11/2024

CUSTOMER SEGMENTATION REPORT

PREPARED BY:

HUNG NGUYEN

Table Of Contents

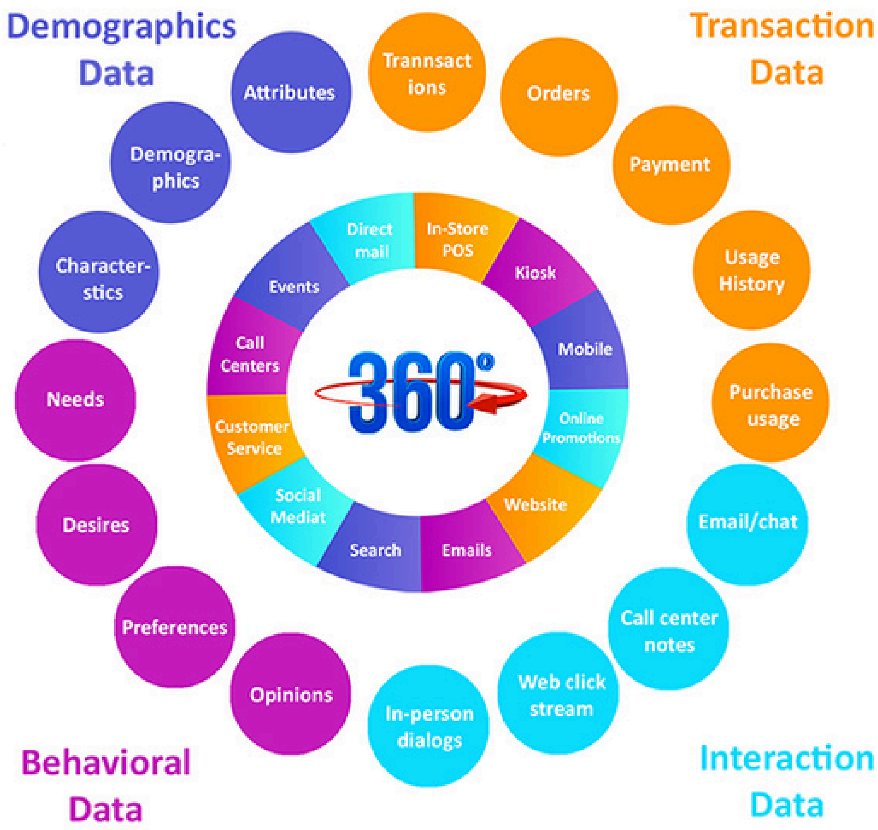
1.Problem definition:

The company's marketing team has just launched a sales-boosting campaign for the end of the year. Now, it is necessary to segment the company's current customer base into smaller, distinct groups to develop tailored policies and strategies that optimize advertising costs while still nurturing customer relationships and achieving the targeted year-end sales goals.

This report is therefore conducted to classify and analyze the company's current customer base using the RFM model metrics. The report will provide the marketing team with a comprehensive and detailed view of the company's customer base, enabling the implementation of appropriate and effective campaigns.

2.Theoretical basis:

2.1.Customer 360:



Customer 360 Model

Customer 360 is a concept in Customer Relationship Management (CRM) that represents a comprehensive and holistic view of a customer based on four key aspects: **Transaction data, Behavioral data, Interaction data, and Demographics data.**

