NOVEMBER, 2024

CUSTOMER SEGMENTATION REPORT

RFM Model

KIEN HUYNH

Table of Contents

I. Introduction

The marketing team has launched a year-end sales campaign and now requires a segmentation of the current customers into distinct groups. This segmentation will provide better insights about our current customers as well as lost customers so that the team can come up with diferent approaches that align with the needs of each segment.

II. Literature Review

1. Customer 360

Customer 360 refers to a 360 degree view of customer data, including all customer interactions, from a website inquiry to a product purchase. Through the view, every group in the company whether it's marketing, sales, customer service, or support has access to the same data about each customer, which includes basic information (name, contact details, and demographics), interactions (purchases, website visits, inquiries, and social media interactions), behavioral data (preferences, opinions, and feedback), and relationships (family members).

Customer 360 can help the company to:

Provide a better understanding of customer insight:

The approach helps elimiating duplicate, fault, and unprecise information. Thus, it establishes a unified accurate data for have a comprehensive view of the business's customers.

Promote cross-selling and upselling:

With such insights, businesses can recommend related products or services. They can also recognize connections between customers in order to expand sales across customer groups within the same ecosystem.

Increase Retention and Loyalty:

Business now can divide their customers into small segments. Through that, they can provide personalized experiences and product suggestions, which helps build long-term loyalty and improves customer retention.

Improve agility:

Businesses can quickly identify emerging trends and shifts in consumer demand allowing faster decision-making and a more responsive approach to market changes.

2. RFM model

The RFM (Recency, Frequency, and Monetary) Model is a behavior-based model used to analyze the behavior of a customer. The model consists of three measures combined into a three-digit score, which is later divided into four equal quartiles (25%) group. It is used to assess when, how often, and how much customers buy, helping businesses predict future purchasing behavior. By analyzing past purchases, companies can identify which customers are worth targeting.

Recency: Customers who have recently made a purchase are more likely to return for future transactions. Marketing efforts can encourage recent customers to revisit, while also targeting lapsed customers with reminders and incentives to re-engage.

Frequency: The number of purchases a customer makes can depend on product type, price, and need for replenishment. Marketing can focus on reminding customers to repurchase items when their needs arise, such as when groceries run low.

Monetary Value: The amount a customer spends is crucial. Businesses often prioritize high-spending customers for marketing and loyalty efforts, but should also consider maintaining relationships with regular, lower-spending customers.

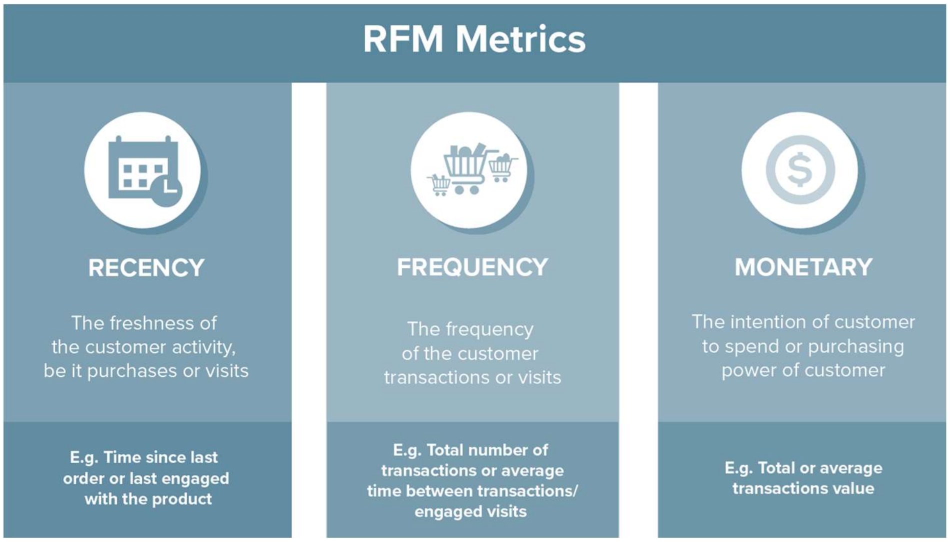


Figure 1: RFM Model

3. Quartiles

Besides median, which divides a distribution in half, we can also split the distribution into smaller equal parts. It can be split into quarters (Quartiles), fifths (Quintiles), tenths (Deciles), or even into hundredths (Percentiles). For quartiles, the first quartiles (Q1), also known as 25^{th} percentile, divides the distribution such that 25% of the observations lie at or below it. Similarly, the Q2 or median, represents 50% of the observations lie at or below it, and Q3 represents the 75^{th} percentile. The position of four quartiles in an array with n entries sorted in ascending order can be calculated from this formula:

$$L_y = (n+1)\frac{y}{100}$$

Where y is the percentage point at which we are dividing the distribution (25%, 50%, and 75% for Q1, Q2, and Q3).

