

RFM analysis of the pharmacy chain's customer base

Introduction

Motivation for analysis: Pharmacy revenue typically declines during the off-season. The company has an SMS messaging system in place and wants to use it more effectively. To ensure the budget is spent wisely, the company plans to send personalized messages to customers.

The purpose of this analysis is to recommend ways to improve the effectiveness of SMS messaging.

Objectives:

1. Analyze the pharmacy chain's customer base using RFM methodology, enabling classification of customers according to purchase behavior.
2. Provide targeted recommendations on product and service offerings for each customer segment.

What is RFM analysis?

RFM analysis is a technique used to segment customers based on their purchasing behavior. It helps identify which customers generate the most revenue for the company.

The abbreviation RFM stands for:

- **Recency** - how recently a customer made a purchase,
- **Frequency** - how often a customer makes purchases,
- **Monetary** - the total value of a customer's purchases (the total amount of purchases made by customers)

For each of these dimensions, customers are divided into three equal groups and assigned a score from 1 to 3.

Recency (time since last purchase)	Frequency (purchase frequency)	Monetary (purchase amount)
1 - recent customers	1 - frequent	1 - large
2 - occasional / irregular	2 - occasional	2 - medium
3 - long ago, possibly lost	3 - very rare	3 - small

After dividing people into groups, we obtain up to *27 segments*.

This analysis highlights customers 1) who purchase frequently and in large volumes, 2) those who make regular small purchases, and 3) those who have not bought anything for a long time.

Targeting messages to specific customer groups leads to much higher conversion rates than generic, non-personalized advertising.

Database structure

The analysis was conducted using the **Bonuscheques** table (***datetime, shop, card, bonus_earned, bonus_spent, summ, summ_with_disc, doc_id***) from the pharmacy database. This table contains *records of transactions* made with bonus cards, which are linked to *regular customers* who provided their phone numbers.

datetime -> date and time of the transaction.

summ -> transaction (check) amount.

If a transaction was processed while the cash register was offline, the system recorded an encrypted sequence of characters instead of the actual card number. Since such records cannot be linked to specific customers, they were excluded from the analysis.

Additionally, customers who made only a single purchase were excluded. Since personalized discounts are intended to encourage repeat purchases and increase the average transaction value, offering them to one-time buyers would not be meaningful.

RFM analysis metrics: segmentation thresholds

This analysis assessed customer distribution by calculating (1) means and medians, (2) percentiles, and (3) the cumulative sum and order count:

- The **mean** is widely used to describe data, but it is sensitive to extreme values, which can skew the result. The **median**, on the other hand, reflects the middle value of the data and provides a more reliable measure when outliers are present.
- **Percentiles** split the dataset into layers, helping to see how customers are grouped across different levels. This approach shows differences more precisely but tends to smooth out outliers. For even finer detail, the distribution was also examined by **deciles**.
- In addition, the **cumulative sum** and **number of orders** were calculated to apply *Pareto's 80/20 rule*.

RFM analysis

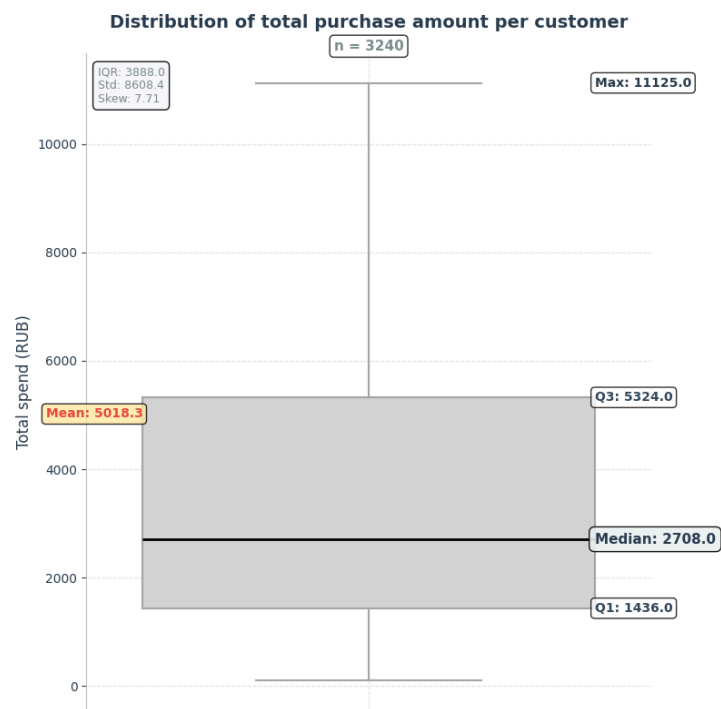
Period covered by the analysis: July 2021 - June 2022.