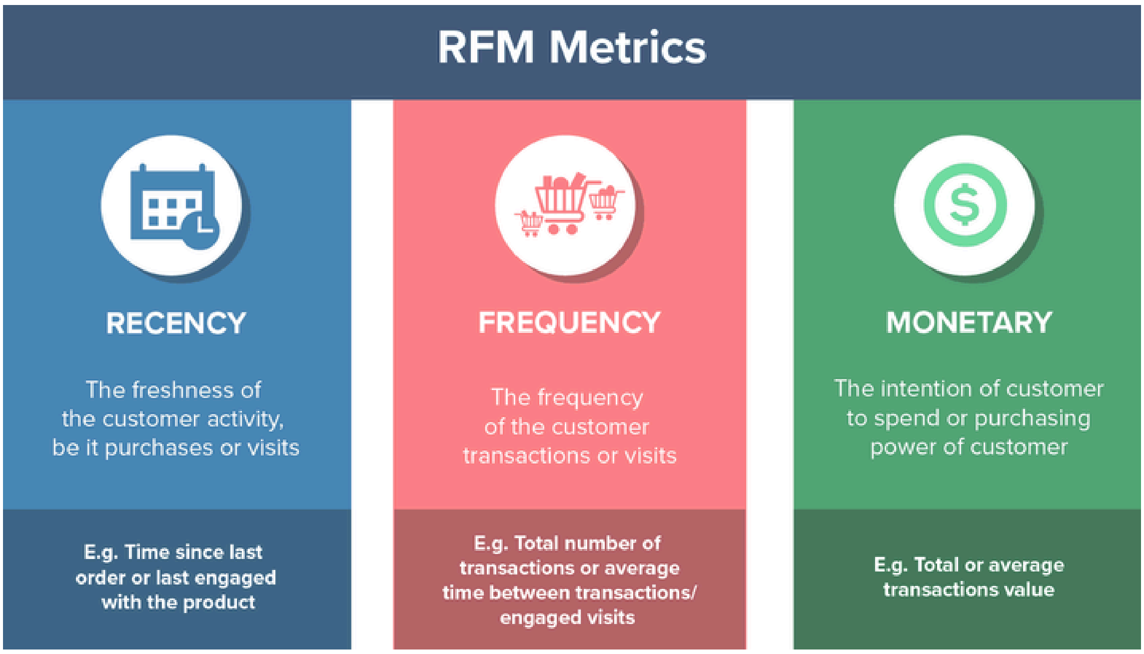
The goal of Customer 360 is to collect data from various sources to create a complete customer profile, enabling businesses to gain deeper insights into customer behaviors, preferences, needs, and characteristics. This understanding allows businesses to market more effectively and deliver better services.

2.2. RFM Model:

RFM is a customer segmentation model based on **Transaction Data** that helps analyze customer behavior and assess their potential value. This enables businesses to develop more effective strategies for engaging with different customer groups.

As a part of the **Customer 360** philosophy, the RFM model focuses on three key factors:

- **Recency**: The time elapsed since the customer's most recent purchase.
- **Frequency**: The number of purchases made by the customer within a specific timeframe.
- **Monetary**: Calculated based on the total spending or the average spending per transaction, depending on the business's analytical goals.



RFM Model

2.3. Inter Quartile:

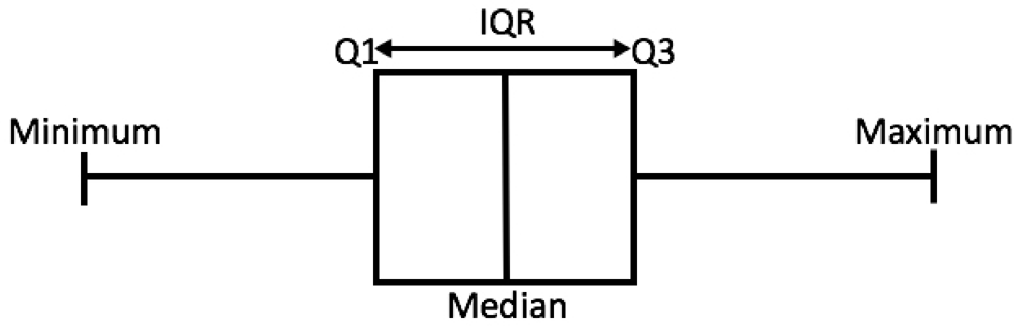
The Interquartile Range (IQR) represents the middle range of a dataset, containing the central 50% of the values. It is commonly used to measure dispersion in cases where the data distribution is skewed. When a dataset is asymmetrical, the median is used instead of the mean to represent the central tendency, making the IQR an appropriate measure for assessing variability.

