer visits Starbucks with six possible values (Never, Rarely, Occasionally, Frequently, Very frequently, Daily).
How do you usually enjoy Starbucks?	The method the customer uses to enjoy Starbucks (Dine-in, Drive-thru, Takeaway)
How much time do you normally spend during your visit?	How much time the customer spends at Starbucks (in minutes)
The nearest Starbucks's outlet to you is...?	The distance of the Starbucks store from the customer's location (Distance from customer's location)
Do you have Starbucks membership card?	Whether the customer is a member of the Starbucks. (Yes or No)

What do you most frequently purchase at Starbucks?	A categorical variable indicates the most frequently purchased item at Starbucks. (Coffee, Cold drinks, Other)
On average, how much would you spend at Starbucks per visit?	How much the customer typically spends per visit to Starbucks in Malaysian Ringgits. (Less than RM20, Around RM20 - RM40, Zero, More than RM40)
How would you rate the quality of Starbucks compared to other brands (Coffee Bean, Old Town White Coffee..) to be	The customer's rating of the quality of Starbucks' products compared to other brands, ranging from 1 (Very Bad) to 5 (Excellent).
How would you rate the price range at Starbucks?	The customer's rating of the quality of Starbucks' products, ranging from 1 (Very Bad) to 5 (Excellent).
How important are sales and promotions in your purchase decision?	The customer's rating of Starbucks' promotional activities, ranging from 1 (Very Bad) to 5 (Excellent).
How would you rate the ambiance at Starbucks? (lighting, music, etc...)	The customer's rating of the ambiance of Starbucks stores, ranging from 1 (Very Bad) to 5 (Excellent).
You rate the WiFi quality at Starbucks as...	The customer's rating of the quality of Starbucks' Wi-Fi, ranging from 1 (Very Bad) to 5 (Excellent).
How would you rate the service at Starbucks? (Promptness, friendliness, etc..)	The customer's rating of the service at Starbucks, ranging from 1 (Very Bad) to 5 (Excellent).
How likely you will choose Starbucks for doing business meetings or hangout with friends?	The likelihood of the customer choosing Starbucks for business meetings or hangouts with friends, ranging from 1 (Very Bad) to 5 (Excellent).
How do you come to hear of promotions at Starbucks? Check all that apply.	A categorical variable indicates how the customer learns about promotions at Starbucks, ranging from 1 (Very Bad) to 5 (Excellent).
Will you continue buying at Starbucks?	A binary variable indicates whether the customer will continue to purchase from Starbucks in the future, ranging from 1 (Very Bad) to 5 (Excellent).

## **Organizational Memory**

CRM software also provides real-time visibility into inventory levels, sales forecasting, and price management, which can be critical for managing fluctuating demand and inventory constraints. By providing online access to customer data, salespeople can access customer history and contact information from anywhere at any time, empowering them to handle difficult questions with confidence and resolve issues instantly.

The food and beverage industry, like any other industry, needs to effectively manage its organizational memory to meet customer demands while operating profitably. This is especially important given the various constraints such as fluctuating consumer demand, narrow margins, strict government regulations, and limited resources that are present in the industry. One tool that can help maintain a strong organizational memory is Customer Relationship Management (CRM). A CRM system brings together data from various sources such as restaurant website, social media, emails, and other web-based content for valuable feedback, all collected into one database, providing better insights. [2]

A CRM software can provide a holistic view of customer data and history, which can be used to improve relationships, increase customer retention, and ultimately drive sales growth. By integrating marketing and operations data with CRM system, businesses can maximize profits by undertaking highly targeted marketing campaigns. Information such as visit frequency and average spending, helps create offers and promotions tailored to their customer. Businesses can create promotional emails for customers who haven't visited in a while, and offer personalized discounts and incentives for regulars and heavy spenders. Online bookings and promotions can also be automatically added to the database to grow the customer list. [3]

Additionally, CRM software provides real-time visibility into inventory levels, sales forecasting, and price management, which can be critical for managing fluctuating demand and inventory constraints. By providing online access to customer data, salespeople can access customer history and contact information from anywhere at any time, empowering them to handle difficult questions with confidence and resolve issues instantly. [3]

In addition to CRM, ERP systems is another valuable tool for the food and beverage industry to manage their organizational memory. By utilizing ERP software, food and beverage companies can integrate data from different departments, including finance, human resources, manufacturing, supply chain, and inventory. This integrated data can be stored in a single database, providing real-time transaction data and eliminating the need for multiple legacy systems with maintenance complexities. Furthermore, ERP systems provide a framework for improving the way food and beverage companies manage their order fulfillment process, which involves various departments and stages, from taking customer orders to delivering products and collecting revenue. This process generates vast amounts of data, which can be stored in a data warehouse and used for trend analysis, helping companies to better understand customer behavior and market trends. [4]

In addition, ERP systems can be integrated with customer relationship management (CRM) software, which enables food and beverage companies to manage interactions and relationships with current and potential customers. By incorporating CRM and ERP software, organizations can improve their business processes and store information in the organizational memory. This information can be leveraged to make better decisions, improve productivity, and provide better customer service, ultimately contributing to the long-term success of the organization. [4]

Furthermore, CRM and ERP software can help organizations in the food and beverage industry to effectively manage the order fulfillment process, from the moment a customer places an order, to processing it, delivering the product, and collecting payment. ERP systems can integrate the different departments involved in this process, from sales, to finance, to manufacturing, to the logistics department that delivers the product. This can help organizations to streamline their processes, reduce costs, and improve customer satisfaction.[4]

In conclusion, incorporating CRM and ERP software can help organizations in the food and beverage industry to improve their business processes, store information in the organizational memory, and effectively manage the order fulfillment process. By doing so, organizations can make better decisions, improve productivity, and provide better customer service, ultimately contributing to the long-term success of the organization.



## **Information Integration**

The food and beverage industry has experienced a significant drop in sales due to the COVID-19 pandemic. Customer satisfaction is a critical factor for the industry's sustainable development, and understanding customer satisfaction is essential. Social media text reviews have become an important source of information for customers making purchasing decisions, particularly for food and beverage industries.

