

Saint Petersburg State University
Graduate School of Management
Master of Business Analytics and Big Data

**DATA-DRIVEN DEVELOPMENT OF LOYALTY PROGRAM FOR
RETAIL COMPANY**

Consulting project

Master's Thesis by the 2nd year student
BM.5783.2020 - Master's in Business Analytics and Big Data

Zavadskaya Ekaterina

Research Advisor:

Zhukova Sofia

Professor, Information Technologies in Management Department

Saint Petersburg

2022

**ЗАЯВЛЕНИЕ О САМОСТОЯТЕЛЬНОМ ХАРАКТЕРЕ ВЫПОЛНЕНИЯ
ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЫ**

Я, Завадская Екатерина Владимировна, студент второго курса магистратуры направления «Менеджмент», заявляю, что в моей магистерской диссертации на тему «Разработка программы лояльности для розничной компании на основе данных», представленной в службу обеспечения программ магистратуры для последующей передачи в государственную аттестационную комиссию для публичной защиты, не содержится элементов плагиата.

Все прямые заимствования из печатных и электронных источников, а также из защищенных ранее выпускных квалификационных работ, кандидатских и докторских диссертаций имеют соответствующие ссылки.

Мне известно содержание п. 9.7.1 Правил обучения по основным образовательным программам высшего и среднего профессионального образования в СПбГУ о том, что «ВКР выполняется индивидуально каждым студентом под руководством назначенного ему научного руководителя», и п. 51 Устава федерального государственного бюджетного образовательного учреждения высшего образования «Санкт-Петербургский государственный университет» о том, что «студент подлежит отчислению из Санкт-Петербургского университета за представление курсовой или выпускной квалификационной работы, выполненной другим лицом (лицами)».

_____ (Подпись студента)
_____ 31.05.2022 (Дата)

**STATEMENT ABOUT THE INDEPENDENT CHARACTER OF
THE MASTER THESIS**

I, Zavadskaya Ekaterina, (second) year master student, program «Management», state that my master thesis on the topic «Data-Driven Development of Loyalty Program for Retail Company», which is presented to the Master Office to be submitted to the Official Defense Committee for the public defense, does not contain any elements of plagiarism.

All direct borrowings from printed and electronic sources, as well as from master theses, PhD and doctorate theses which were defended earlier, have appropriate references.

I am aware that according to paragraph 9.7.1. of Guidelines for instruction in major curriculum programs of higher and secondary professional education at St.Petersburg University «A master thesis must be completed by each of the degree candidates individually under the supervision of his or her advisor», and according to paragraph 51 of Charter of the Federal State Institution of Higher Education Saint-Petersburg State University «a student can be expelled from St.Petersburg University for submitting of the course or graduation qualification work developed by other person (persons)».

_____ (Student's signature)
_____ 31.05.2022 (Date)

АННОТАЦИЯ

Автор	Завадская Екатерина
Название магистерской диссертации	Разработка программы лояльности для розничной компании на основе данных
Факультет	Высшая школа менеджмента
Направление подготовки	Бизнес-аналитика и большие данные
Год	2022
Научный руководитель	Жукова София
Описание цели, задач, основных результатов	<p>Товары, представленные на рынке, как правило, схожи по ассортименту и ценам, что делает конкуренцию между розничными торговцами очень высокой. Существуют различные программы лояльности, призванные привлечь клиентов к конкретной розничной компании. Целью данного исследования является сегментация потребителей на основе данных за период с 2020 по 2021 год с целью обновления существующей программы лояльности. С помощью ИТ-инструментов были подготовлены данные, полученные от компании, и проведен дальнейший RFM анализ. Были выявлены существующие проблемы и ключевые изменения в поведении клиентов. Полученные результаты, такие как распределение клиентов на семь групп, выявление новых пользователей, а также основные тенденции в изменении поведения клиентов, были использованы для выработки рекомендаций по улучшению существующей программы лояльности. Рекомендации включают разработку многоуровневой бонусной программы, которая сделает обслуживание клиентов более персонализированным.</p>
Ключевые слова	программа лояльности, сегментация клиентов, RFM анализ

ABSTRACT

Master Student Name	Zavadskaya Ekaterina
Master Thesis Title	Data-Driven Development of Loyalty Program for Retail Company
Title Faculty	Graduate School of Management
Main field of study	Business Analytics and Big Data
Year	2022
Academic Advisor's Name	Zhukova Sofia
Description of the goals, tasks, and main results	<p>Goods presented on the market, generally are similar in the assortment and prices, that makes the competition between retailers very challenging. Different loyalty programmes exist to attract customers to give a favour to a specific retailer. The goal of this research is to segment consumers based on data for the period from 2020 to 2021 in order to update the existing loyalty programme. With the help of IT instruments, the data received from the company were prepared and further RFM analysis was carried out. Existing problems and key changes in customer behaviour were identified. The results obtained, such as the distribution of customers into seven groups, the identification of new users, as well as the main trends in changing customer behaviour were used to offer recommendations for improving the existing loyalty programme. Recommendations include the development of a multi-level bonus program, which will make the customer experience more personalised.</p>
Keywords	loyalty programme, data segmentation, RFM-analysis

Table of content

List of Figures.....	6
List of Tables	7
Introduction	8
Chapter 1. Literature review	11
1.1. Overview.....	11
1.2 Theoretical foundations of the development of a loyalty programme.....	11
1.3 Segmentation as an approach to improve loyalty programmes	16
1.4 Case of supermarket chain	20
1.4.1 The Company profile	20
1.4.2 Problem statement and business goal.....	20
1.5 Summary of Chapter 2.....	21
Chapter 2 Research methodology.....	23
2.1 Overview.....	23
2.2 IT instruments for data processing.....	23
2.3 Data description and exploration	23
2.3.1. Data specification and description	23
2.3.2 Data problems	24
2.4 Data preparation.....	24
2.5 Practices of RFM analysis application.....	27
2.6 Description of the analysis methodology.....	29
2.7 Conclusions on chapter 3	33
Chapter 3. Loyalty program improvement	34
3.1 Overview.....	34
3.2. Identification of problems in customer behaviour based on segmentation.....	34
3.2.1 Results for City A.....	34
3.2.2 Results for City B.....	38
3.2.3 Results for City C.....	41

3.2.4 Results for City D.....	44
3.2.5 Comparison of segmentation results for all cities	47
3.3 Development of a loyalty programme based on the received customer segmentation	49
3.4 Conclusions on chapter 3	54
Conclusion	55
References	56

List of Figures

Figure 1. Research framework.....	9
Figure 1. The classification of loyalty programmes	13
Figure 3. Visualisation of outliers	25
Figure 4. Box plot of customers purchases.....	26
Figure 5. Stages of RFM analysis.....	30
Figure 6. The average monthly sum of purchases in city A, rub.....	34
Figure 7. Dynamics of the segment share in City A.....	35
Figure 8. Dynamics of RFM indicators in City A	36
Figure 9. Dynamics of the main segments and groups of the city A.....	37
Figure 10. The shares of groups that entered the outflow in 2021 in the city A	37
Figure 11. The average monthly sum of purchases in city A, rub.....	38
Figure 12. Dynamics of the segment share in City B.....	38
Figure 13. Dynamics of RFM indicators in City B	40
Figure 14. Dynamics of the main segments and groups of the city B	40
Figure 15. The shares of groups that entered the outflow in 2021 in the city B	41
Figure 16. Average monthly sum of purchases in city C, rub	41
Figure 17. Dynamics of the segment share in City C.....	42
Figure 18. Dynamics of RFM indicators in City C	43
Figure 19. Dynamics of the main segments and groups of the city C.....	43
Figure 20. The shares of groups that entered the outflow in 2021 in the city C	44
Figure 21. Average monthly sum of purchases in city D, rub.....	44
Figure 22. Dynamics of the segment share in City D.....	45
Figure 23. Dynamics of RFM indicators in City D	46
Figure 24. Dynamics of the main segments and groups of the city D.....	46
Figure 25. The shares of groups that entered the outflow in 2021 in the city D	47
Figure 26. Dynamics of the average monthly cost of purchases in all cities.....	47
Figure 27. Dynamics of segments for all cities	48
Figure 28. The process of accrual of points under the proposed loyalty program	52

List of Tables

Table 1. Customer loyalty levels	12
Table 2. Comparison of main types of loyalty programmes.....	14
Table 3. Comparison of customers segmentation methods	19
Table 4. Characteristics of cities.....	20
Table 5. IT instruments	23
Table 6. Data specification	23
Table 7. Description of prepared data.....	26
Table 8. Number of unique users.....	30
Table 9. Comparison of the main statistical indicators.....	31
Table 10. Segmentation Model Parameters	32
Table 11. Groups of consumer segments.....	32
Table 12. Average purchase amount in City A.....	35
Table 13. Average monthly purchase amount in City B.....	39
Table 14. Average purchase amount in City C.....	42
Table 15. Average monthly purchase amount in City D	45
Table 16. Average monthly purchase amount in all cities.....	48
Table 17. Characteristics of main customer groups.....	49
Table 18. Comparison of customer groups	50
Table 19. Comparison of proposed types of loyalty programs.....	51
Table 20. Conditions of the proposed loyalty program	52

Introduction

Nowadays, many supermarket chains with a similar assortment and approximately the same prices are present on the market. Therefore, customers with such a wide variety can afford to choose where to shop and in what quantity. Methods such as special offers and loyalty programmes are generally used to attract customers.

Loyalty programmes are designed to build a long-term relationship with the customer to highlight a particular supermarket chain among others and motivate customers to increase the number of visits and the average check-up. The loyalty programme can help to:

- retain customers who bring the greatest profit,
- increase the average check-up and the frequency of purchases
- reduce the cost for regular customers
- establish constant communication with customers
- collect a high-quality database
- get valuable information about customer behaviour
- increase brand credibility and awareness
- monitor changes in consumer behaviour and effectively influence it
- optimise marketing expenses through personal offers.