The IQR is determined based on three key values:

- **Q1 (Lower Quartile):** The value at the 25th percentile of the dataset.
- **Q2 (Median):** The central value that divides the data into two equal halves.
- **Q3 (Upper Quartile):** The value at the 75th percentile of the dataset.
-

The formula for calculating the IQR is:

IQR=Q3–Q1



IQR Method box plot

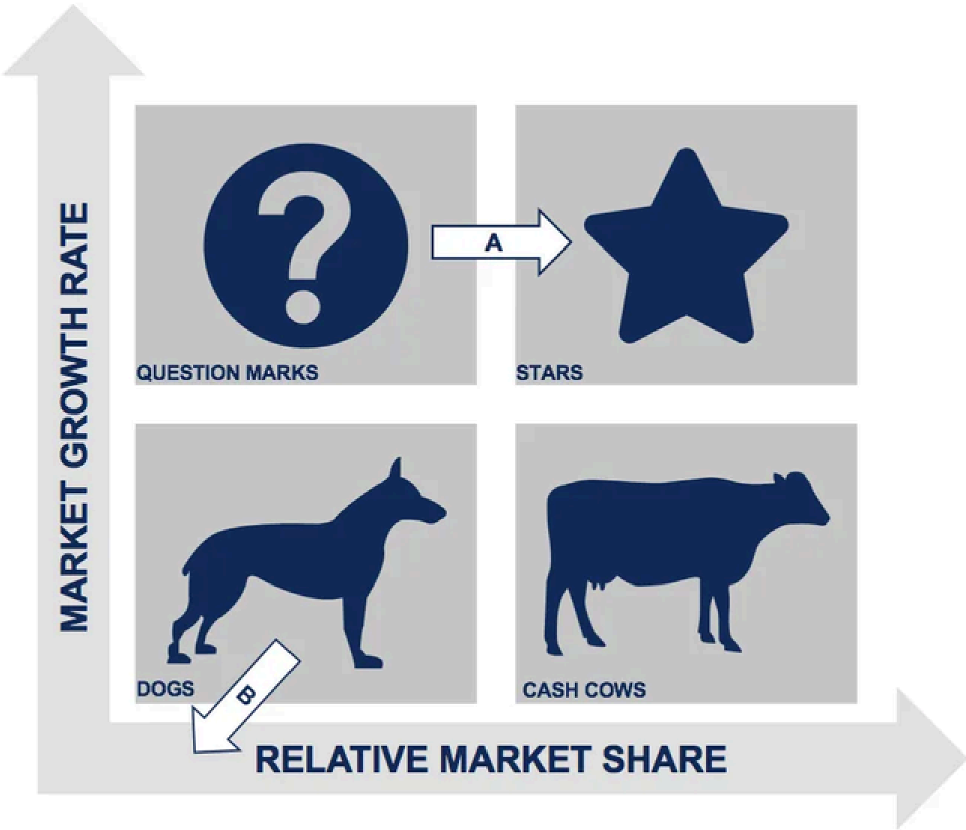
2.4. BCG Matrix:

The BCG Matrix (Boston Consulting Group Matrix) is a strategic tool used to analyze a company's product portfolio. It helps identify the position and potential of products or business units based on two key criteria: **market growth rate** and **relative market share.**

The BCG Matrix categorizes products or business units into four groups within four quadrants:

- **Stars**: Products with high market share in a fast-growing market.
- **Cash Cows:** Products with a large market share in a slow-growing market, providing stable revenue.
- **Question Marks**: Products with low market share in a fast-growing market, requiring investment to develop.
- **Dogs:** Products with low market share in a slow-growing market, often generating limited profit.

In this analysis, **BCG Matrix** will be applied to segment our customers based **RFM Model** introduced above.