Text mining is a powerful tool for processing large volumes of unstructured or semi-structured data, such as online reviews. By converting text into numerical data, it allows businesses to identify important information and uncover hidden knowledge. As the number of online reviews continues to increase rapidly, text mining has become crucial for analyzing these data and extracting insights. It involves applying natural language processing (NLP) techniques and machine learning algorithms to automatically identify patterns, trends, and relationships within text data.[5]

In Text mining, technique can be classified into four main layers:

Information Retrieval (IR): In the food and beverage industry, IR techniques can be used to index and search through large volumes of online reviews to find relevant information about specific products, brands, or restaurants. For example, a company may use IR to search through customer reviews to identify common complaints or issues with a particular product, allowing them to address these issues and improve customer satisfaction.[5]

Information Extraction (IE): IE techniques can be used to extract specific information from unstructured text data, such as online reviews. In the food and beverage industry, this can include extracting information about the quality of ingredients, the level of customer service, or the overall dining experience. For example, a restaurant may use IE to extract information about the most commonly mentioned dishes in customer reviews, allowing them to make data-driven decisions about menu offerings.[5]

Text Categorization (TC): TC techniques can be used to categorize large volumes of text data into specific categories or topics. In the food and beverage industry, this can include categorizing customer reviews based on the type of cuisine, the price range, or the location of the restaurant. For example, a company may use TC to categorize customer reviews of their restaurant chain based on geographic location, allowing them to identify regional trends and tailor their offerings accordingly.[5]

Information Summarization (IS): IS techniques can be used to summarize large volumes of text data into a more concise and easily digestible format. In the food and beverage industry, this can include summarizing customer reviews into a few key points or themes. For example, a restaurant may use IS to summarize customer reviews into a list of the most common compliments and complaints, allowing them to quickly identify areas for improvement.[5]

## Text Mining Process

This section introduces the presented text mining-based method, to discover the most essential factors for customers' satisfaction from unstructured online customer reviews. The process contained 6 steps according to Sullivan (2001):

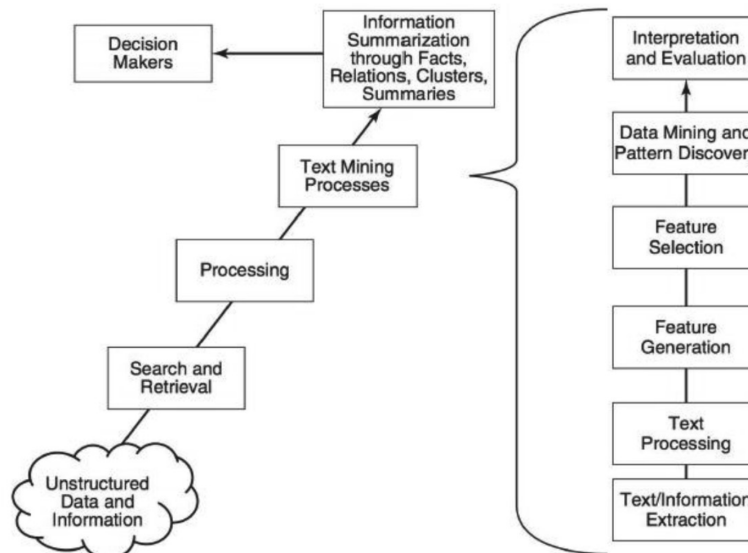


Figure1: text mining process

1. Text Extraction: Collect a large dataset of online customer reviews from relevant platforms. Extract the relevant text data, including reviews, ratings, and other metadata.
2. Text Processing: Preprocess the raw text data by removing irrelevant information such as stop words, punctuations, and non-alphabetic characters, performing stemming, and normalizing the text to a consistent format.
3. Feature Generation: Identify and extract important features from the preprocessed text data, such as specific product or service features mentioned in the reviews, sentiment analysis of the reviews, and other relevant metadata like location, date, and reviewer demographics.
4. Feature Selection: Select the most relevant features based on their frequency of occurrence and importance for customer satisfaction using statistical analysis and machine learning algorithms.
5. Data Mining and Pattern Discovery: Apply various various data mining and machine learning techniques, such as clustering, association rule mining, and classification, to uncover patterns and relationships among the selected features.
6. Interpretation and Evaluation: Interpret the results obtained from the previous steps and evaluate their relevance and accuracy for discovering the most essential factors for customers' satisfaction.

## Insight Creation

### The Machine Learning Canvas

<b>Decisions</b> <p>How are predictions used to make decisions that provide the proposed value to the end-user?</p> <p>by providing insights into the key factors that influence customer loyalty. Based on the results of the model, businesses can prioritize their efforts and resources to improve the factors that have the greatest impact on customer loyalty.</p>	<b>ML task</b> <p>Input, output to predict, type of problem</p> <p>Question: What factors are most likely to influence customer loyalty?</p> <p>I/P: Customer Survey data</p> <p>O/P: The level of influence of various factors on customer loyalty, ranked in order</p> <p>Problem: Regression model</p>	<b>Value Propositions</b> <p>What are we trying to do for the end-user(s) of the predictive system? what objectives are we serving</p> <p>To help the end-user(s) make informed decisions that will improve customer loyalty and drive business performance. By identifying the key factors that influence customer loyalty, the system can help businesses to prioritize their efforts and resources to maximize their impact on customer satisfaction and retention.</p>	<b>Data Sources</b> <p>Which raw data source can we use (internal and external)?</p> <p>Internal:</p> <ul style="list-style-type: none"> <li>• CRM tool</li> <li>• Payments database</li> <li>• Customer complaints, inquiries, and feedback.</li> <li>• Survey data</li> </ul> <p>External</p> <ul style="list-style-type: none"> <li>• Social media data</li> <li>• Market research data</li> </ul>	<b>Collecting Data</b> <p>How do we get new data to learn from (inputs and outputs)?</p> <p>Use survey form tools like Google Forms that can be used for a variety of purposes and easy to use to get survey from customer or people in the coffee shop. We will get the survey data from database every hour</p>
<b>Marketing Predictions</b> <p>When do we make predictions on new inputs? How long do we have to features a new input and make a prediction?</p> <p>Utilize a model to predict the loyalty and impact of various factors every hour, enabling us to detect any changes and make necessary adjustments.</p>	<b>Offline Evaluation</b> <p>Methods and metrics to evaluate the System before deployment.</p> <ol style="list-style-type: none"> <li>1. Correlation matrix</li> <li>2. R-squared (R<sup>2</sup>) value</li> <li>3. Mean squared error</li> <li>4. Mean absolute error</li> <li>5. Cross-validation</li> </ol>	<p>Objectives: to identify the factors that are most likely to influence customer loyalty in order to improve customer satisfaction and retention.</p>	<b>Features</b> <p>Input representations extracted from raw data sources.</p> <ul style="list-style-type: none"> <li>• Demographic data: age, gender, income, location, etc.</li> <li>• Transactional data: purchase history, frequency of visits, purchase amount, etc.</li> <li>• Customer feedback data: ratings, comments, etc.</li> </ul>	<b>Building Models</b> <p>When do we create/update models with new training data? How long do we have to featureise training input and create a model?</p> <p>Update the model daily by using the data from last 14 days</p>
<b>Live Evaluation and Monitoring</b> <p>Methods and metrics to evaluate the system after deployment, and to quantify value creation</p> <ul style="list-style-type: none"> <li>• Customer loyalty</li> <li>• Monitor churn rate</li> <li>• Customer satisfaction surveys</li> <li>• Check the Evaluation measurement</li> <li>• A/B testing</li> </ul>				