Most well-known retailers have their loyalty programmes, generally similar to each other; hence, it is necessary to improve the user experience by personalising the offers. To implement a more personalised loyalty programme, collecting as much information as possible about customers, their behaviour, and habits is vital. Unfortunately, the information supermarkets can collect about the participants of loyalty programmes is often quite limited, mainly about the number and the date of purchases. Nevertheless, based on such data, it is possible to segment customers using RFM analysis that allows to successfully divide customers into groups and, depending on the results, determine which groups of customers can be stimulated and how.

This research was conducted for a Russian regional supermarket chain that also has a loyalty programme, which, however, does not correspond to current trends in this industry and does not show high efficiency. The research has a practical value offering recommendations on improving the existing loyalty programme.

The main research goal of this project is to analyse customer data for the period from 2020 to 2021 provided by the Company and perform segmentation of consumers into

different groups. This allows to visualise customers' behaviour and to improve the existing loyalty programme strengthening the Company's position in the competitive market.

To achieve the goal, the following objectives are defined:

- Data preparation for further analysis
- Conducting customer segmentation and identifying key changes in customer behaviour
- Data visualisation of results
- Problem identification and loyalty programme development.

In pursuing these tasks, various tools of data analysis were employed, including a JupyterHub environment with Python libraries, Microsoft Excel and business intelligence software (Microsoft Power BI). Work with data is organised based on CRISP-DM methodology. The present work includes both theoretical aspects of studying the development of loyalty programmes and the CRISP-DM methodology for data processing and analysis. The research methodology is shown in **Figure 1**.

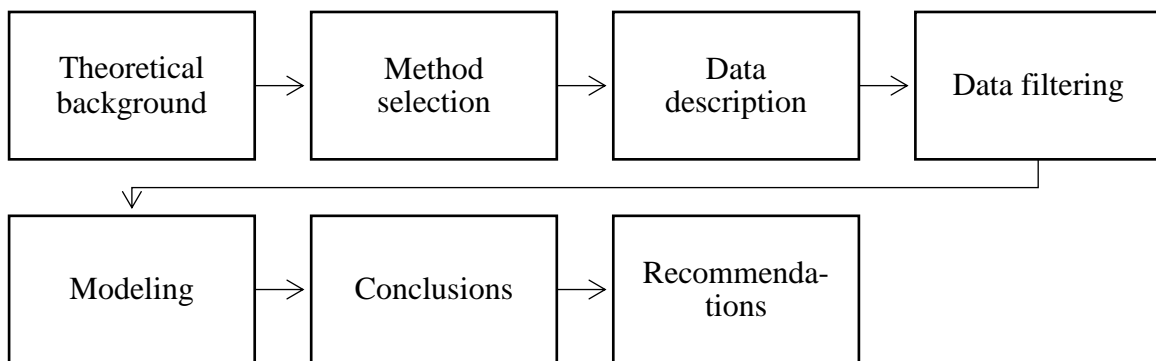


Figure 1. Research framework

Source: [author research]

As a result of these stages, it is planned to obtain the following results:

- Determination of the number of loyal consumers
- Dynamics of changes in customer segmentation
- Identification of customers who may fall into the outflow.

The second chapter describes the theoretical foundations and features of developing loyalty programmes, provides various types of loyalty programmes and justifies the choice of a multi-level bonus programme as the basis of recommendations. It also explains the chosen method of RFM analysis for further segmentation and describes the company providing the data.

In the third chapter, the selected method of RFM analysis is considered in detail, data are described and prepared, and criteria for the application of the method are developed.

The fourth chapter examines visualised results for each city, where the analysed supermarket chain is located. Furthermore, based on the results of segmentation, a new loyalty programme is proposed with the possible results of its implementation described.

The study is concluded with a summary of findings and results with recommendations for future work.

Chapter 1. Literature review

1.1. Overview

The chapter gives different concepts of the loyalty programme and its classification. It also describes the company for which the analysis will be carried out. By comparing different types of loyalty programmes, the choice of a certain type for updating the company's loyalty programme as well as the choice of segmentation and method for further analysis is justified.

1.2 Theoretical foundations of the development of a loyalty programme

Currently, in the academic literature there is no universal interpretation of the concept of a «customer loyalty programme» due to the relative novelty of the topic of loyalty as such. Nevertheless, it is possible to identify several approaches to developing the customer loyalty programme.

According to Zeithaml et al., customer loyalty can be understood as «favourable behaviour towards a company, evidenced through a customer's likeliness to do repeat business with a given retailer, a preference towards a certain brand and word-of-mouth advocacy» (Zeithaml et al. 1996)

Businesses continue to utilise loyalty programmes to influence consumers' purchase behaviour over time (Breugelmans et al. 2015), enhance loyalty by giving incentives (Gorlier & Michel 2020), and sustain a competitive advantage (Nastasoiu & Vandenbosch 2019).

Historically, the objectives of the customer loyalty programme have often been reduced solely to stimulating sales by encouraging an increase in the frequency of visits to the store and the volume of purchases, thus reducing the loyalty programme to encourage the volume of demand. (Bombajj et al. 2020)

Usually, the main attribute of a loyalty programme is a loyal customer card. Regular customer cards, whether they are physical or electronic, allow retailers to receive information about the customer, his buying behaviour, preferences, and shopping repertoire. Based on such unique customer knowledge, it is possible to conduct a quantitative and then qualitative analysis of the customer profile and generate targeted offers for each customer segment. (Chernysheva and Yakubova 2019)

The characteristics of customer loyalty are defined by researchers as follows:

- the consumer prefers this company or brand to other competitors on a regular basis
- the consumer regularly makes repeated purchases from this company and is not changing it
- the customer feels a sense of satisfaction from his relationship with the company and the brand

- the actions of competitors do not have an impact on him, or this influence is minimal
- loyalty is primarily explained not so much by the rational as by the emotional component.

It is possible to distinguish several levels of loyalty, which are presented in Table 1.

Table 1. Customer loyalty levels

<u>Name of the level</u>	<u>Characteristic</u>
Potential buyer	This type of customer does not realise the need for a specific company or brand product. The main task is to convince such a consumer to purchase a product or use services through advertising and promotions
Random (new) customer	This type can make a one-time purchase of goods. For this type of customer to become loyal, it is necessary to provide a high level of service, the quality of the product and the convenience of making a purchase.
Customer	Such customers regularly purchase goods or services from the company. The main goal when working with this group is to make a long-term relationship with the customer, for this, it is necessary to create additional emotional value received by the customer from each purchase.
A loyal customer	This type of customer often purchases a product of this brand. Therefore, it is important to identify the needs of such customers and personalise their experience.
An adherent customer	Such customers are the most loyal and are ready to forgive some of the shortcomings of their beloved company if they are temporary, but if problems from an accidental misunderstanding develop into permanent, even the most loyal buyer is likely to switch to competitors' products.

Source: [Chernysheva and Yakubova 2019]

There are different types of loyalty depending on different attributes. Butcher (Butcher 2004) offers two classifications of loyalty programmes. First, he divides all programmes into closed and open ones. Closed programmes assume some conditions for entry – membership fees and filling out entrance questionnaires. Open programmes do not contain conditions, «but as a result, they often involve many clients who do not bring the company profit» Second, depending on the nature of the target groups and orientation, there are programmes designed for end-users, entrepreneurs, and distributors.

Company «RBC Capital Markets» offers the following classification of loyalty programmes, which is shown in Figure 2

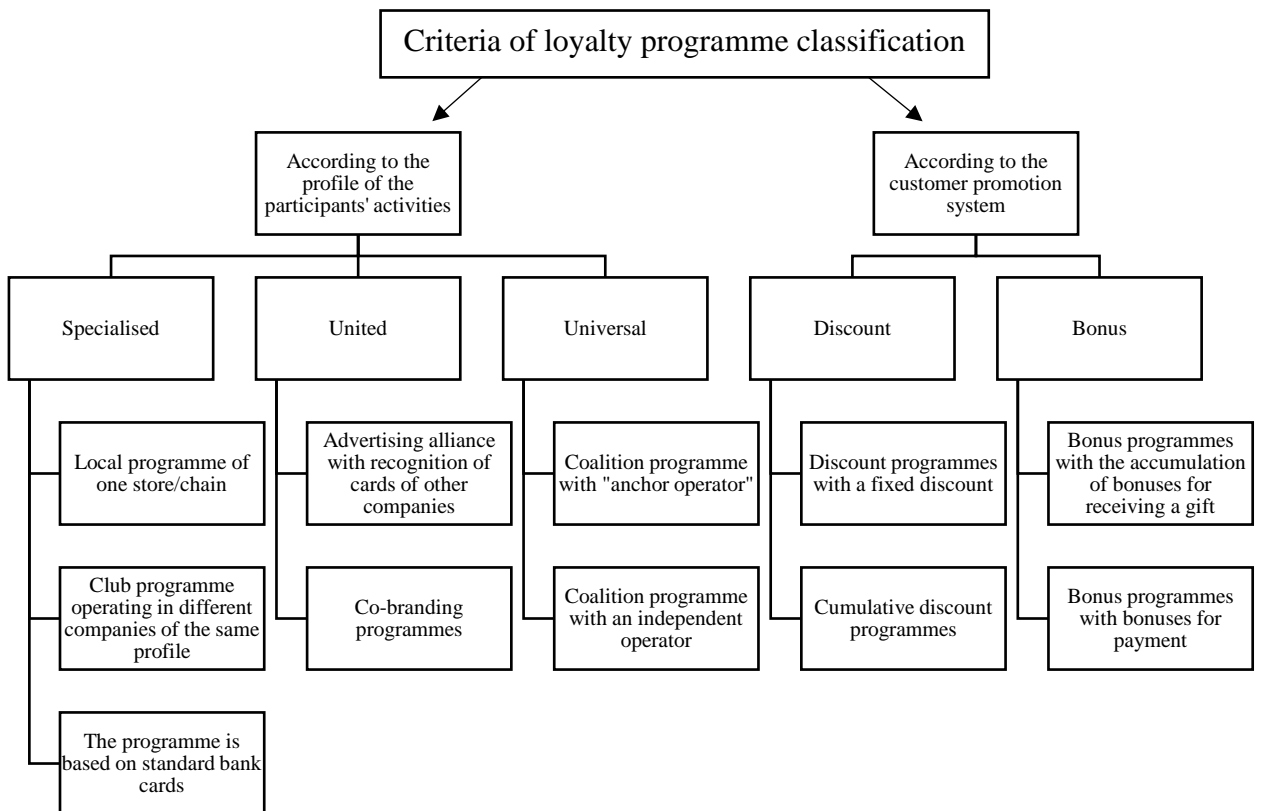


Figure 2. The classification of loyalty programmes

Source: [«RBC Capital Markets»]

In addition, there are types of loyalty programmes such as intangible loyalty programmes. This is the provision of gifts and/or services when certain conditions are met, or the accrual of bonuses with the opportunity to exchange them for awards.