BCG Model

3. Dataset overview:

Our dataset includes 2 tables: **customer_transactions** and **customer_registered.**

+ Overview of dim table *customer_registered:*

Column name	Data Type	Description
ID	int	The ID of each customer (Primary Key)
Contract	varchar(50)	The ID of each contract (Unique)
LocationID	int	The ID of location
BranchCode	int	The code of branch
Status	int	The code of status
Created_date	date	The date the contract was created
Stopdate	date	The date the contract stops.

+ Overview of fact table *customer_transaction:*

Column name	Data Type	Description
TransactionID	int	The ID of the transaction.
CustomerID	int	The ID of each customer (Foreign Key)
Purchase_Date	date	The date of transaction conducted
GMV	int	Gross Merchandise value (total amount of sales)

4. RFM marks and customer segmentation:

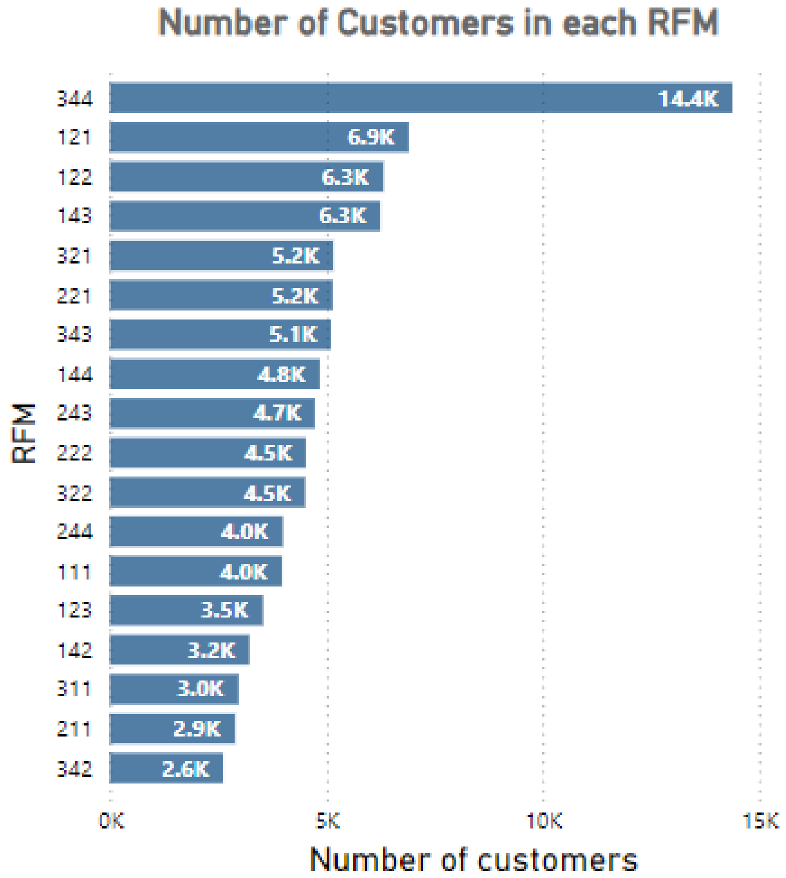
4.1. RFM Marks:

We will use quartiles to create 4 intervals to segment our customer based on 3 criteria R-F-M of RFM models. Each standard will be divided into 4 levels from 1 - 4, proportional to the greatness.

During calculation, the indexes of frequency and monetary are divided by the customer lifetime to ensure the fairness in segmenting customers.

After calculation, we got the table below:

Mark	1	2	3	4
Recency	>=92 days	from 63 to 92 days	from 32 to 62 days	from 1 to 31 days
Frequency	<= 0.00046	above 0.00046 to 0.00063	above 0.00063 to 0.00065	above 0.00065
Monetary	<= 42.81	above 42.81 to 52.53	above 52.53 to 65.29	above 65.29