The selected timeframe enables a reliable retrospective analysis and helps ensure that marketing efforts are not directed at customers who may no longer be reachable due to:

1. relocation,
2. a change in workplace that reduces the likelihood of visiting the pharmacy, or
3. updated contact information (e.g., a new phone number)

In our case, the mean purchase amount is 5018 rubles, while the median is 2708 rubles.

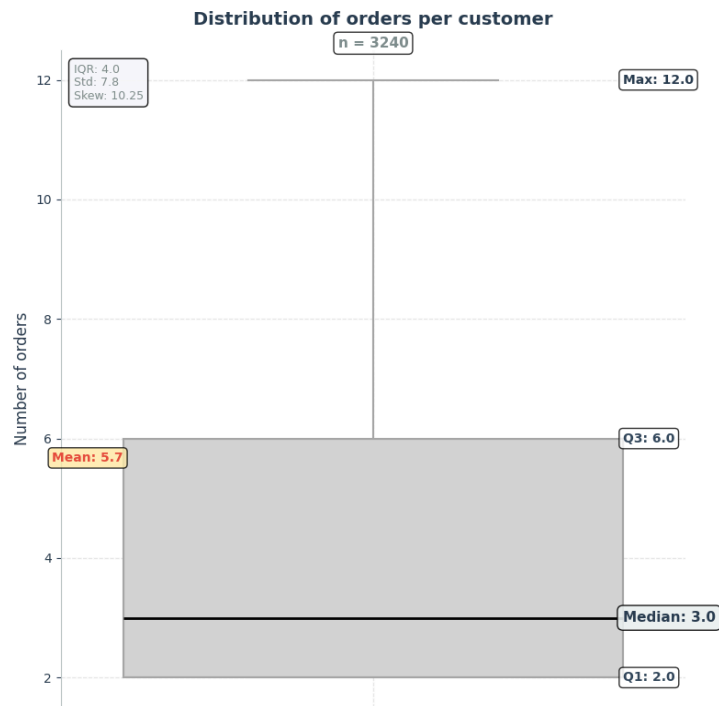
The large gap between these figures indicates the presence of outliers, likely resulting from a small number of customers making exceptionally large purchases.



Picture 1: Distribution of total purchase amount per customer (box plot)

The mean number of purchases per customer is 5.7, while the median is 3, indicating a right-skewed distribution.

This suggests that a substantial portion of customers make only a single purchase, creating an opportunity to focus retention efforts on converting these one-time buyers into regular customers.



Picture 2: Distribution of orders per customer (box plot)

Recency

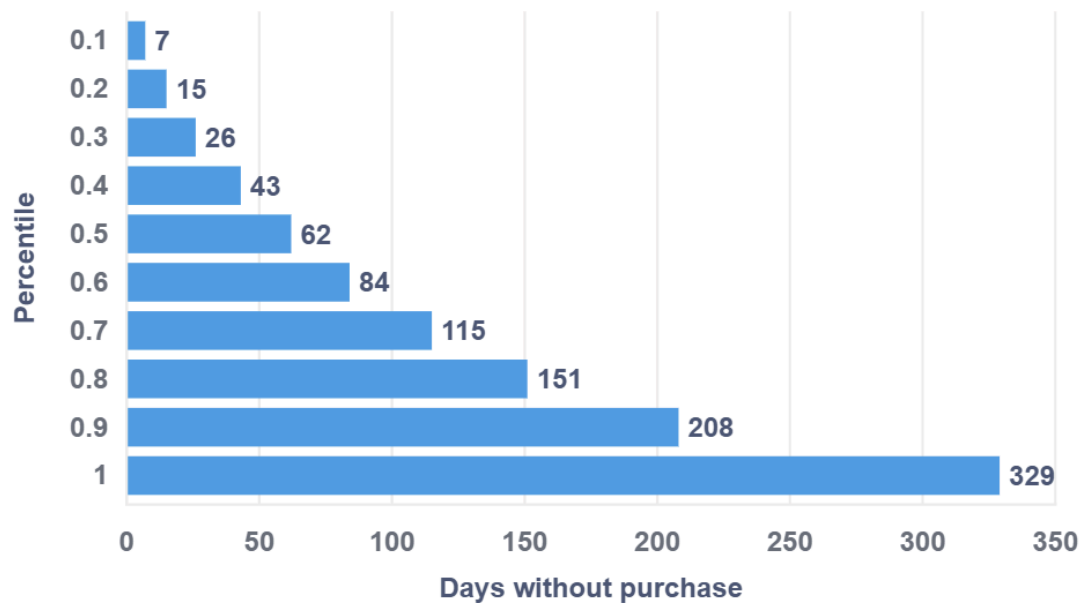
The decile distribution indicates that 30% of customers made their most recent purchase within the last 26 days. This segment includes both repeat and newly acquired customers who have recently engaged with the pharmacy chain. It represents the most recent customer group (**Group 1**).