As applied to retail, one of the most popular programmes is the discount and bonus programme.

The bonus programme is the most common and simple model in which regular customers are awarded points for purchases, and the next time they contact the company, these points turn into physical benefits: discounts, gifts and additional services.

When implementing such a loyalty system, it is important not to complicate it. Complex conditions for awarding bonuses can confuse not only buyers but also the staff responsible for providing remuneration. Here it is advisable to adhere to the principle: the simpler and clearer, the better.

To decide on choosing a loyalty programme, the main features of bonus and discount types were identified and presented in Table 2.

Table 2. Comparison of main types of loyalty programmes

Type of loyalty programme	Advantages	Disadvantages
Discount	<ul style="list-style-type: none"> - simple and clear conditions - the buyer always knows the size of his discount - it does not require large implementation costs — only for the issue of loyalty cards 	<ul style="list-style-type: none"> - there is no difference from the programme of the neighbouring store - does not motivate the buyer to return to the store and make the next purchase - does not consider the interests and individual needs of the buyer.
Bonus	<ul style="list-style-type: none"> - motivates the buyer to make the next purchase in the same store instead of a competing one - serves as an additional incentive if bonus points burn out after a certain period - the company does not lose revenue in vain, providing a discount to customers who rarely make purchases. 	<ul style="list-style-type: none"> - it is difficult to explain the conditions not only to customers but also to staff - the customer does not know exactly his discount - difficulties with settings, as well as possible errors in the accrual and deduction of points.

Source: [author research]

As can be seen from Table 2, each of the types has its advantages and disadvantages. It is also worth noting that from the consumer point of view, the loyalty programme based on discounts is still the most profitable and comfortable, since the customer benefits in any case. (Bombaij et al. 2020)

At the same time, the bonus accrual system is frequently either quite difficult for the consumer to perceive or unprofitable to use, since very often retailers offer a small percentage of accrual or a high discount rate (for example, 10 points = 1 ruble). In such cases, the accumulation of bonuses becomes very long and unprofitable, therefore, if there

are many competitive shops with a similar product range, offering discounts, the consumer is likely to go there.

At the same time, in recent years, the bonus system has been the most popular with retailers, since unlike the discount system, it is more focused on creating a long-term relationship with the customer. Also, with a well-chosen ratio of bonuses to money and the percentage of bonus accrual, it is possible to increase the number of purchases and the number of visits. (Chaudhuri et al. 2020)

Bonus and discount programmes can also be multi-level. There is no fixed discount for the buyer in this programme. The scheme of its operation is the following:

1. At the first purchase, a loyalty card with a minimum bonus is issued to the customer.
2. The amount of all subsequent purchases are recorded on the card.
3. When a certain threshold is reached, the discount is increased.

This option is suitable for stores where customers do not shop daily, but often enough. In this case, they are interested in increasing the discount.

There are also different advantages and disadvantages of such type. To begin with advantages:

- It is an honest loyalty programme. The customers know exactly what discount they will receive.
- The opportunity to encourage exactly those customers who buy often and a lot, which means they are especially important for the company.
- Motivates customers to make further purchases.

However, there are also different disadvantages, such as:

- Complexity in implementation due to the need for a comprehensive accounting of the accumulated number of purchases.
- Complexity in restoring the previously received discount amount when the card is lost can be very insulting.

When developing and changing a loyalty programme, it is also important to track certain metrics. There are several indicators of the effectiveness of the loyalty programme that requires mandatory tracking:

- Lifetime Value, that is, the average life cycle of a customer in the company. After the implementation of a successful programme, it will grow, as people will stay with the company longer.

- Percentage of user engagement. This metric is quite simple to calculate. It is necessary to take the total number of customers for a certain period (for example, by receipts)

and the number of those who are members of the loyalty programme. It is also worth paying attention to how intensively buyers use the programme. If many cards are held, but the movement of bonuses is slow, the efficiency is not high enough.

- Churn rate or otherwise an indicator of user churn. Unlike Lifetime Value, it should decrease, because customers stay with the company longer.

- Net Promoter Score — this indicator demonstrates how well word of mouth works in the company. The more people are willing to recommend the company to others, the better for the business.

Developing a loyalty programme is a complex process. To effectively implement a loyalty programme, it is important to understand customers' needs and behavioural characteristics.

1.3 Segmentation as an approach to improve loyalty programmes

Modern loyalty programmes usually have a high level of personalisation. The business strives to give customers an individual offer that will be interesting to them and make them pay attention to the goods of the brand. The widely used approach to understand customer interests is to segment them based on their personal information.

Customer segmentation can be performed using a variety of unique customer characteristics to help businesses to customise marketing plans, identify trends, plan product development, advertise campaigns and deliver relevant products. The most common attributes used in customer segmentation are location, age, and previous purchase behaviour. (Christya et al 2021)

Loyalty programmes and especially segmentation of customers is a modern direction of development not only in international but also in Russian retail. More and more retail chains are announcing the launch of customer loyalty programmes, and some of them have already succeeded so much in this that they communicate directly with their customers, providing individual offers.

By implementing loyalty programmes, retail chains are primarily interested in achieving two main goals:

- increasing the flow of customers
- an increase in the average amount of the check-ups.

Achieving these goals strongly depends on a strategically and methodologically correct approach to customer clustering. Relevant customer segmentation enables retail chains to increase the share of loyal customers and prevents a drop in traffic among buyers who are at risk of switching. (Chaudhuri et al. 2020)

The purpose of dividing customers into groups is to increase the efficiency of the company's interaction with the market, consumers, and partners. If the segmentation is carried out correctly, within the same group, customers have similar preferences, requirements for goods tastes and financial capabilities. Working with an ordered audience is simple according to the following features:

- Having identified priority groups of customers, the company can focus its efforts and investments on achieving the fastest possible payback, increasing profits, and popularising the brand. This is achieved by meeting the needs of customers.
- The choice of the company's global marketing concept.
- Knowing consumer preferences, a business can determine current market trends and form business priorities.
- When the company's management knows how many percent of buyers belong to different categories, they can make quantitative calculations regarding the return on investment, profit forecast, and further development of the company.

The main aspect of dividing consumers into categories is the criterion by which segmentation is performed. The choice of the criterion is very important, as it affects the subsequent payback of the loyalty programme. (Kim et al. 2021) The most popular criteria are the following:

- Demographic feature – in this case, customers can be divided into categories according to their place of residence, education, and income level. Although it used to be one of the key factors for determining people's needs, it is no longer significant.
- Geographical feature – the country, region and city of residence also determines customers interests.
- Psychographic sign – this factor is now considered a very effective principle of segmentation of consumers for the motivational system. This group of factors combines lifestyle, personal qualities of character and social status of consumers. Knowing the psychological characteristics of the customers, it is easier to convince them to buy the company's goods, but it is very difficult to get this information.
- Behavioural feature – the defining aspect of this principle of categorisation of buyers – the study of consumer experience. A personalised loyalty programme allows the company to get data on consumer experience. By tracking the purchases, tastes, priorities, and preferences of the customer, the business can form individual offers. Such offers are highly likely to be of interest to the buyer.
- Gender – there are differences in needs and priorities between men and women.

Behavioural segmentation of clients is the most relevant for this study. The main goal of behavioural segmentation is to understand the various buying, consumption, and usage habits that consumers have, which should be used by companies to preserve their brands and products, promote their purchases and encourage them to enter a strong competition in the market.

Behavioural segmentation is dynamic, it is constantly evolving because customer profiles are constantly changing.

The behaviour of each segment needs to be evaluated periodically, as customer behaviour changes due to the influence of factors such as time, location, and economic moment.

In behavioural segmentation, the following criteria are distinguished:

1. The reason for making a purchase. It is also the principle of using the product and is great for behavioural segmentation.

2. Consumer benefits. Each buyer sees his benefits in the purchase, having identified them, customers can be divided into groups and each of them can offer additional products with similar benefits.

3. Client status. All buyers can be divided into groups based on the different groups, such as new customers, regular customers etc. Each of these segments requires a special approach and strategy of attraction with the help of mailings in different directions.

4. Degree of commitment. Some customers prefer a certain brand of a product and will make a choice in its favour regardless of the price. Most often they purchase products of one trademark ignoring others, such consumers are called unconditional adherents. In addition to them, there is a segment of tolerant adherents who prefer several brands. Fickle adherents constantly change their preferences and move from one brand to another.

6. The attitude to the product determines how much time should be spent working with each segment. They are divided into enthusiastic, positive, neutral, negative, and hostile.

7. The degree of readiness for perception, which is also the degree of awareness, the criterion segments customers into people who are informed about the product to varying degrees.

There are various ways to segment customers. (Kim et al. 2021)

1. ABC XYZ method.

This segmentation method is based on the Pareto principle: 20% of customers provide 80% of all profits. It is suitable for all companies that sell more than one product.