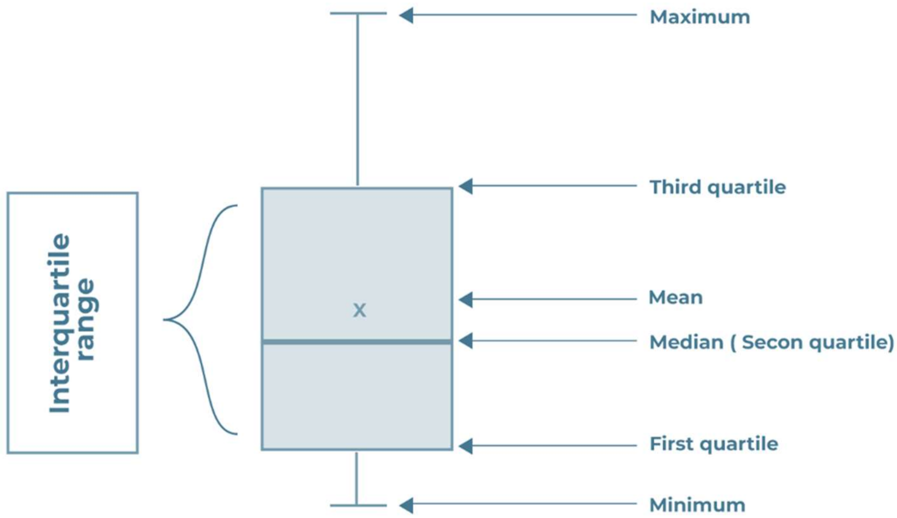


Figure 2: Quartiles Box Plot

4. BCG Matrix (Boston Consulting Group Matrix)

BCG Matrix, also known as the product portfolio matrix, is a strategic planning tool designed to evaluate the strategic position of a firm's brand portfolio. As one of the most widely used methods for portfolio analysis, the BCG Matrix classifies a firm's product and/or services into a two-by-two matrix. The classified quadrants are called **Stars**, **Question Marks**, **Cash Cows**, and **Dogs**.

Though, BCG matrix is typically used for analyzing products or business units. It can be adapted for customer segmentation. When applied to customers, the matrix helps businesses categorize their customer base into different segments based on two key dimensions: customer loyalty (market share) and profitability or growth potential (market growth).

In this context:

- **Stars** are high-value customers who are both loyal and have high potential for growth or revenue generation.
- **Cash Cows** are customers who are very loyal and consistently profitable but may not have significant growth potential.
- **Question Marks** are customers with high growth potential, but they are not yet loyal or profitable, requiring more investment or attention to convert them into loyal, high-value customers.

➢ **Dogs** are customers who are neither loyal nor profitable, and they may need to be reevaluated or even considered for removal from the customer base.