Attribution 4.0 International (CC BY 4.0) You are free to: Share & Adapt. Under the following terms: Attribution – You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

Designed by: Jeremy Corman

[www.digitalmarketingcanvas.co](http://www.digitalmarketingcanvas.co)

Figure 2: The Machine Learning Canvas

**Value Propositions:** To help the end-user(s) make informed decisions that will improve customer loyalty and drive business performance. By identifying the key factors that influence customer loyalty, the system can help businesses to prioritize their efforts and resources to maximize their impact on customer satisfaction and retention.

**Objectives:** to identify the factors that are most likely to influence customer loyalty in order to improve customer satisfaction and retention.

#### Data Sources:

Internal	External
<ul style="list-style-type: none"> <li>• CRM tool</li> <li>• Payments database</li> <li>• Customer complaints, inquiries, and feedback.</li> <li>• Survey data</li> </ul>	<ul style="list-style-type: none"> <li>• Social media data</li> <li>• Market research data</li> </ul>

**Collecting Data:** Use survey form tools like Google Forms that can be used for a variety of purposes and easy to use to get surveys from customers or people in the coffee shop. We will get the survey data from the database every month.

**Features:** Input representations extracted from raw data sources.

- Demographic data: age, gender, income, location, etc.
- Transactional data: purchase history, frequency of visits, purchase amount, etc.
- Customer feedback data: ratings, comments, etc.

**Building Models:** Update the model monthly by using the data from the previous month.

**Decisions:** By providing insights into the key factors that influence customer loyalty. Based on the results of the model, businesses can prioritize their efforts and resources to improve the factors that have the greatest impact on customer loyalty.

**ML task:** Question: What factors are most likely to influence customer loyalty?

I/P: Customer Survey data

O/P: The level of influence of various factors on customer loyalty, ranked in order

Problem: Regression model

**Marketing Predictions:** Utilize a model to predict the loyalty and impact of various factors every month, enabling us to detect any changes and make necessary adjustments.

**Offline Evaluation:** Methods and metrics to evaluate the System before deployment.

1. Correlation matrix
2. R-squared ( $R^2$ ) value
3. Mean squared error
4. Mean absolute error
5. Cross-validation

**Live Evaluation and Monitoring:** Methods and metrics to evaluate the system after deployment, and to quantify value creation

- Customer loyalty
- Monitor churn rate
- Customer satisfaction surveys
- Check the Evaluation measurement
- A/B testing

## Presentation

**Dashboard 1:** What is the level of customer satisfaction with the pricing of our products based on their characteristics?

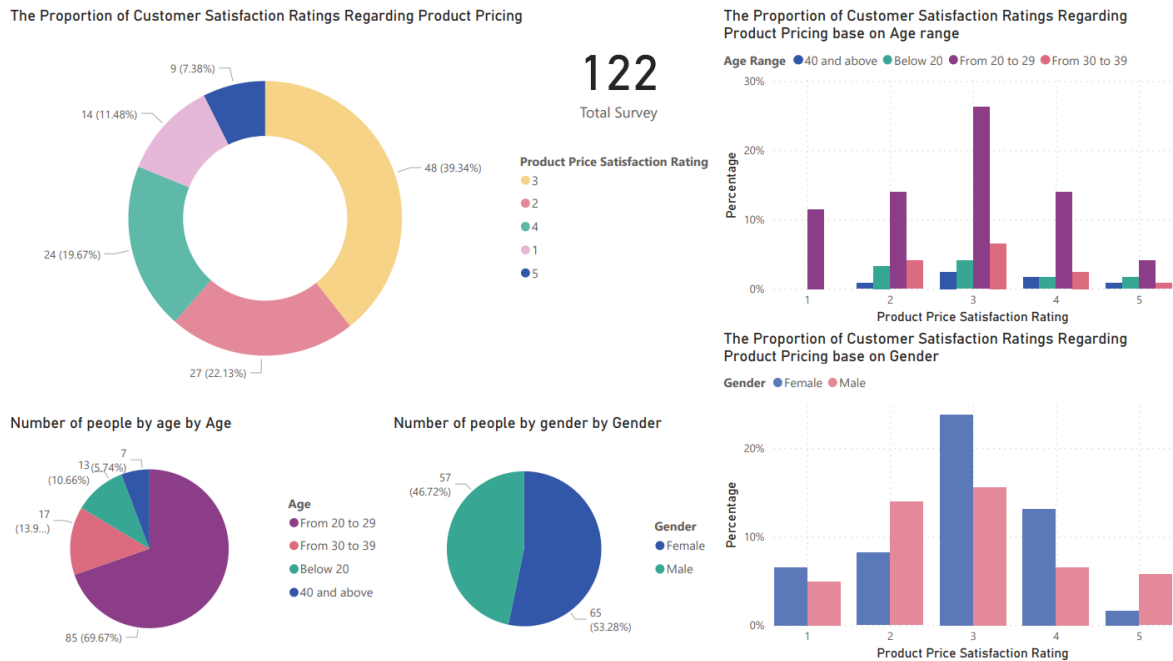


Figure 3: Dashboard 1

Figure 3 illustrates a dashboard that provides insight into customer satisfaction levels with our product pricing based on customer characteristics. The purpose of analyzing the satisfaction levels of different customer groups (e.g., teens under 20 and people over 40) is to develop targeted strategies to meet the specific needs and preferences of each group. By understanding the satisfaction levels of different customer groups, we can adjust our pricing strategies to ensure that each group stays engaged and continues to purchase our products.