The method is carried out in three stages:

The first stage is ABC analysis. It is necessary to allocate "segment A", that is, those 20% of customers who will generate the main profit which is 80%. During segmentation, we will also get group B, which gives 15% of revenue, and group C — 5%.

The second stage is XYZ analysis. It is necessary to find out the stability of demand from different groups from the indicators for the month. To do this, the coefficient of variation is calculated.

Customers who consistently bring revenue to the company — AH, and CZ bring the company the least revenue and their demand is not predictable. Customers from the AY group have a fluctuating demand, but they also bring high income, so they need to pay attention and make special offers for them. Customers from Group B bring good revenue, but their demand is hard to predict. Perhaps they can be shown ads for seasonal goods before the holidays.

2. RFM analysis

The base is segmented by customer activity. It is necessary to divide it into groups to find out how often customers buy and how much they spend. There are three parameters:

- time since the last purchase — “recency”
- number of purchases — “frequency”
- the number of purchases is “monetary”.

3. LTV segmentation

The LTV method answers the question of how much profit a company can get from one client. To calculate LTV, the average customer receipt, the number of purchases per month and the time during which they continue to buy are considered. The comparison of these methods is presented in Table 3.

Table 3. Comparison of customers segmentation methods

Method	Advantages	Disadvantages
ABC XYZ	- Simplicity - Only data about the sum of purchase is required	- Requires to delete some groups of customers (who have made one-time or large purchases, new customers)
Lifetime Value	- Focusing on a long-term customer relationship that allows the company to focus on retention rather than attraction	- Complexity of sales volume forecast

RFM segmentation	- Effective optimisation of work with segments	- Focusing on past customer actions
------------------	--	-------------------------------------

Source: [author research]

Based on the table, the RFM analysis is the most convenient and comprehensive. Therefore, the RFM analysis is employed for customer segmentation.

1.4 Case of supermarket chain

1.4.1 The Company profile

Consulting project is developed for the Russian regional supermarket chain, which is the leader of retail trade in the region and has 80 stores in six cities, offering a wide selection of high-quality food and non-food products.

In addition to the analysed supermarket chain, there are also other local supermarkets, federal supermarket chains such as «Magnit», «Okey», as well as foreign supermarkets «Metro Cash & Carry» and «SPAR».

In 2021, X5 Retail Group N.V. entered the market of the region and opened 8 supermarkets «Pyaterochka» in 4 cities. The chain also plans to develop in the region and open more supermarkets, that may lead to increased competition in the market and an outflow of customers from other supermarkets.

Based on increased competition, to continue to retain customers in the analysed supermarket chain, it is important to identify current problems in customer behaviour.

To attract customers and increase loyalty, this supermarket chain utilises methods such as discounts, as well as a loyalty programme.

The loyalty programme consists of accruing bonuses in the amount of 1% of the cost of all purchases. It is possible to pay for the following purchases with these bonuses, but there are certain restrictions in the programme. Bonuses are not awarded for certain groups of goods, for example, for own-brand goods and discounted goods. The share of users participating in the programme in revenue is more than 70%, therefore, the behaviour of the programme participants reflects the trend of the entire supermarket chain.

1.4.2 Problem statement and business goal

Current trends in the development of federal supermarket chains can have a negative impact on customer behaviour. Within the framework of the project, there is also data for four cities in the region. These cities have different characteristics, which are presented in Table 4.

Table 4. Characteristics of cities

City	Characteristic
A	- Location near the regional centre - Industrial satellite of the regional centre - Many federal and regional supermarket chains
B	- Company town - Location away from other cities - A large number of federal and regional supermarket chains
C	- Regional Centre - A large number of federal and regional supermarket chains
D	- A small satellite town of the regional centre - A small number of supermarket chains

Source: [author research]

The main problem of this supermarket chain is the presence in all cities of the following negative trends:

- Reducing the average sum of purchases.
- Reducing the number of purchases.

Based on this, it is important to determine exactly how the behaviour of customers has changed, as well as to suggest ways to improve the current loyalty programme.

To sum up, the business goal is to identify problems and critical changes in customer behaviour of loyalty programme participants.

1.5 Summary of Chapter 2

Analysis of the literature has shown that it is convenient to use segmentation to identify the characteristics of customer behaviour to track the dynamics of the transition of consumers from one segment to another.

Summarising the above, the development of loyalty programmes is an urgent problem, both at the beginning of the century and nowadays. Literature review shows that a lot of work has been devoted to the study of the types and effectiveness of methods of stimulating loyalty programmes.

To develop loyalty in retail, the use of the bonus programme is more effective. Moreover, one of the most effective ways to measure changes and behavioural patterns is customer segmentation. Segmentation also helps in developing a marketing strategy and improving the existing loyalty programme. Among the various methods, RFM analysis stands out, since it helps to segment customers even in the absence of a deep understanding of the customer's characteristics, that is, only by the number of purchases and the prescription

of purchases. Therefore, the bonus programme is used as a basis for developing recommendations, and segmentation based on RFM analysis is used to identify existing problems, as well as for greater personalisation of the loyalty programme.

Chapter 2 Research methodology

2.1 Overview

The following chapter discusses the methodology utilised in the present work. It describes IT tools with the help of which the analysis was performed. It also describes the array of data received from the company, as well as the features and problems of the data. The chapter discusses RFM analysis in more detail and describes its application for this analysis.

2.2 IT instruments for data processing

To achieve the goal, it is necessary to include both data mining and BI development tasks. Table 5 describes the main IT instruments, that were used to solve these tasks.

Table 5. IT instruments

Resource	Description	Where it was used
Jupyter Hub with Spark environment	Interactive development environment (IDE) for Jupiter notebooks for data mining tasks in Python programming language	- Primary data processing - Analysis of the city D
Microsoft Power BI	Business intelligence and interactive visualization software	- Data visualization
Microsoft Excel	A program for working with spreadsheets	- Building a RFM model - Segmentation of customers across all cities

Source: [author research]

Summing up all the above, the combination of all instruments makes it possible to effectively solve the segmentation task for the further development of the loyalty program and the offer of recommendations.

2.3 Data description and exploration

2.3.1. Data specification and description

The project includes an analysis of supermarket chain data provided by the company. Data included information on purchases of customers participating in the loyalty programme. The general characteristics of the data are presented in Table 6.

Table 6. Data specification

Files	2 Excel files: a) Data on the amount of purchases b) Correspondence of the phone number to the city
-------	---

Size	a) 25 sheets, 2 parameters, 4856459 records (on average, 200 thousand per sheet) b) 1 sheet, 2 parameters, 339057 records
------	--

Source: [author research]

Data with information about purchases presented on 25 sheets for the period from December 2019 to December 2021 by month and contains the following parameters:

1. Customer's phone number
2. Number of purchases
3. Sum of purchases

An additional set of data was also provided with the correspondence of the phone number to a specific city or region. The data is presented for four cities, each of which has its own specifics, which will affect the results of segmentation.

2.3.2 Data problems

In the process of working with data, a data problem was identified, which is important to solve when preparing data for the analysis.

This problem is associated with atypical purchase rates, in particular with unusually high purchase rates. Such indicators are erroneous and are fraudulent transactions, the reasons for which may be as follows:

- Using a number participating in the loyalty programme by a large number of people
- Using purchases of people not participating in the loyalty programme.

Such data must be correctly identified and deleted, as they can lead to erroneous results.

2.4 Data preparation

Data preparation consists of several stages:

- Data filtering
- Data transformation for further analysis.

At the first stage, outliers and other erroneous data are usually detected. In statistics, outliers are values that are quite different from other values. Outliers can indicate anomalies in the data distribution or errors; therefore, outliers are often excluded from the data set.

For the primary determination of outliers, it is possible to use graphical method, particularly, dot plots and boxplots.

One of the simplest methods for determining emissions is to create a dot plot. The figure below shows a dot plot for the values of the purchases sum and the number of

purchases (Figure 3). The graph clearly shows abnormal values that can affect the averages and interfere with segmentation.

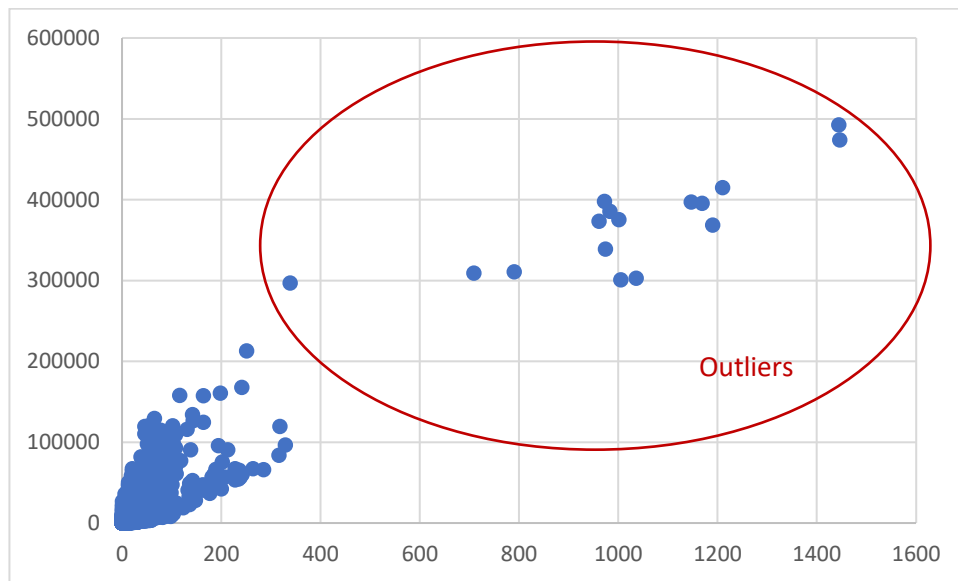


Figure 3. Visualisation of outliers

Source: [author research]

Nevertheless, as can be seen from the graph, there are quite a lot of such values, so it is difficult and impractical to manually remove all unnecessary.