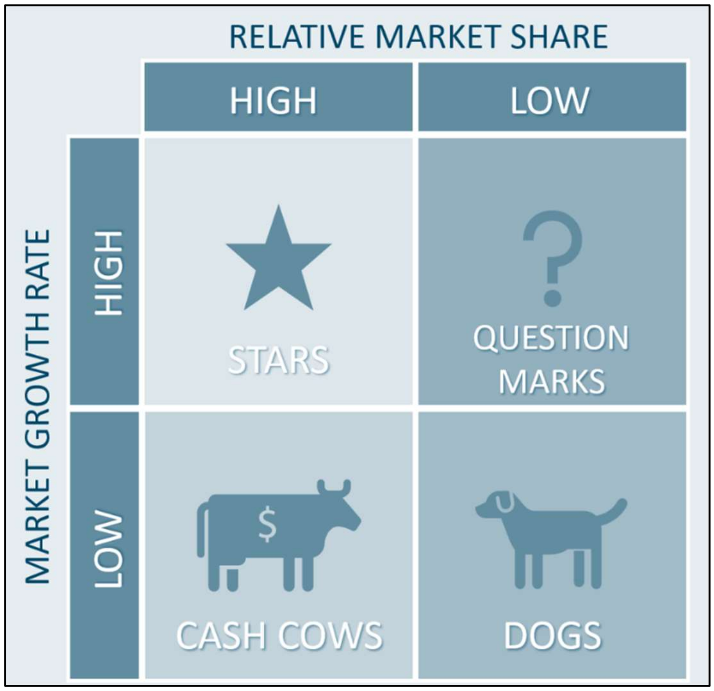


Figure 3: BCG Matrix

III. Methodology

The following steps were taken to segment customers based on Recency, Frequency, and Monetary values:

Recency: To calculate recency, the number of days between each customer's last purchase date and September 1, 2022, was determined. This time difference reflects how recently a customer made a purchase. These values were then ranked into four groups (1 to 4) based on quartiles, with customers in the lower quartiles representing those who made recent purchases.

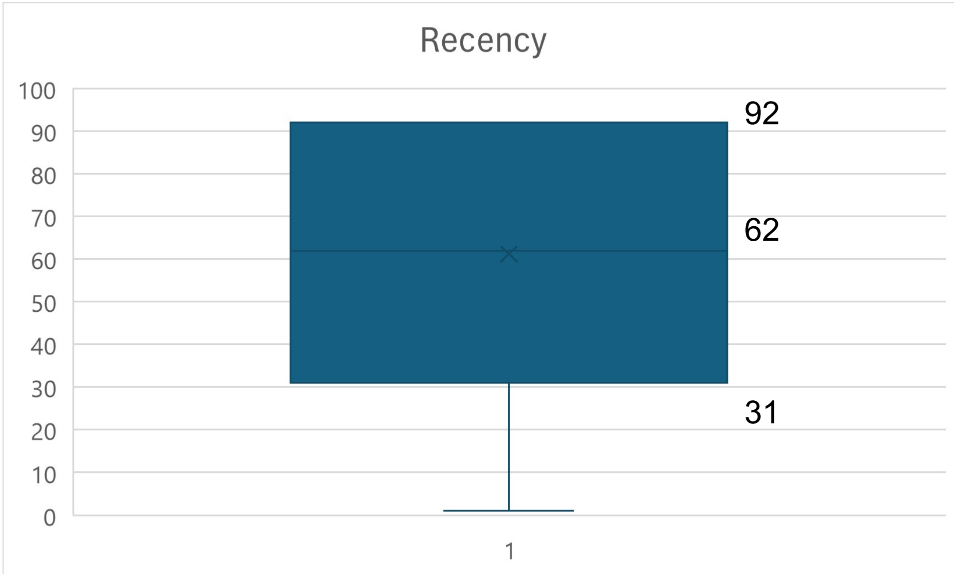


Figure 4: Recency Box Plot

Frequency: For frequency, the total number of transactions for each customer was counted and then divided by the customer's age (calculated as the difference between the customer's account creation date and September 1, 2022). This provided the average number of transactions per year for each customer. The frequency values were then ranked into four quartiles, with customers in higher quartiles being those who made more frequent purchases.