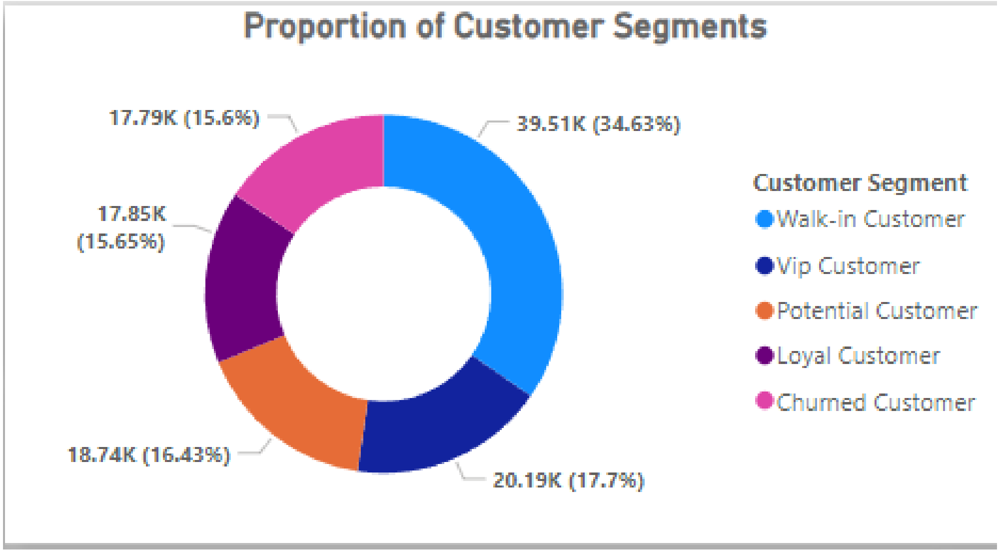
The chart showing the clustering results.

4.2. Customer Segmentation:

Group	Identification	Description	RFM
Star	VIP Customer	Three criteria stay at high level.	'344', '343', '334', '444', '443', '434', '433'
Cash Cow	Loyal Customer	Recency and Frequency often stay at high level while monetary index is considerable	'424', '243', '324', '244', '342', '242', '333', '332', '234', '341', '241', '441', '331', '231', '442', '431', '432', '423'
Question Mark	Potential Customer	Customers who just conducted the purchase with a considerable monetary level while the frequency is still low.	'222', '322', '223', '323', '224', '132', '233', '312', '232', '313', '214', '314', '421', '414', '422', '413'
	Churned Customer	The customers who do not conduct the purchase recently.	'143', '142', '124', '131', '141','133', '144', '134'
Dog	Walk-in Customer	The customers who conducted a few purchase, the recency is low regardless of the monetary level	Leftover

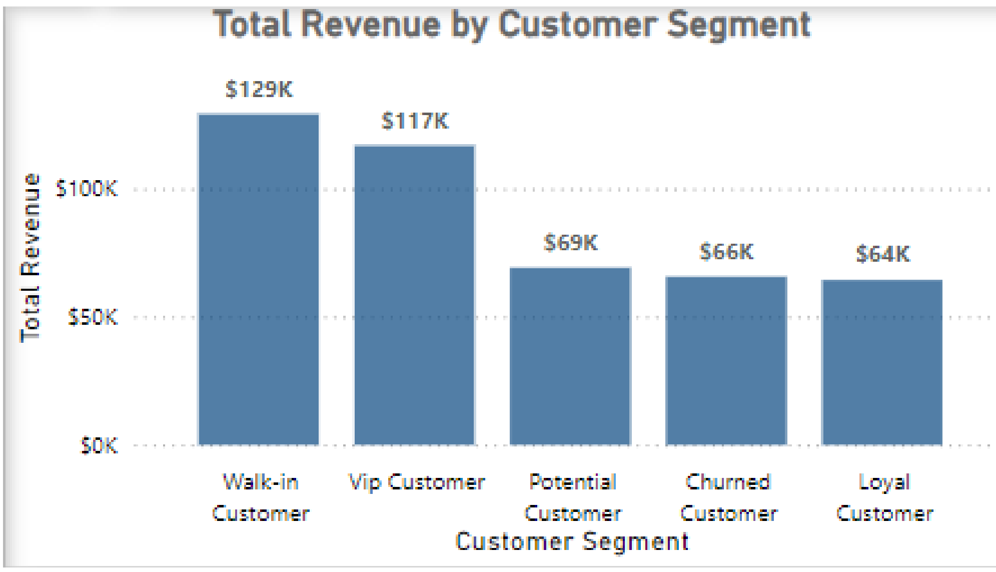
5. Data Analysis:

5.1. Overview of the number of customers in groups

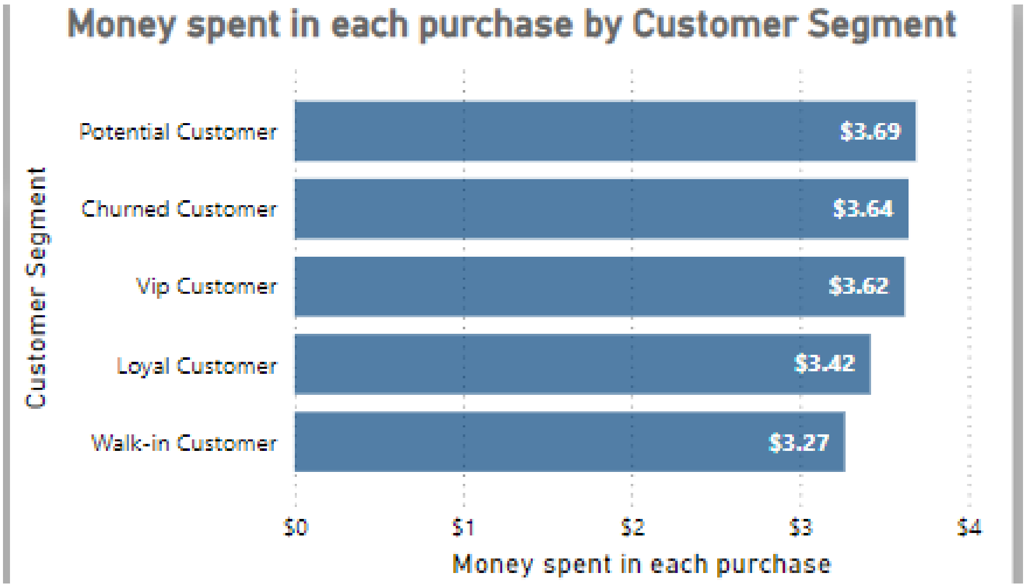


The largest slice is the **"Walk-in Customer"** segment, which makes up **39.51%** of the total customer count. This suggests that a significant portion of the customer base is made up of walk-in or one-time customers, rather than more loyal or recurring customer types.
The next largest segment is the **"Vip Customer"** group, accounting for **20.19% of customers**. This indicates the business has a meaningful base of premium or high-value customers.
The **"Potential Customer"** and **"Loyal Customer"** segments make up **17.85%** and **16.43%** respectively. These seem to represent customers at different stages of the sales funnel or relationship with the business.
The **"Churned Customer"** and **"Loyal Customer"** slices are the smallest, at **5.63%** and **0.39%** respectively. This suggests customer retention may be an area of focus, as the business seems to have a relatively low percentage of truly loyal, repeat customers compared to other segments.