One of the most popular automatic methods for determining emissions is determination of emissions based on the interquartile range (IQR) (Han 2012).

This method was chosen because it has the advantage of not affecting the values of extreme outliers. It is since quartiles are defined values that divide the aggregate into 4 equal parts. Each of the four parts contains 25% of the data. To determine these parts, the data must be arranged in ascending order. Next, the interquartile range (IQR) is calculated as the difference between the third and first quartile, that is, between 25% (Q1) and 75% (Q3) of the data.

Based on this, the following values were determined for the data:

$$\text{IQR} = \text{Q3} - \text{Q1} = 5902 - 1202 = 4700$$

This value can be interpreted so that the average 50% of the values in the dataset have a spread of 4700 rubles.

This method was chosen because it has the advantage of not affecting the values of extreme outliers. That is, in this way, emissions will be determined as follows:

$$\text{Upper internal boundaries} = \text{Q3} + \text{IQR} * 1,5 = 5902 + 4700 * 1,5 = 12\ 952$$

$$\text{Upper outer boundaries} = \text{Q3} + \text{IQR} * 3 = 5902 + 4700 * 3 = 20\ 002$$

Usually, the value 1,5 is used to determine the boundary. This value is a constant of the Tukey method. However, for the outer limits to determine extreme outliers, it is also

possible to increase the indicator to 3. In the case of this study, indicator 3 is just used to exclude the removal of important indicators for the case.

Further, all values exceeding the upper outer limits were removed. IQR identified about 30,000 such anomalous values, which is about 1% of the total amount of data. Therefore, the number of outliers does not significantly affect the total amount of data, but significantly affects other indicators.

Visually, this method can be represented by a box plot shown in Figure 4.

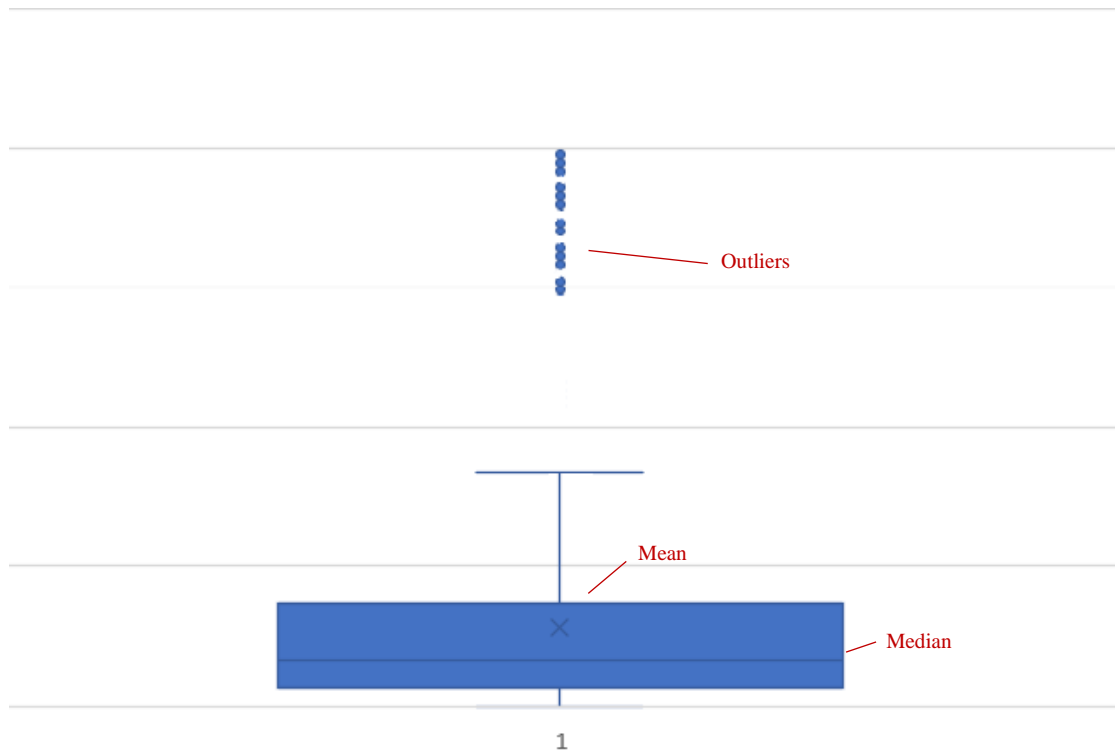


Figure 4. Box plot of customers purchases

Source: [author research]

After clearing the data, it was decided to divide all the data into two years and four cities. It was done due to the need to check the specifics of consumer behaviour depending on the city, since cities differ in size, the presence of competitors and the level of wages. Based on this, the following data sets are shown in Table 7.

Table 7. Description of prepared data

City	Year	Number of rows
City A	2020	287 255
	2021	258 691
City B	2020	251 331
	2021	217 589
City C	2020	244 016

	2021	230 470
City D	2020	1 617 055
	2021	1 448 513

Source: [author research]

An interesting fact is that the smallest city has the largest number of records, which is explained by lower competition relative to other cities and the high popularity of the analysed supermarkets.

2.5 Practices of RFM analysis application

RFM (Recency Frequency Monetary) model is very renowned marketing tool that uses transaction history to define behaviour-based customer segmentation. (Martinez R. G. et al 2019). The RFM model was proposed by Hughes in 1994 year.

Despite the long-term appearance, this model has not lost its relevance, because marketers know that instead of covering 100% of the audience, it is more effective to focus on specific groups of customers who are the most profitable. The RFM model becomes an excellent tool for launching direct marketing actions to retain, activate and return customers. This model is also used in determining the strategy for creating engagement and loyalty in the customer base. (Cheng 2009)

The work on the analysis is carried out in three stages:

Stage 1. Collecting information. First of all, it is necessary to determine the reporting period. As a rule, this is one calendar year. To fill in the table, the following data needed:

- Full name of the customer (it can be contact phone number or email address);
- dates of purchases or other interactions with the company;
- total number of orders during the specified reporting period;
- the sum of all purchases.

Stage 2. Grouping and evaluation. The essence of RFM analysis is the segmentation of customers depending on the prescription of the purchase, the frequency and the amount of money spent. In the classical version, 3 groups are allocated in each evaluation criteria. Then each group is assigned a numerical designation from 1 to 3.

By prescription of the order (Recency):

- 3 — recent customers
- 2 — relatively recent customers;
- 1 — long-time customers.

By frequency of purchases (Frequency):

- 3 — buy often;

2 — buy infrequently;

1 — single purchases.

By the amount of purchases (Monetary):

3 — a large amount;

2 — average amount;

1 — a small amount.

The ranges for values 1, 2 and 3 can be set independently. Each customer receives a score from 1 to 3 according to all criteria, depending on the activity. For example, the user 333 bought once a very long time ago and for a small amount. And customer 111 buys often, spends a lot of money, and the last purchase was a day ago. The result may be a 3 x 9 matrix. There will be up to 27 groups with designations 121, 231, 311.

Stage 3. RFM evaluation and marketing strategy selection. Personal offers and recommendations are generated for each segment to increase loyalty and sales level. Over time, customers can move from one segment to another.

Like any other marketing analytics tool of the company, RFM analysis has a number of pros and cons. (Coussement et al 2014)

Advantages:

- increase the loyalty of the target audience and reduce the cost of advertising campaigns due to accurate targeting;
- simplicity and compatibility with other tools for working with consumers;
- minimum time for analysis. To group the database even for 100 thousand contacts, it will take about half an hour — it is enough to download the information from the CRM;
- the possibility of use in small, medium and large businesses in the fields of HoReCa, e-comm, offline retail and others;
- the ability to profile each RFM analysis group to more accurately determine the target audience;
- targeting improvement — the ability to find a new audience similar to the RFM analysis group by using look-alike technology or other targeting.

Disadvantages:

- the effectiveness is affected by the volume of the customer base (to get an accurate result, it is desirable to collect at least 10,000 contacts per year);
- not suitable for companies with one-time sales;
- the customer base is updated, so you need to analyse at least 1 time a year.

Recency (prescription of purchase) is one of the important criteria in RFM analysis. If a customer does not perform a target action for a long time (does not buy), we can talk about a high risk of his transition to competitors. By offering the buyer some benefits, it is possible to achieve an increase in the frequency of purchases and their average receipt, but to return the lost customer is not an easy task. And the earlier the store sends him a reminder about himself, the higher the probability of repeated treatment. (Martinez R. G. et al 2019).

Experienced business owners recommend performing RFM analysis several times a season. This allows to increase the number of repeated requests and control the inflow/outflow of customers.

2.6 Description of the analysis methodology

As mentioned earlier, the RFM analysis was chosen as a method of customer segmentation, since it corresponds to the research goals and the data provided.

Since the data for two years were obtained, it was decided to conduct an RFM analysis for 2020 and 2021 to be able to compare the results and identify the changes that have occurred, specifically, how the distribution of customers by segment has changed, which groups of customers left the most, which groups of customers stopped going to supermarkets in 2021.