THE PROPORTION OF CUSTOMER SATISFACTION RATINGS REGARDING PRODUCT PRICING

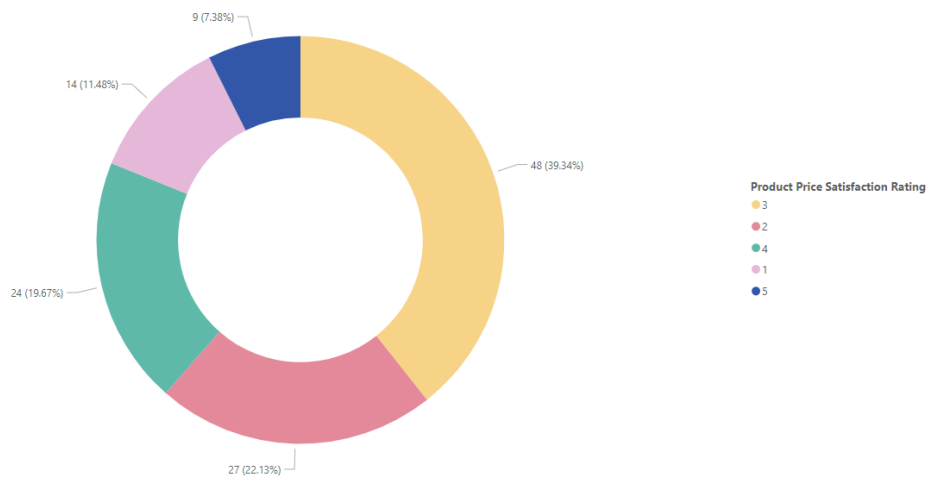


Figure 3.1: Proportion of Customer Satisfaction Ratings for Product Pricing in Donut Graph

The donut graph illustrates the proportion of customer satisfaction ratings for product pricing. The graph consists of five segments, each representing a product rating on a scale of 1 to 5, indicating the level of satisfaction with our product's price. The scale is arranged in order of satisfaction, with 1 representing the least satisfaction with the price of our products and 5 representing the highest satisfaction with our products. The graph clearly shows that most customers rated the product pricing favorably, with over half of the customers (72%) giving a rating of 3 or higher. Overall, the graph shows that the majority of customers are satisfied with the product pricing, although there is still room for improvement, particularly in the lowest rating category.

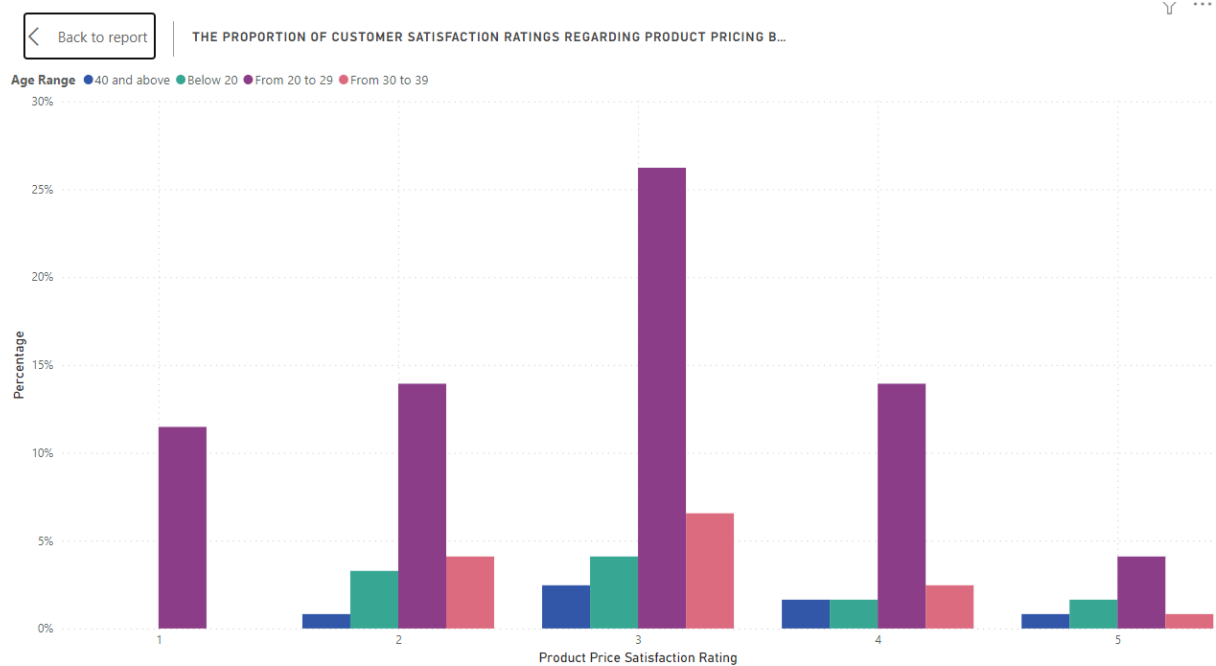


Figure 3.2: Customer Satisfaction Ratings for Product Pricing by Age Range in Bar Chart

Figure 3.2 shows the distribution of customer satisfaction by age group, which is a key customer characteristic in our input data. Each group bar in the figure represents a specific product rating, while the sections within the bar represent different age ranges. The age ranges include Below 20, From 20 to 29, From 30 to 39, and 40 and above. Interestingly, a consistent pattern emerged across the different product ratings. Most customers, regardless of age group, tend to give a rating of 3, indicating a moderate level of satisfaction. After this, the percentages of 2, 4, and 5 ratings are relatively similar across all age groups. However, it is worth noting that only a small percentage of customers in the 20-29 age group gave a rating of 1.

By analyzing the distribution of customer satisfaction ratings by age groups, we can gain valuable insights into how different age groups perceive the product pricing and identify areas of strength and potential areas for improvement.