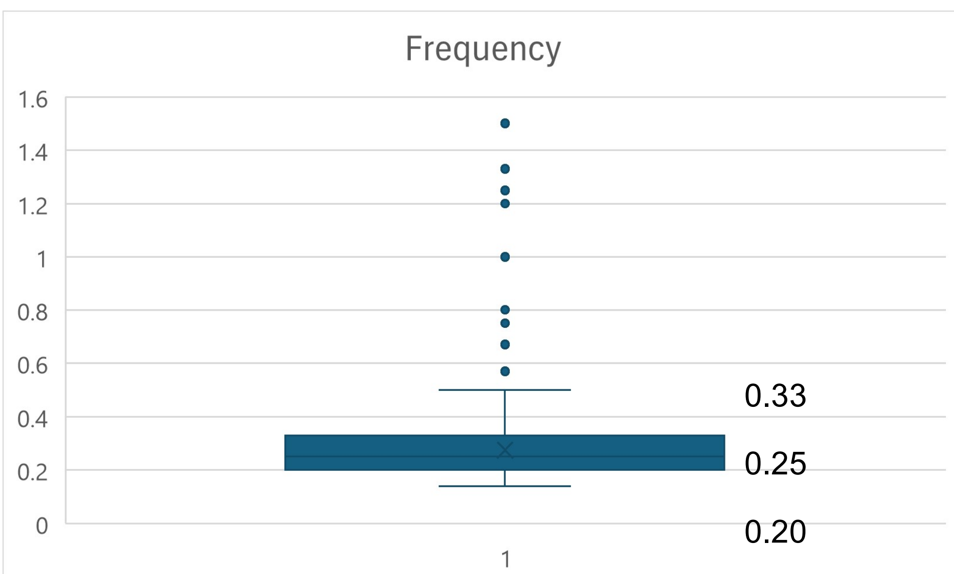


Figure 5: Frequency Box Plot

Monetary: To calculate monetary value, the total Gross Merchandise Value (GMV) spent by each customer was calculated, divided by the customer's age in years (since account creation). This provided the average yearly spend per customer. These values were also ranked into four quartiles, with customers in higher quartiles representing those who spent more on average annually.

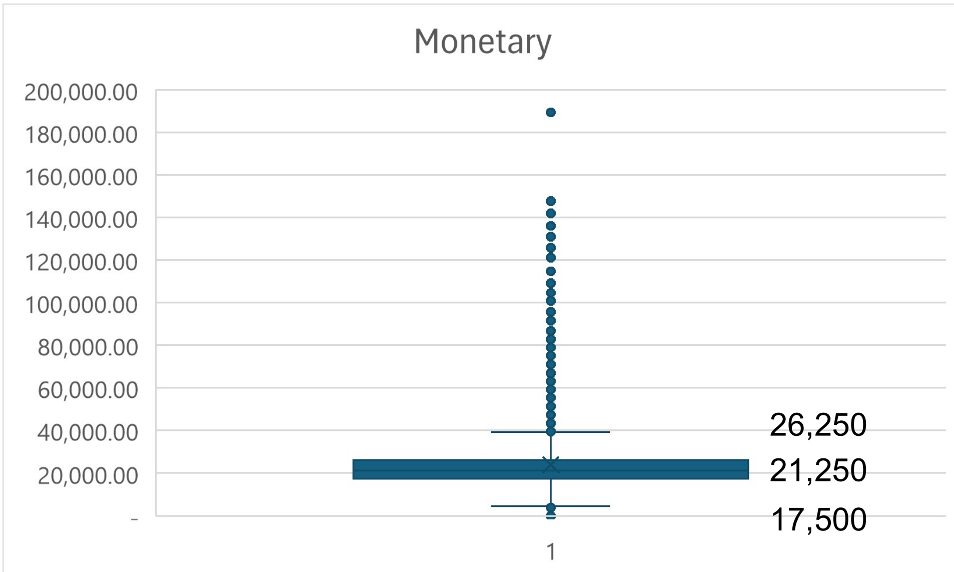


Figure 6: Monetary Box Plot

By ranking customers into quartiles for each of the three dimensions (Recency, Frequency, and Monetary), customers were segmented based on their purchasing behavior, which allowed for the identification of valuable customer groups for targeted marketing strategies.

IV. Summary Statistic

	Recency	Frequency	Monetary
Mean	61.25	0.27	23,940.78
St.dev	26.12	0.10	10,834.51
Max	92	1.5	189,394.00
Min	1	0.14	-
Range	91	1.36	189,394.00
Median	62	0.25	21,250.00
Quartile 25%	31	0.2	17,500.00
Quartile 75%	92	0.33	26,250.00

Table 1: Summary Statistic

The table presents key statistical data on Recency, Frequency, and Monetary values for customer transactions. On average, customers made their last purchase 61.25 days ago, with a wide variation in recency, ranging from 1 day to 92 days. Most customers made their last purchase within the last two months (median of 62 days). The frequency of customer purchases is low, with a mean of 0.275 transactions, indicating that, on average, customers purchase roughly once every four years. The frequency ranges from 0.14 to 1.5, reflecting that while most customers purchase infrequently, some outliers have a much higher frequency of transactions, as seen with the 1.5 max value. In terms of spending, the average customer spends 23,940.78 per annum, with values ranging from 0 to 189,394. The median spending is 21,250, showing that most customers spend less than this amount, while the top 25% of spenders contribute more, with a 75th percentile value of 26,250. The high standard deviations in recency, frequency, and monetary values suggest significant variability in customer behavior, emphasizing the need for targeting segmentation and targeting marketing strategies.