The median recency is 62 days, while the mean recency is 85 days -> roughly 30% of customers fall into **Group 2**, which corresponds to customers with completion times clustered around or above the mean.

Those who have not visited a pharmacy in over 85 days (approximately three months) fall into **Group 3**. This group likely includes

- infrequent pharmacy users (e.g., younger individuals),
- those for whom visiting a pharmacy has become geographically inconvenient, or
- those who have switched to competitors.

Distribution of days since last purchase



Made with  Metabase

Picture 3: Distribution of days since last purchase

Frequency

One-time customers represent 32.6% of the customer base but account for only 11.5% of total visits. This low level of engagement classifies them as the least active segment (**Group 3**).

The median purchase frequency is 3, representing 50.99% of customers and forming the largest segment. The mean is 6. The cumulative distribution indicates that 76.6% of customers made between 3 and 6 purchases during the period. This segment is classified as **Group 2**.

During the review period (July 2021 - June 2022), 23.43% of customers made 7 or more visits. This group accounted for 57.36% of all purchases. These loyal and frequent customers are designated as **Group 1**.

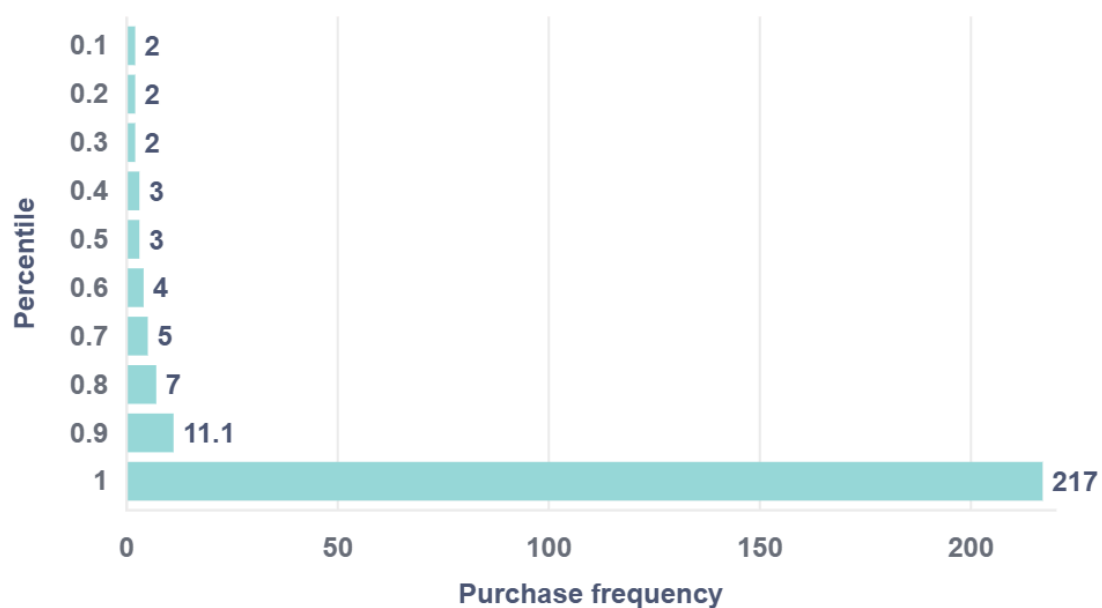
Cumulative contribution by purchase count

total_purchases	customers_count	total_visits	cum_customers_percent	cum_visits_percent
2	1,056	2,112	32.59	11.49
3	596	1,788	50.99	21.21
4	374	1,496	62.53	29.35
5	285	1,425	71.33	37.09
6	170	1,020	76.57	42.64
7	151	1,057	81.23	48.39
8	95	760	84.17	52.52
9	82	738	86.7	56.54
10	59	590	88.52	59.75
11	48	528	90	62.62

58 rows

Table 1: Partial view of 'Cumulative contribution by purchase count' [obtained from Metabase]

Distribution of the number of purchases per customer



Made with  Metabase

Picture 4: Distribution of the number of purchases per customer

Monetary

The median-to-mean ratio points to outliers, so the mean is not fully reliable. Analysis of the distribution and cumulative totals shows that 20.03% of customers (spending ≥ 6346 rubles) generate 60.44% of total sales. These customers are designated as **Group 1**.