NUMBER OF PEOPLE BY AGE BY AGE

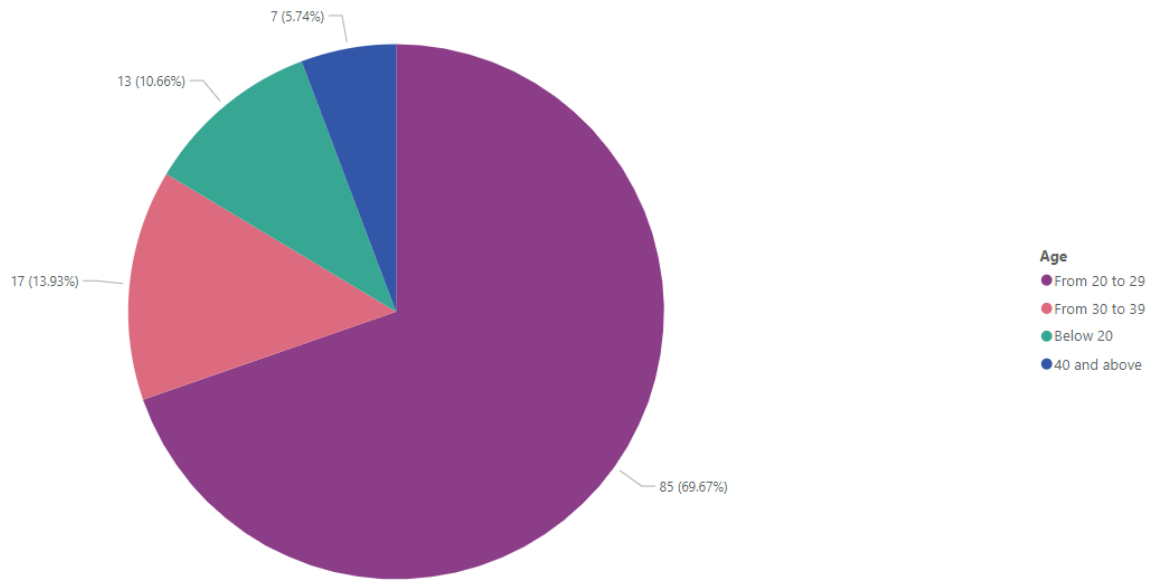


Figure 3.3: Distribution of People by Age Range in Pie Graph

The pie chart clearly illustrates the distribution of the survey participants in different age groups. It is clear that the largest segment, about 85% of the participants, belong to the 20-29 age group. In contrast, the 30-39 age group accounts for about 17% of the participants, while the under-20 age group accounts for about 10.66%. The smallest percentage, about 5.74%, represents individuals aged 40 and older.



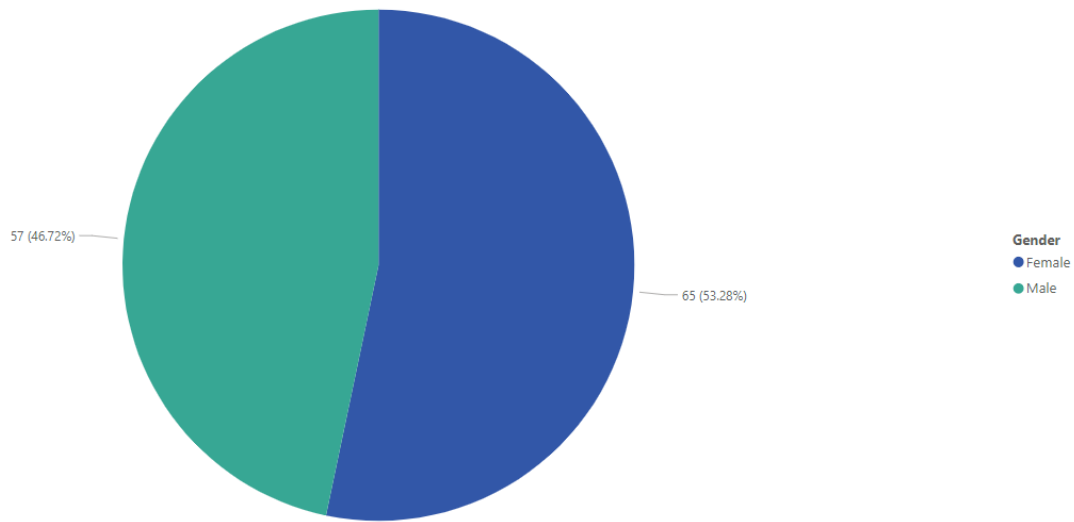


Figure 3.4: Distribution of People by Gender in Pie Graph

This pie graph displays the distribution of survey participants based on their gender, which consist of two segments; females and males. It is noteworthy that the female segment accounts for a slightly larger percentage, 53.28% of participants, while male account for 46.72%.

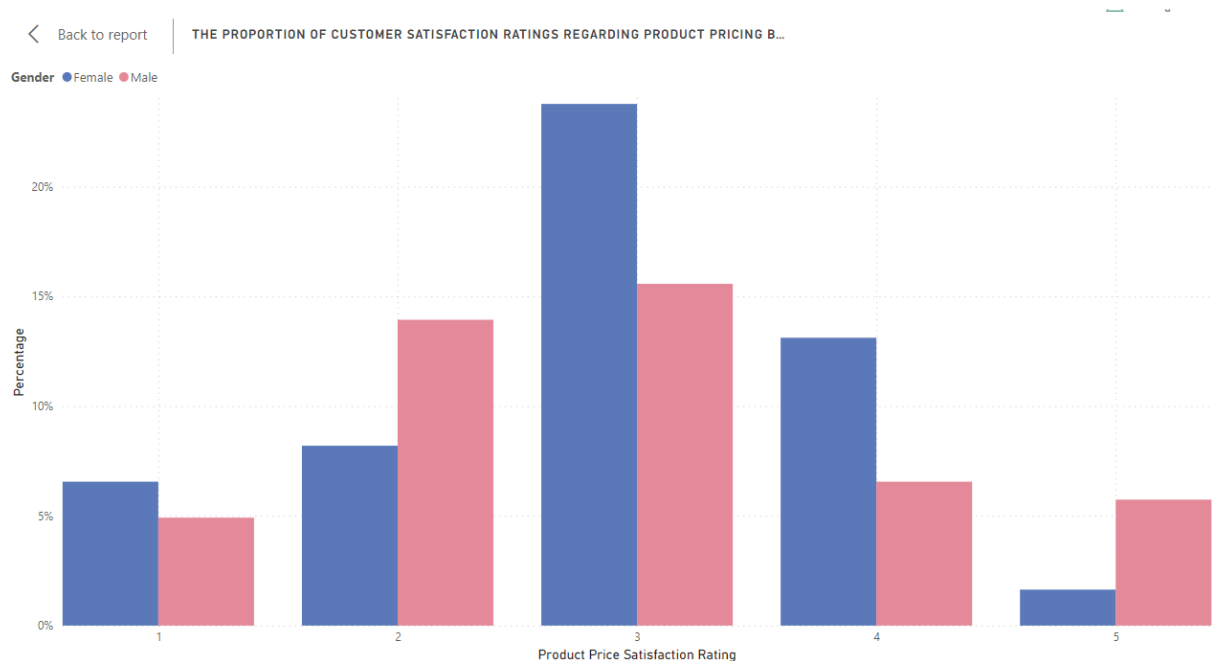


Figure 3.5: the proportion of customer satisfaction ratings regarding the product pricing by gender in bar chart

The bar chart is segmented into five group bars, each representing a different product price rating. Within each group bar, there are two sections representing male and female customers. The height of each bar represents the percentage of customers who gave a particular rating, and the color of each bar represents the gender of the customers who gave the rating.