A single algorithm was used to analyse all the cities of each year to compare the results. So, the analysis can be divided into the following steps: (Figure 5)

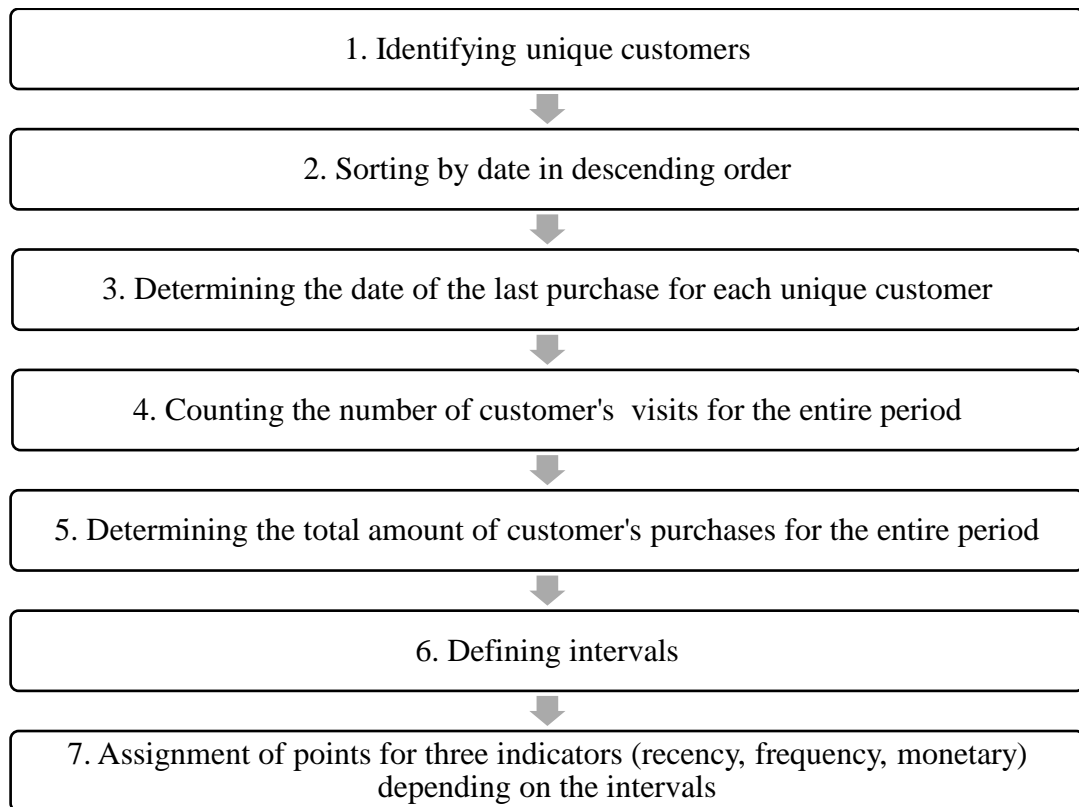


Figure 5. Stages of RFM analysis

Source: [author research]

At the first stage, the number of unique users is determined, according to which the sum of purchases for the entire year is determined. These values could be determined using pivot tables in Excel, as well as using the Pandas library in Python. Based on this, there turns out the following number of unique users by cities. (Table 8)

Table 8. Number of unique users

City	Number of unique users
City A	36 502
City B	31 139
City C	26 147
City D	169 452
Sum	263 240

Source: [author research]

The most important problem of the RFM analysis is to determine the criteria for dividing into segments. The intervals may depend on the number of ratings (3, 4, 5), as well as largely on the specifics of the company's activities. For example, for some companies, a good result will be a purchase once every 4 months, and for some once a week. Nevertheless,

for segmentation can be often used statistical indicators such as median, average, maximum and minimum values.

For this analysis, a rating system from 1 to 5 was used, where 1 is the worst result, and 5 is the best.

To determine the limitation, the maximum number of days was found, and then values were obtained in the amount of 20%, 40%, 60% and 80% of this number. The obtained values became intervals for estimates. In the case of this project, the maximum number of days is 366. Since this project examines the behaviour of customers in the supermarket chain, on the one hand, the indicator of 31 days since the last purchase is usually not very good. However, in the absence of more detailed information, this value is the best, since it means that the customer was in the supermarket last month.

Frequency was defined as the number of months when a customer visited the supermarket chain stores. Based on it, the maximum number will be 12. Such visitors are evaluated for the maximum score. Consequently, the lowest ratings will be for customers who visit the store from 1 to 3 times a year.

To determine the boundaries for estimating the monetary, there was selected such statistical indicator as the median. The reason for such decision was a small number of buyers with very high purchase sums. According to this, the calculation of the average indicator may give an incorrect estimate. In particular, the comparison of indicators is given in Table 9.

Table 9. Comparison of the main statistical indicators

Statistical indicator	Indicator values
Mean	24168,4
Median	13240,4
Maximum	189383,8

Source: [author research]

In order not to complicate the work with data and not to get too many segments, that have different sizes, not five, but three criteria were chosen for the monetary for the period:

- 1 – the lowest sum of purchases for the whole year
- 3 – the average sum of purchases for the whole year
- 5 – the highest sum of purchases for the whole year

The choice of the amount for a smaller number of groups was due to the problem of the lack of data on the behaviour of the client by the days of the month and the need for a more accurate assessment of the number of visits by the client.

To determine the upper limit of the interval, the maximum number was divided by 5 and rounded up and then it is summed up with the median value.

Summarising all the above, Table 10 was compiled where all the segmentation parameters are listed.

Table 10. Segmentation Model Parameters

	Segment indicator values	Parameters
Recency	5	$R < 73$ days
	4	74 days $< R < 146$ days
	3	147 days $< R < 220$ days
	2	221 days $< R < 293$ days
	1	$R > 294$ days
Frequency	5	> 12 visits
	4	10 – 11 visits
	3	7 – 9 visits
	2	4 – 6 visits
	1	1 – 3 visits
Monetary	5	$M > \text{Median} + 40000$
	3	$\text{Median} < M < \text{Median} + 40000$
	1	$M < \text{Median}$

Source: [author research]

Based on this, a table was compiled with parameters that were used for all cities.

As a result of segmentation, 42 segments were obtained, however, for the convenience of interpretation, it was decided to group these segments, which is presented in Table 11.

Table 11. Groups of consumer segments

Group name	Description of the group	Segment indicator values		
		R	F	M
The best customers	Customers, who come regularly and buy for large sums	5	5	5
Loyal customers	Customers, who also come quite often and buy on an average check	5	4, 5	3, 5
Low check loyal customers	Customers, who also come quite often and buy on an above average check	5	4, 5	1

Potentially loyal customers	Customers, who came recently and made several purchases for an average or above average amount	5	1, 2, 3	3, 5
Ordinary customers	Customers, who sometimes come and mostly make purchases for an average or below average amount (rarely on above average)	3, 4	3, 4	1, 3, 5
At risk	Customers who have been around for a relatively long time, have made several purchases and for average or below average amounts	3, 4	1, 2	1, 3, 5
Lost customers	Customers who have made single purchases for a very long time	1, 2	1,2	1, 3, 5

Source: [author research]

Therefore, now it is possible to get an idea of the behaviour of consumers and their segmentation even without knowing the principles of RFM analysis and how the estimates on the indicators of recency, frequency and monetary were distributed. Nevertheless, within the framework of this study, for a deeper understanding of changes in consumer behaviour, it is also important to consider segments without grouping.

2.7 Conclusions on chapter 3

To solve the business goal and the research goal, there was used a combination of theoretical aspects in the development of loyalty programmes, as well as customer segmentation, and a CRISP–DM methodology for data analysis. After preparing the data for segmentation, the RFM method was applied, which allows to identify the main problems in consumer behaviour. It was also decided to group the indicators into 7 groups for ease of perception of information.

The application of this method to all the cities under study will allow us to determine the following tasks:

- which indicator will change the most in 2021 compared to 2020
- which share of segments will have the largest number of participants, as well as the largest share in revenue
- how the volumes of segments have changed in 2021.

Chapter 3. Loyalty program improvement

3.1 Overview

The chapter visualises the results obtained after customers segmentation. The main problems in customer behaviour for the period 2020 - 2021 are identified both for each city and for the supermarket chain as a whole. Based on the obtained characteristics of the groups, an update of the current loyalty programme is proposed, and the possible effect of the implementation is described.

3.2. Identification of problems in customer behaviour based on segmentation

3.2.1 Results for City A

The city A is the third largest city in the region. This city is located about an hour's drive from the regional centre, so competition from both regional and federal networks is quite high.

The primary analysis of the average amount of purchases in this city is shown in the graph below. (Figure 6)

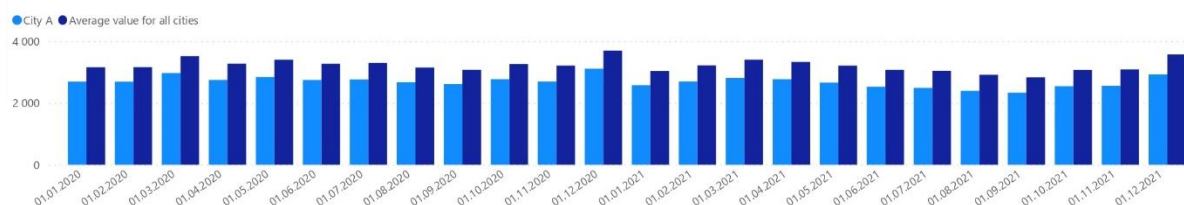


Figure 6. The average monthly sum of purchases in city A, rub

Source: [author research]

The graph shows that the sum of purchases per month for the entire period is less than the average for all cities. There is also a noticeable increase in purchases in December and a decline in summer and September.