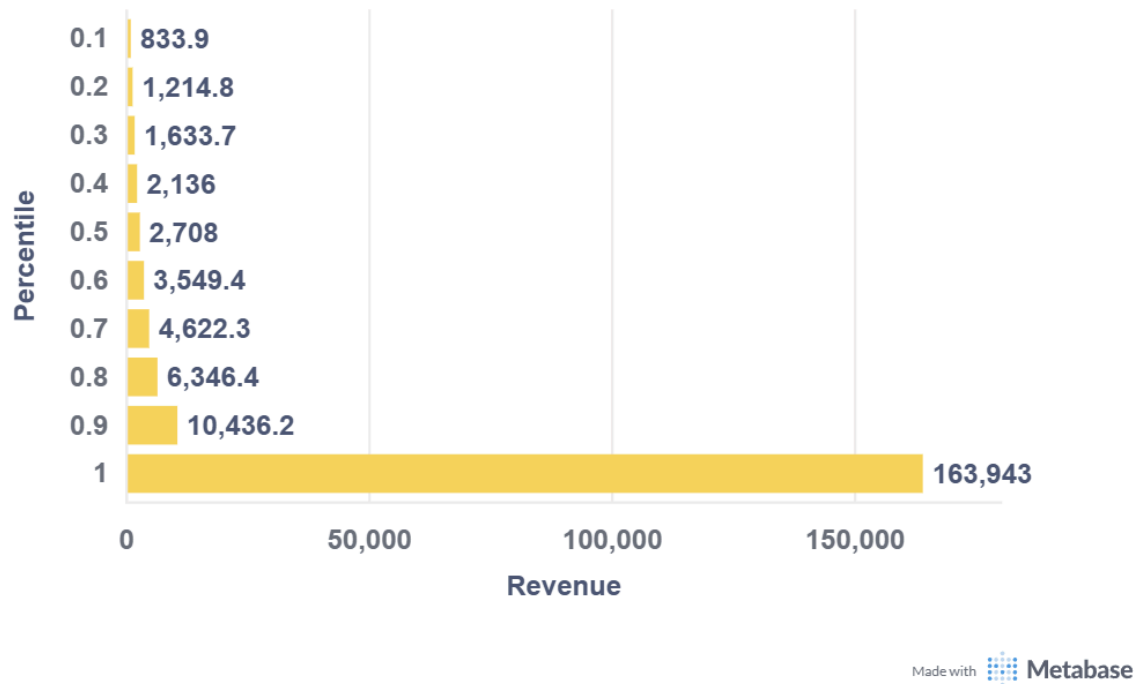
Group 2 includes the majority of customers, with spending between 2136 and 6345 rubles. Together with Group 1, they generate 90% of total sales (cumulative) and represent 60% of the customer base (fourth decile). Both the median (2708 rubles) and the mean (5018 rubles) fall within this range.

Group 3 represents the remaining 40% of customers, contributing only 10% of total sales.

Contribution of customers to revenue		
customers_share_percent	revenue_share_percent	revenue_threshold
20.03	60.44	6,346
1 row		
Contribution of customers to revenue		
customers_share_percent	revenue_share_percent	revenue_threshold
60	90.21	2,136
1 row		

Table 2: Partial view of ‘Contribution of customers to revenue’ [obtained from Metabase]

Distribution of total revenue per customer



Picture 5: Distribution of total revenue per customer

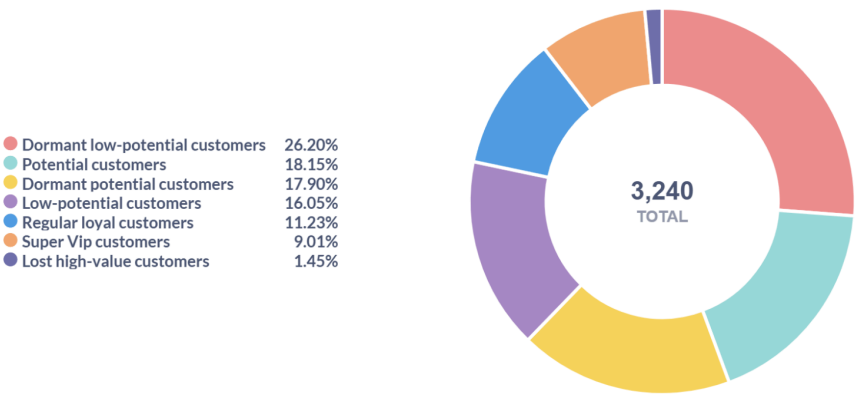
Segmentation thresholds

In the analysis, customers were segmented into three groups according to each of the three parameters. The resulting segments are:

Group	Recency	Frequency	Monetary
1	≤ 26 days since last purchase	≥ 7 purchases	≥ 6346 rubles
2	≤ 84 days since last purchase	3 - 6 purchases	2136 - 6345 rubles
3	> 84 days since last purchase	2 purchases (very infrequent)	≤ 2135 rubles

To simplify customer segmentation, very small segments or those requiring similar marketing strategies were merged. Managing 27 distinct segments with only 10-50 customers each would be resource-intensive and increase the risk of messaging errors. To streamline operations, customers were consolidated into seven logical groups.

RFM analysis: segment shares (%)



Made with  Metabase

Picture 6: RFM analysis - segment shares (%)

logic_group	rfm_group	customers_count	group_total
Super Vip customers	111	292	292
logic_group	rfm_group	customers_count	group_total
Regular loyal customers	112	123	364
Regular loyal customers	121	38	364
Regular loyal customers	122	203	364
logic_group	rfm_group	customers_count	group_total
Lost high-value customers	311	47	47

logic_group	rfm_group	customers_count	group_total
Potential customers	211	143	588
Potential customers	212	86	588
Potential customers	221	50	588
Potential customers	222	240	588
Potential customers	231	5	588
Potential customers	232	64	588

logic_group	rfm_group	customers_count	group_total
Low-potential customers	113	5	520
Low-potential customers	123	114	520
Low-potential customers	131	3	520
Low-potential customers	132	55	520
Low-potential customers	133	149	520
Low-potential customers	213	9	520
Low-potential customers	223	185	520

logic_group	rfm_group	customers_count	group_total
Dormant potential customers	312	51	580
Dormant potential customers	321	53	580
Dormant potential customers	322	317	580
Dormant potential customers	331	7	580
Dormant potential customers	332	152	580

logic_group	rfm_group	customers_count	group_total
Dormant low-potential customers	233	188	849
Dormant low-potential customers	313	3	849
Dormant low-potential customers	323	225	849
Dormant low-potential customers	333	433	849

Table 3: RFM analysis - logic group and RFM groups [obtained from Metabase]

Analysis summary: key recommendations and insights

- The RFM analysis of the pharmacy chain's customer base resulted in seven customer segments for targeted SMS campaigns, aimed at increasing average purchase value and purchase frequency.

The resulting distribution of customers closely aligns with a Pareto-like pattern, which is particularly effective for optimizing limited human and financial resources.

- **Strategic focus:**
 - Increase pharmacy visit frequency
 - Increase average transaction value (average check)

This can be achieved by guiding customers through the loyalty marketing funnel:

- Turn “cold” customers into “warm” regulars.
- Convert “warm” customers into “hot” ones.
- Encourage “hot” customers to make purchases and transition them into VIPs.
- Transform VIPs into brand ambassadors who attract their relatives and friends to the pharmacy.

Now, in more detail:

1. **Build broad top-of-funnel coverage by re-engaging two groups of dormant customers and attracting new ones.**

The RFM analysis shows that the pharmacy chain has many infrequent customers. Converting these customers into regular customers should be a key priority, as even those with lower average purchase amounts contribute to revenue and represent potential loyal customers.

Dormant potential customers (17.9%)

This group consists of customers who previously made purchases of significant amounts but have since become inactive. It is one of the most important groups, as they have the potential to be reactivated as active customers. To re-engage them, the company should send targeted reminders, such as:

- Promotions like “double bonuses until [date X]”.
- Notifications that additional bonuses have been credited to their card and will expire on [date Y].
- Offers to reduce the cost of ABC medicines until [date X].

In other words, it is important to encourage these customers to visit the pharmacy before the specified deadlines.

Dormant low-potential customers (26.2%)

This group consists mainly of one-time customers who have not made purchases at the pharmacy for a long time, aside from some minor transactions.