V. Results

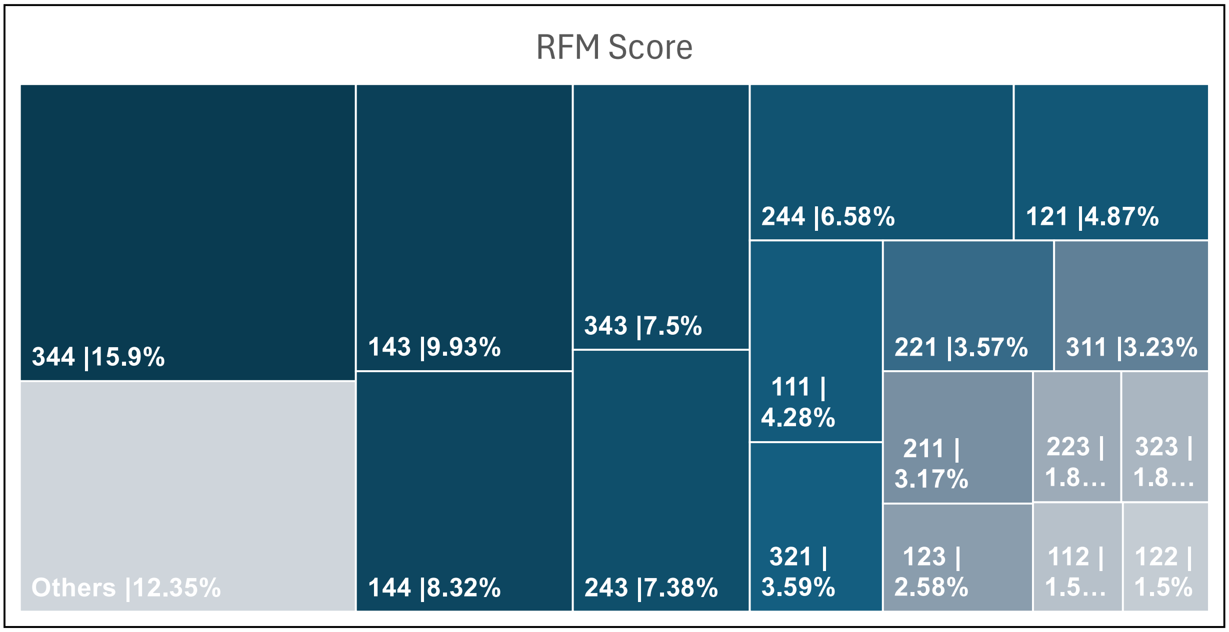


Figure 7: RFM Score Treemap

The RFM score table reveals a highly uneven distribution of customers across segments, with the top segments accounting for a significant portion of the customer base. For example, the largest segment, 344, comprises 15.90% (18,142 customers), followed by 143 (9.93%), 144 (8.32%), and 343 (7.50%), collectively making up a substantial share of the total customers. These high RFM segments represent loyal and high-value customers who are likely to drive revenue and should be prioritized for retention strategies. Meanwhile, low RFM scores like 111,112, and 121, though significant in count, may require reactivation efforts to prevent further attrition. Overall, the data reflects a classic Pareto Principle, where a few top-performing segments drive most of the value, emphasizing the need for an approach that maximizes returns from key groups while addressing weaknesses in lower-performing ones.

Applying the adapted BCG Matrix, we have:

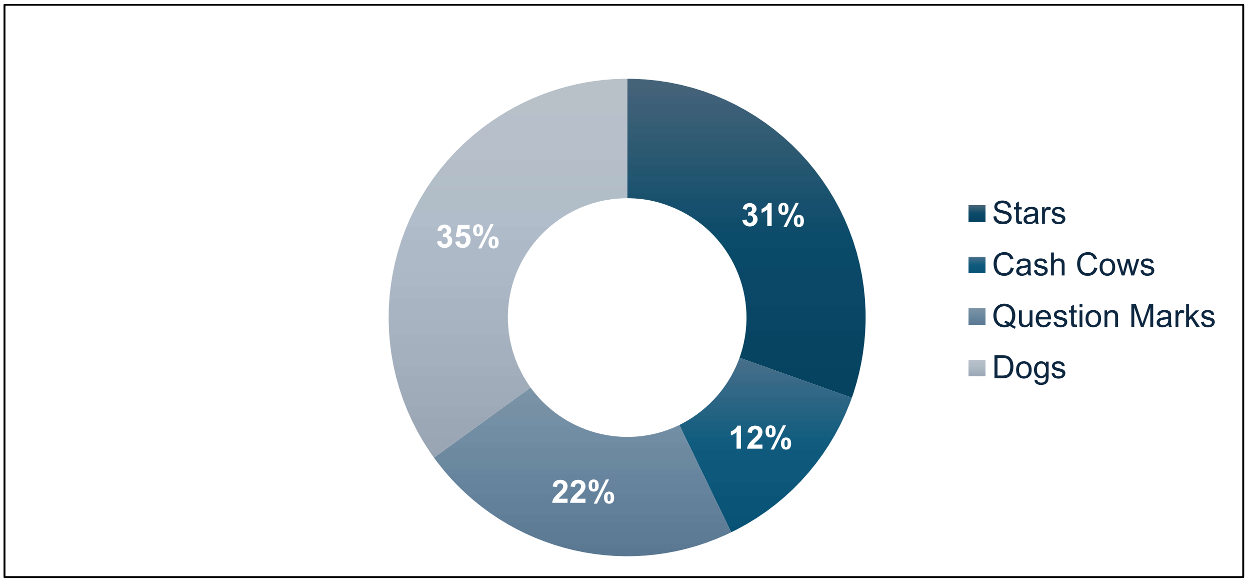
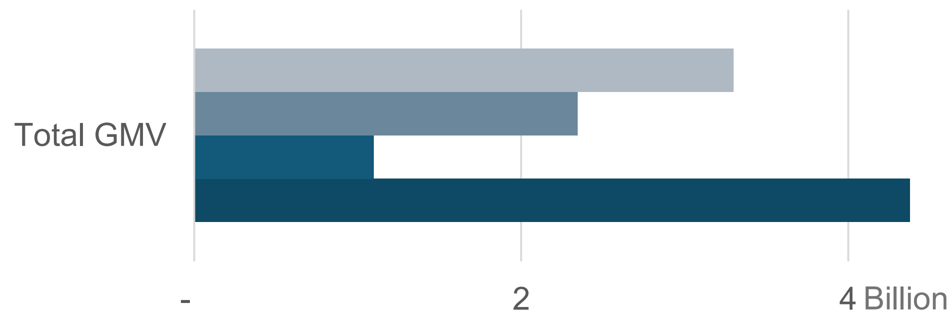


Figure 8:BCG Customer Segment Percentage

The customer segment breakdown offers valuable insights about our current customers, which shows potential opportunities for targeted strategies. The largest segment, Stars (30.46%), the second-largest segment, represents highly valuable and loyal customers who contribute significantly to revenue. These customers should be targeted with premium offerings or loyalty programs to maintain their engagement. Question Marks (22.22%), highlights customers with potential for growth but needing additional investment to become more engaged. Hence, targeted marketing is needed to convert them into loyal Stars. Cash Cows (12.35%) consist of stable, profitable customers, while Dogs (34.97%) represent customers with low engagement and growth potential, requiring reactivation efforts to prevent further attrition. Therefore, by providing equal attention to Stars and Cash Cows while focusing more on Question Marks and carefully managing or reducing the Dogs segment, businesses can optimize both short-term revenue as well as long-term growth.



	Total GMV
Dogs	3,299,707,711
Question Marks	2,346,688,712
Cash Cows	1,098,281,558
Stars	4,379,583,217

Figure 9: Total GMV Group By BCG Segments

The total GMV (Gross Merchandise Value) distribution sorted by the customer segments indicates important insights into revenue generation. Stars generates the highest GMV, contributing a substantial 4.38 billion, highlighting their position as the most valuable segment in terms of revenue. Question Marks follows with 2.35 billion, indicating that although they have potential for growth, they currently contribute less in GMV but may provide significant opportunities with the right investments. Dogs, despite representing a large portion of the customer base, contribute 3.30 billion in GMV, showing that while they may have lower growth potential, they still play a significant role in the revenue stream. Cash Cows, with 1.10 billion, provide steady and reliable revenue, although their growth potential is limited.