As mentioned earlier, for ease of perception, it was decided to group the RFM indicators into groups. A comparison of the values of groups by the number of users is presented below. (Figure 7)

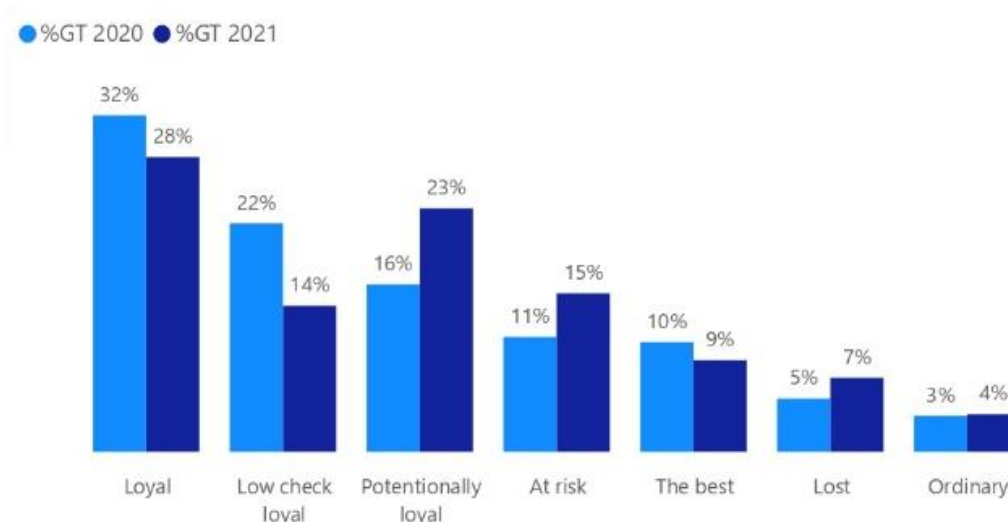


Figure 7. Dynamics of the segment share in City A

Source: [author research]

The graph shows that the main group of loyal consumers, which is one of the most important, decreased by 4%, which is a negative signal. This also confirms the decrease in the number of people in the most important group of customers (The best customers), since on the one hand, the decrease in the volume of loyal customers could be attributed to an increase in the number of the best customers. There have also been significant changes in a group of low check loyal customers that have also moved to a group of potentially loyal customers. On the one hand, the positive aspect is the increase in the number of purchases, but the lack of frequent appearance in the stores of the supermarket chain does not greatly improve the total revenue.

The average amount of purchases that customers of each group spent per year also underwent changes, which is presented in the table. (Table 12)

Table 12. Average purchase amount in City A

Group	Average monthly purchase amount, rub		Share in the total amount, %		Growth rate
	2020	2021	2020	2021	
The best customers	6998,1	6674,1	36,38	34,76	0,95
Loyal customers	2736,1	2469,1	43,71	41,46	0,90
Low check loyal customers	627,9	471,9	6,81	3,92	0,75
Potentially loyal customers	980,8	902,5	7,79	12,52	0,92

Ordinary customers	1401,5	1498,8	2,40	3,19	1,06
At risk	398,2	354,9	2,17	3,20	0,89
Lost customers	292,9	226,1	0,74	0,95	0,77

Source: [author research]

As can be seen from the table, there is a noticeable decrease in the average amount of purchases per year for each group of customers. The only exceptions are ordinary users, but, nevertheless, their impact on the total amount of revenue is insignificant, and the number of participants is also small. Therefore, when developing recommendations, it does not make much sense to allocate special privileges for this group.

To determine more precisely what exactly the most critical changes are, the difference between the RFM indicators in 2021 and 2020 was found. That is, if:

- the value is positive, then there is degradation,
- the value is negative, then there is an improvement,
- 0 means that the value has not changed.

Clearly, the distribution of changes is shown in the graph below. (Figure 8)

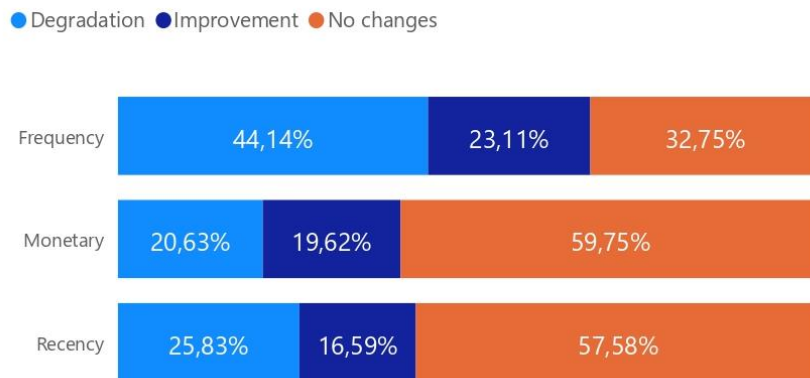


Figure 8. Dynamics of RFM indicators in City A

Source: [author research]

It can be noted that the frequency index has undergone major changes, and for the worse. Even though other indicators have not changed to a greater extent, it is worth noting a tendency to worsen.

Also, to understand the results, not only the analysis of groups was used, but also the 5 largest RFM segments in 2020. The graph below shows the dynamics of people's movements from one segment to another. (Figure 8)

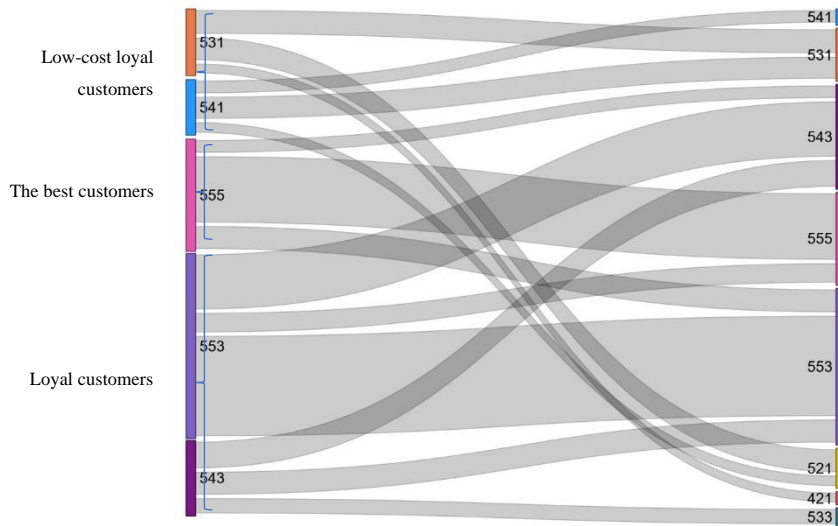


Figure 9. Dynamics of the main segments and groups of the city A

Source: [author research]

The graph shows that mostly customers from the main segments moved to segments with a lower frequency. On the one hand, these transitions occur within the same group, but nevertheless, if such a decline continues, then, for example, even more loyal customers may become either potentially loyal or loyal with a low check.

In 2021, in addition to changing the structure of customer segmentation, some customers stopped visiting supermarkets of this company, that is, they fell into outflow. In City A, among the total number of unique users in two years, namely 32 979 people, 2 770 or 8.6% were in the outflow.

To understand the further strategy, it is important to determine from which groups customers mainly fell into the outflow. This is shown in the pie chart below. (Figure 10)

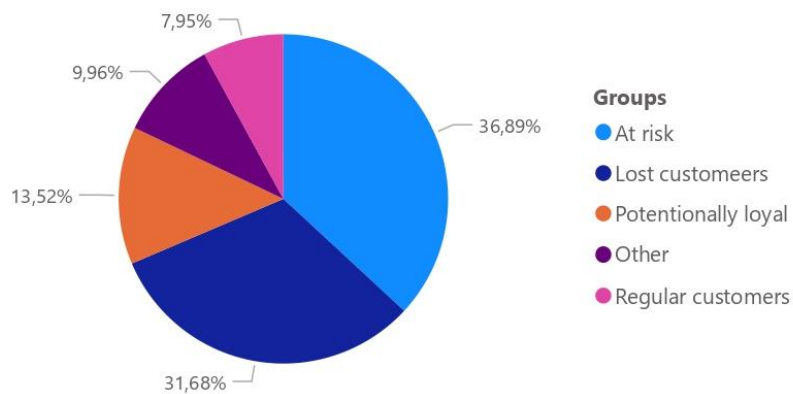


Figure 10. The shares of groups that entered the outflow in 2021 in the city A

Source: [author research]

On the pie chart, the interesting fact is that most of the outflow was not lost customers, but customers at risk. On the pie chart, the interesting fact is that most of the outflow was not lost customers, but customers at risk.

In general, consumers of all segments somehow got into this outflow. That is, even high loyalty in one period does not guarantee 100% that the client will stay in the next. However, since the percentage of the number of groups is not so large, it can be assumed that customers could get into the flow not because of dissatisfaction with the supermarket chain stores, but because of objective reasons, for example, due to moving to another city.

3.2.2 Results for City B

The city B is company town. The size of the city is approximately equal to the city A, but the main feature of the city is its great distance from other cities.

The primary analysis of the average amount of purchases in this city is shown in the graph below. (Figure 11)

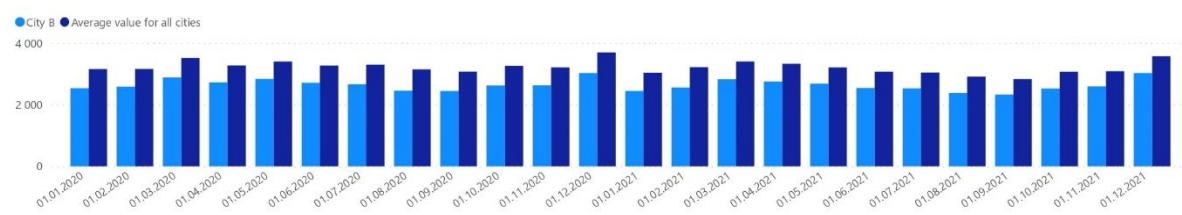


Figure 11. The average monthly sum of purchases in city A, rub

Source: [author research]

As for the average cost of purchases, City B has approximately similar amounts as City A. This is due to their similar size.