Several factors may explain their long inactivity:

- Relocation - customers who have moved may be unwilling to travel to a pharmacy far from home or work.
- Contact details changed - if customers' phone numbers have changed, marketing efforts may be wasted.
- Negative customer experience - some customers may have been dissatisfied with the services provided by the pharmacy.
- Demographics - dormant low-potential customers may be young and healthy individuals who rarely need to visit pharmacies.

Risks: re-engaging this group would require a significant advertising budget, with uncertain returns. Even if some customers do come back, it is unclear whether the investment would be justified.

Opportunities: There is still potential to convert this group into regular customers. Many of them have made two purchases at the pharmacy - essentially casual customers. However, if a customer has returned even once, are they truly casual? Several scenarios are possible:

- **Proximity and loyalty** - the customer may live near the pharmacy but typically visits another one. Loyalty could be developed through promotions on parapharmaceuticals, discounts for registering in the pharmacy app, or welcome points

for joining the loyalty program.

- **Lack of awareness** - the customer may not realize that the pharmacy sells not only medicines but also parapharmaceuticals. Raising awareness could be effective, as pharmacies are generally more trusted than unknown marketplace sellers.
- **Price sensitivity** - the customer may perceive the pharmacy as expensive. In this case, an SMS about promotions and more affordable prices could improve engagement.
- **Relocation** - if the customer has moved, they may have stopped visiting the original pharmacy. Here, it is worth highlighting that the company operates a network of pharmacies, potentially offering a more convenient location.

* Medicines and parapharmaceuticals are specific products, and their promotion requires a more thoughtful approach to convey their value to potential customers. For example, distributing news updates or popular science articles can be effective. This type of communication feels less commercial than traditional marketing content, making it a more appealing way to stay connected with customers and encourage them to visit the website, read product reviews, or request additional information.

2. Three groups of potential “warm” customers

Potential customers (18.15%)

This customer segment includes financially stable customers who shop at the pharmacy irregularly. The key objective is to increase their purchase frequency. Recommended actions include providing regular reminders about the company’s products, announcing new arrivals, and sharing timely information on promotions. Emphasizing deadlines - such as bonus expirations or limited-time offers - can further motivate repeat visits.

Low-potential customers (16.05%)

This customer segment consists of customers whose spending is lower compared to “Potential customers”. The key objective is to increase both purchase frequency and mean transaction value. Effective strategies include multi-buy promotions (e.g., 'buy two, get the third free'), price discounts, and bonus points on selected product categories.

Lost high-value customers (1.45%)

This customer segment includes customers who have not made a purchase for an extended period but typically spend large amounts when they do. The key objective is to reactivate them and restore their loyalty through tailored, high-impact promotions. Despite its small size, this group deserves focused attention, as it has the potential to transition into the Super VIP category and generate substantial profits. Moreover, reactivation initiatives are relatively low-cost due to the low number of customers involved.

3. Loyal customers

Super VIP customers (9.01%)

These are the company's most valuable customers, contributing a significant share of revenue and serving as a strong source of word-of-mouth promotion.

To strengthen their loyalty, the company should provide tailored experiences - such as personalized offers, access to products they frequently purchase, and exclusive rewards (e.g., a VIP membership card). Creating moments that exceed their expectations can encourage them to share their positive experiences, which in turn increases the likelihood of attracting new customers and accelerates conversion rates.

Regular loyal customers (11.23%)

This customer segment consists of established loyal customers who already generate the pharmacy chain's core revenue and present opportunities for additional growth.

The key objective in working with this group is to increase their mean purchase value by introducing them to complementary or new products. Effective strategies may include personalized product recommendations, encouraging engagement through social media, and involving them in customer surveys to strengthen their sense of participation.

Communication with this segment should be carefully managed. Frequent mailings should be avoided to prevent perceptions of spam, as these customers already visit regularly and are familiar with the brand.

A referral program may also be effective for this group - for example, offering discounts to customers who bring in new clients, as well as extending benefits to their friends and relatives.

Conducting surveys provides two key opportunities:

- Demonstrating to customers that their opinions are valued.

- Gaining actionable insights into customer preferences and behaviors, which can help refine the definition of the target audience and design promotions with a higher probability of conversion.

Implementing improvements based on survey results is equally important. If customers identify pain points - such as shopping inconveniences, product range gaps, or dissatisfaction with pharmacy staff - these issues should be addressed. Doing so not only enhances the customer experience but can also drive broader strategic outcomes: increased visit frequency and higher average transaction values.