From the chart, we can see that both male and female customers were most satisfied with the product pricing when rated 3. The proportion of customers who rated the product pricing 3 was higher than other ratings. Furthermore, upon closer examination of the chart, it can be noted that there is a difference between male and female customers in terms of ratings. Specifically, fewer women rated the price of the product as 5 compared to men. This difference suggests that males may perceive pricing as more favorable compared to females.

## Dashboard 2: What are the differences in spending patterns of each visiting frequency?



Figure 4: Dashboard 2

The second dashboard aims to address the third business question. This dashboard provides insights and analysis related to the average consumption patterns based on frequency of visits. By using these insights into these variations, businesses can create personalized offers and tailor their marketing strategies to effectively engage each segment. This includes implementing loyalty programs or incentives to encourage more frequent visits or higher spending from specific groups.

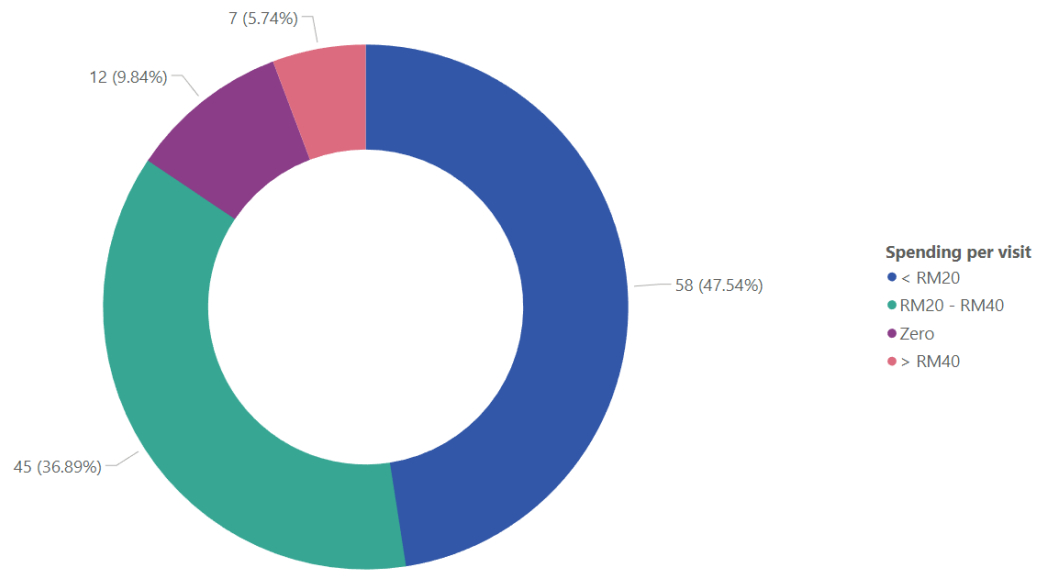


Figure 4.1: Distribution of Spending by Average Spend in Donut Chart

This chart is divided into 4 segments, each representing a group of people in the different spending amounts per visit class. The size of each segment is proportional to the number of people in that group. The chart includes labels or legends that indicate the name of each segment, along with the number of people and corresponding percentage of people in that group.

From the chart, the majority of the people are spending less than 20RM or 20-40RM, which could indicate that they are either on a tight budget or are relatively young and have not yet started earning a substantial income like students or people with lower-paying jobs.

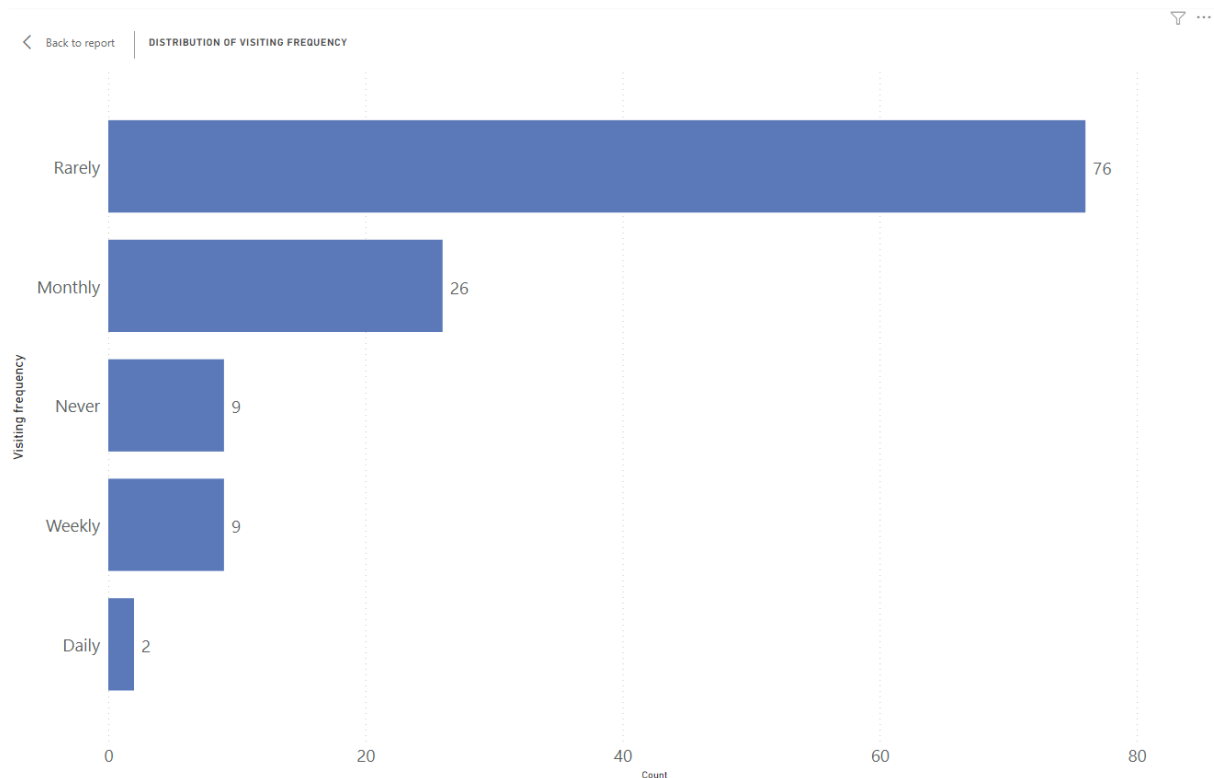


Figure 4.2: Distribution of visiting frequency in Bar chart

The bar chart has 5 horizontal bars, each representing a different visit frequency class: rarely, monthly, never, weekly, and daily. The length of each bar corresponds to the number of people in that class, with the y-axis showing the number of people and the x-axis showing the visit frequency class. The rarely bar is the longest at 76 people, followed by monthly at 26 people, and then weekly, never, and daily with 9 and 2 people respectively.