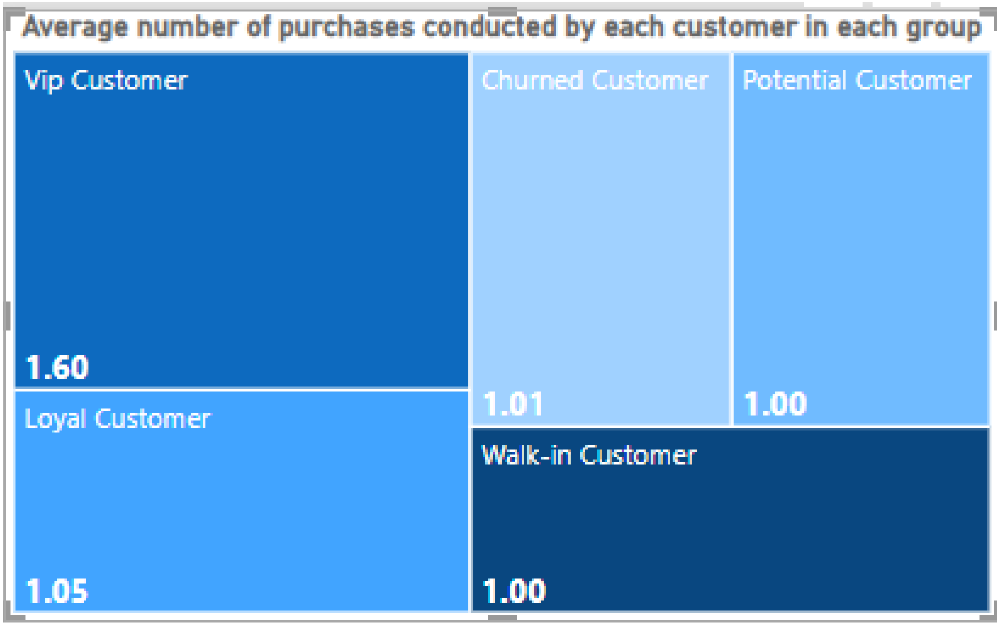
5.2. Overview of the revenue in each group:



The same trend is witnessed in the total revenue bar chart since **walk-in customers** are those who bring about the **largest revenue (129k USD),** followed by **VIP customers.** staying at **117k USD.** This is aligned with the proportion of customers described above. Meanwhile, revenue from potential customers, churned customers along with loyal customers are **apparently the same**, around **65K USD**. Although loyal customer are those conduct considerable purchases, sales from this group is the lowest, requiring company to implement further marketing campaign to boost the purchasing power of this group.



The chart highlights the average money spent per purchase across different customer segments. **Potential Customers** lead with the highest spending at **$3.69,** followed closely by **Churned Customers** at **$3.64** and **VIP Customers** at **$3.62**. **Loyal Customers**, despite their engagement, spend slightly less at **$3.42,** while **Walk-in Customers** record the lowest average spending at $3.27. This suggests that segments with lower engagement, such as **Potential** or **Churned Customers,** tend to make higher-value purchases, whereas **Loyal** and **Walk-in Customers** contribute smaller purchase amounts on average.



The **"VIP Customer"** segment has the highest average at **1.60** purchases per customer. This indicates that while the number of VIP customer could not be high, their monetary values are large and each of customer in the category is likely to conduct more than **1.00 purchase**. Meanwhile, the number for other groups are quite the same, approximately **1.00.** Especially, the **"Walk-in Customer"** segment also has an average of **1.00 purchases** per customer, the lowest among the groups shown. This aligns with the earlier insight that a significant portion of the customer base are one-time or infrequent walk-in customers.

6. Suggestions:

1. Leverage the Potential of Walk-in Customers

- To convert walk-in customers into more profitable customer groups, the following strategies can be implemented:
- Develop an appealing loyalty points system to collect customer data and track their behavior.
- Offer regular exclusive perks for registered members to encourage long-term engagement.

2. Enhance the Experience of VIP and Loyal Customers

- To maximize the value from VIP and loyal customers:
- Personalize customer service to cater to individual preferences.
- Create exclusive privileges, such as early access to new services.
- Host periodic appreciation events to strengthen customer relationships.
- Objective: Increase the average spending of these groups by 15%.

3. Retention Strategies for At-Risk Customers

- To address customers likely to churn:
- Conduct surveys to understand dissatisfaction and identify improvement areas.
- Provide attractive win-back offers to incentivize their return.

4. Activate Potential Customers

- To engage potential customers and encourage their conversion:
- Design affordable trial packages to let them experience the product or service.
- Create targeted content marketing campaigns tailored to this segment's characteristics.
- Implement a referral program (member-get-member) to attract new customers through existing ones.

7. SQL Code:

Code was updated in the link here.