The insight from the graph is that most respondents visit rarely, with 76 people falling into this category. The next most common frequency is monthly, with 26 people visiting once a month. The percentage of those 2 categories is more than 75 percent and the sum of the remaining category is only 20 people. This suggests that the majority of people do not visit on a regular basis.

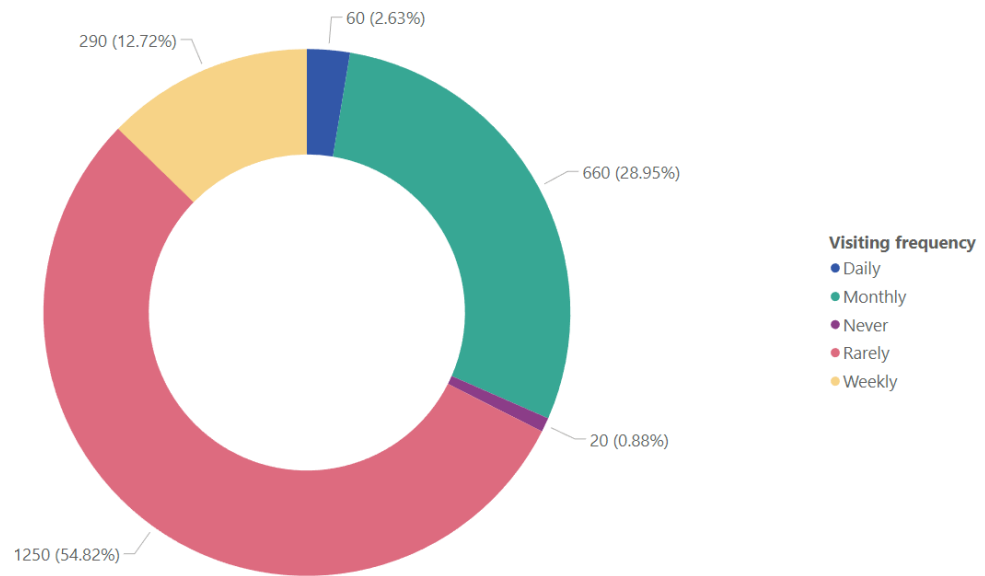


Figure 4.3: Total Sum of Average Spending per Visit Frequency in Donut Chart

The donut chart is divided into five sections, each representing a visit frequency class: rarely, daily, monthly, weekly, and never. The section for the rarely class has the largest portion of the chart, followed by monthly, weekly, daily, and never. Each section is color with the corresponding visit frequency class.

The insight from this chart is that while the most common frequency of visits is rarely, this group actually spends the most, with a total spending of 1250. The daily and monthly groups have a much lower total spending per visit, with only 60 and 660, respectively. From Figure 9 the weekly group has a higher average spending per visit than the rarely group, their total spending is much lower, with a total of only 290. The never group has the lowest total spending per visit, with only 20. Therefore, rarely group have the high potential to generate the most revenue per visit.

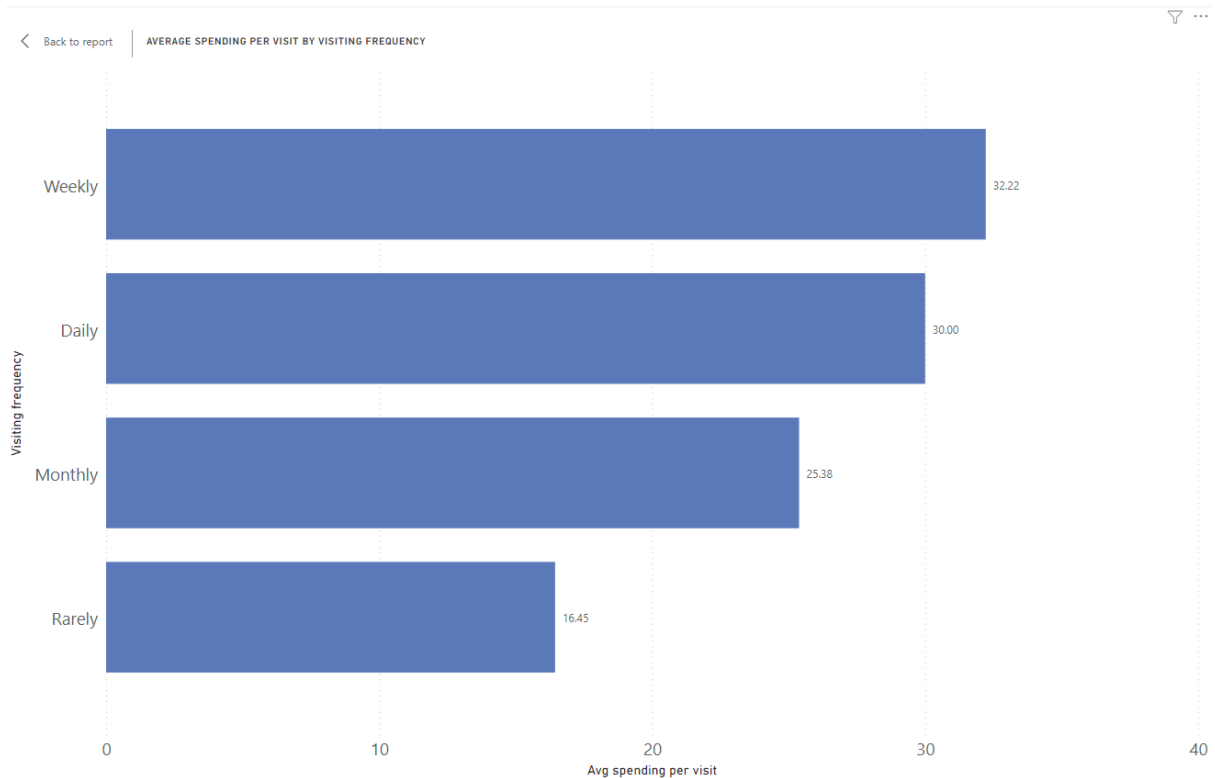


Figure 4.4: Average spending per visit by visiting frequency in Bar chart

The bar chart has four bars representing four different visit frequency classes: weekly, daily, monthly, and rarely. The length of each bar represents the average spending of people in that class. The bar for weekly is the tallest at 32.22, followed by daily at 30, monthly at 25.38, and rarely at 16.45.

The insight from this graph is that while the frequency of visits may not directly correlate with the amount spent per visit, the data suggests that those who visit more frequently tend to spend more on average.

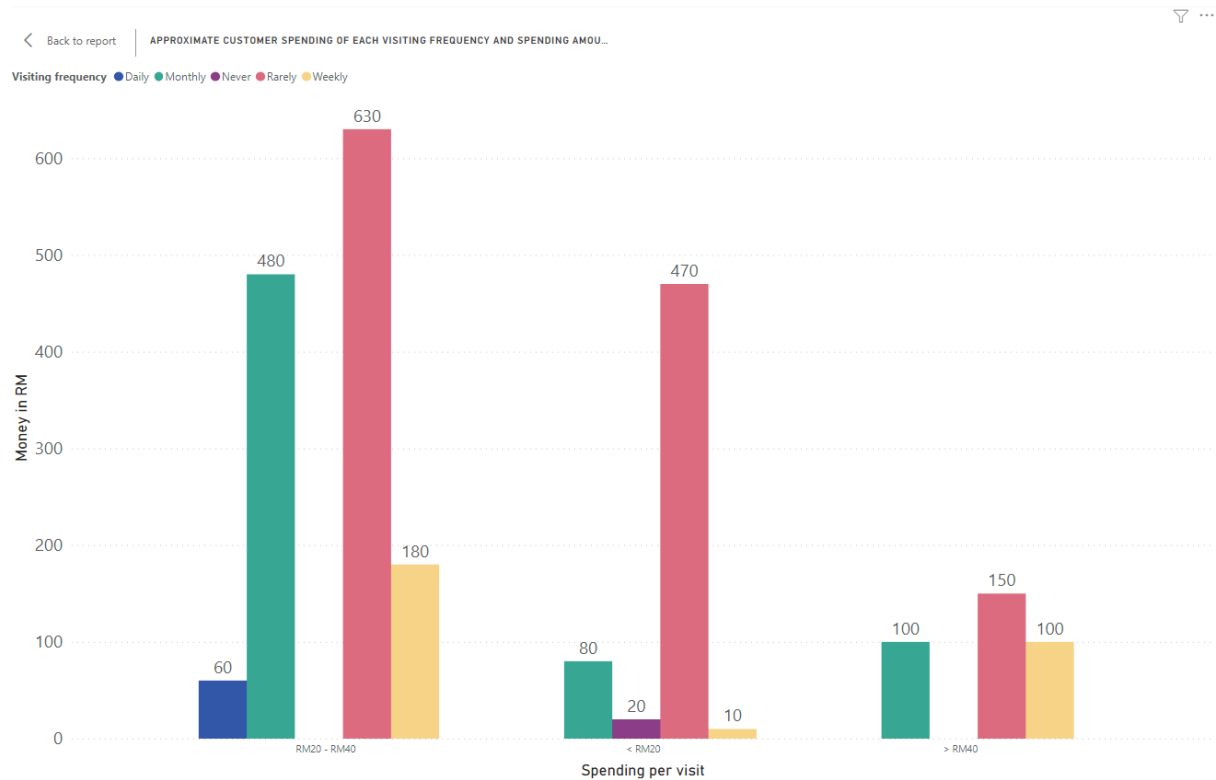


Figure 4.5: Approximate customer spending of each visiting frequency and spending amount in Column chart

The bar chart is divided into three clusters representing the three ranges of customer spending based on their frequency of visiting. The X-axis shows the three ranges of average spending, while the Y-axis represents the sum of spend. The legend shows the five frequencies of visiting.

From the chart, we can see that the highest-spending customers are those who visit rarely and spend 20-40RM per visit. However, there is also a significant number of customers who visit monthly and spend 20-40RM per visit and customers who visit rarely and spend less than 20RM per visit, indicating a potential opportunity for the business to target this group with marketing strategies. Additionally, the business should focus on providing services and products that appeal to customers who prioritize spending less than 20RM per visit, as they make up a significant portion of the customer base.

## Summary

Organizational Memory: To enhance business processes and store information effectively, we use Customer Relationship Management (CRM) and Enterprise Resource Planning (ERP) systems. These systems help us manage and utilize data such as demographic information, transactional records, and customer feedback, which are crucial for building a comprehensive organizational memory. (I/P: Demographic, transactional, and customer feedback data, O/P: Customer-structured data)

Information Integration: In order to handle large volumes of unstructured or semi-structured data like online reviews, we employ Text Mining as an integration tool. This allows us to extract valuable insights and integrate this information with other data sources, including customer surveys, for further analysis. (I/P: Customer Surveys, O/P: Integrated data for analysis)

Insight Creation: Logistic regression is employed to predict the likelihood of customer loyalty or disloyalty based on independent variables that impact loyalty. By utilizing integrated data, we can determine the level of influence various factors have on customer loyalty, ranked in order of significance. (I/P: Integrated data, O/P: Level of influence of various factors on customer loyalty, ranked in order)

Presentation: To effectively address business queries and gain valuable insights, we leverage visualization technology, specifically PowerBI. By visualizing integrated data, we are able to present information in a visually appealing and easily understandable manner, aiding in decision-making processes (I/P: Integrated data, O/P: Visualization)



## Reference

- [1.] Marketing Team. (2021, July 5). The importance of customer loyalty in the food industry. Doxee.  
<https://www.doxee.com/blog/marketing/importance-of-customer-loyalty-in-food-industry/>
- [2.] A, A. M., & A, A. M. (2023, February 16). Everything You Need to Know About CRM for Beverages Industry. Martech360.  
<https://martech360.com/crm/everything-you-need-to-know-about-crm-for-beverages-industry/>
- [3.] SapphireOne. (2022, August 28). Food and Beverage Industry is Catered to by SapphireOne CRM. SapphireOne: The ERP, CRM and Accounting Software Solution for Business. <https://www.sapphireone.com/crm-software/food-beverage-industry-crm/>
- [4.] *Understanding the Need for CRM & ERP Systems Integration* | Cleo. (n.d.). Cleo.  
<https://www.cleo.com/blog/knowledge-base-erp-and-crm-integration>
- [5.] Dang, S., & Ahmad, P. H. (2014). Text Mining: Techniques and its Application. IJETI International Journal of Engineering & Technology Innovations, 1(4), 42.
- [6.] Differentiating between data-mining and text-mining terminology. (n.d.).  
[https://www.researchgate.net/publication/272644253\\_Differentiating\\_between\\_data-mining\\_and\\_text-mining\\_terminology](https://www.researchgate.net/publication/272644253_Differentiating_between_data-mining_and_text-mining_terminology)