The clients of City B were also divided according to the RFM method into segments, and then divided into groups. The result is presented in histogram below. (Figure 12)

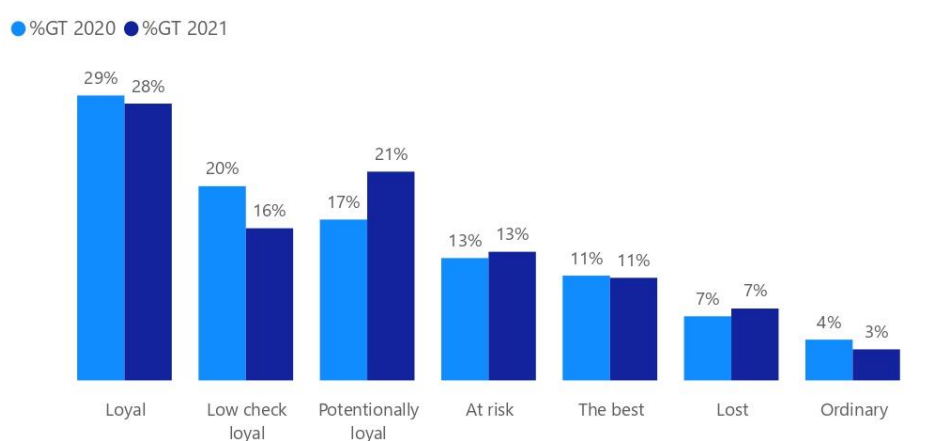


Figure 12. Dynamics of the segment share in City B

Source: [author research]

The graph shows that the greatest changes occurred in the group of loyal consumers with a low check and potentially loyal customers. That is, mostly all the clients of the first group moved to the second. This situation suggests that with a more detailed analysis, the frequency of visits will significantly decrease.

It is also important to look in more detail at the average monthly value of purchases, which is presented in the table below. (Table 13)

Table 13. Average monthly purchase amount in City B

Group	Average monthly purchase amount, rub		Share in the total amount, %		Growth rate
	2020	2021	2020	2021	
The best customers	5598,2	5437,1	31,99	33,55	0,97
Loyal customers	2680,9	2480,6	41,68	41,29	0,92
Low check loyal customers	703,6	577,6	7,46	5,29	0,82
Potentially loyal customers	1189,5	983,1	10,44	12,34	0,83
Ordinary customers	1754,3	1785,6	3,90	3,33	1,01
At risk	534,4	409,9	3,57	3,17	0,77
Lost customers	275,6	237,4	0,96	1,03	0,86

Source: [author research]

Despite the overall significant decrease in the average cost of purchases, which is a negative factor, there are still positive aspects. For example, this is an increase in the average receipt of ordinary users who visit stores quite infrequently. This is a positive trend, since if such users visit stores more often, they could move into a group of potentially loyal customers.

The characteristics of changes for each indicator of the RFM analysis are shown in the graph below. (Figure 13)

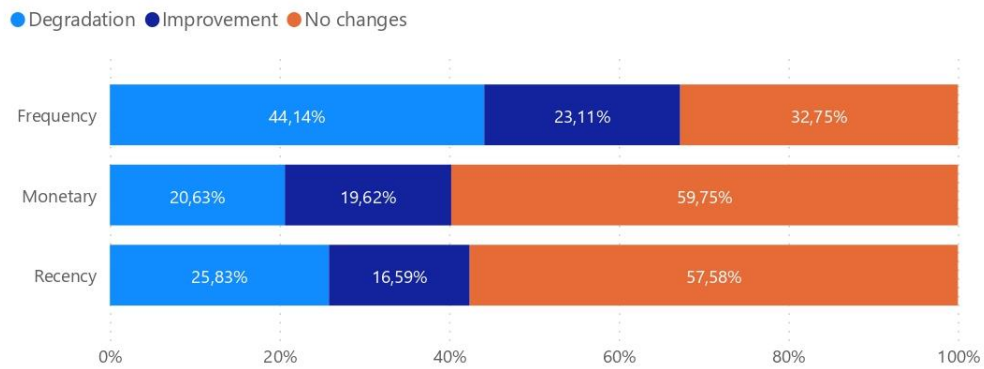


Figure 13. Dynamics of RFM indicators in City B

Source: [author research]

In this city, the frequency of visits has also changed the most. The monetary indicator changed the least, while customers both increased and decreased the purchase price equally. The recency indicator has also deteriorated to a greater extent, which is a serious negative indicator.

In more detail, the dynamics of the transition of clicks from one RFM segment to another can be presented in the following graph. (Figure 14)

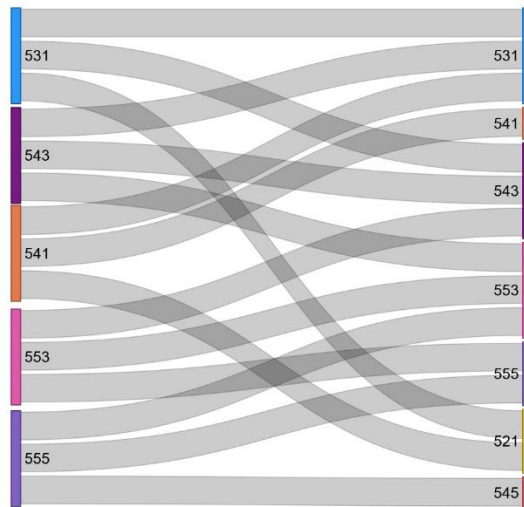


Figure 14. Dynamics of the main segments and groups of the city B

Source: [author research]

The total number of unique users for two years is exactly 31,139 people. Of these, 4.4% of new customers, or 1370 people, came in 2021 compared to 2020. However, in 2021, a significant number of customers fell into the outflow, that is, 11,5% or 3425 people. The outflow structure is shown in the pie chart below. (Figure 15)

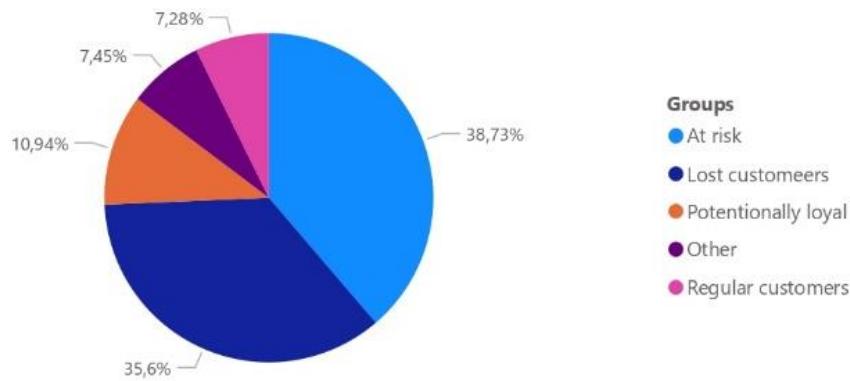


Figure 15. The shares of groups that entered the outflow in 2021 in the city B

Source: [author research]

In general, the situation is natural that more than 70% of the clicks that got into the stream belonged to groups.

3.2.3 Results for City C

The city C is the centre of the region, which imposes certain specifics on the interpretation of the results. Since this is the largest city where the analyzed supermarket chain is represented, there are also federal supermarket chains, that increases competition in the market.

The primary analysis of the average amount of purchases in this city is shown in the graph below. (Figure 16)

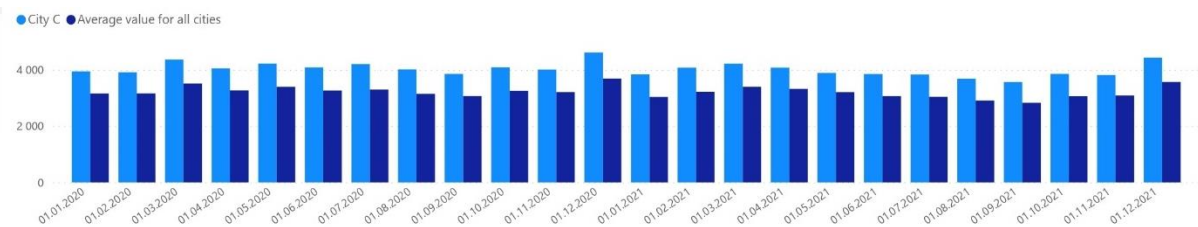


Figure 16. Average monthly sum of purchases in city C, rub

Source: [author research]

Significantly higher than the average cost for the city C in comparison with the results for all cities indicates a higher standard of living for customers.

The results of the distribution of customers into groups in both 2021 and 2020 are shown in the graph below. (Figure 17)

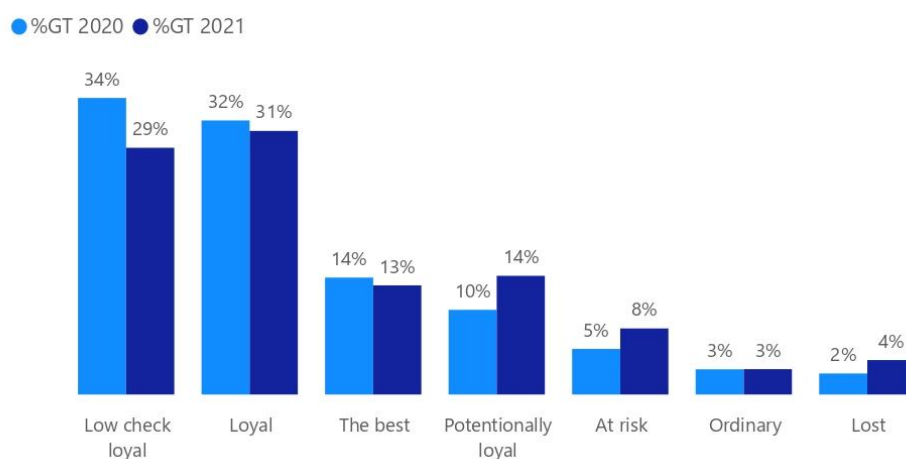


Figure 17. Dynamics of the segment share in City C

Source: [author research]

The graph shows that the basis of consumers in City C are groups of loyal consumers. The share of the best customers is also quite high. In general, despite the increase in groups at risk and lost, the situation is quite positive.

The table below contains more detailed information about the groups. (Table 14)

Table 14. Average purchase amount in City C

Group	Average purchase amount, rub		Share in the total amount, %		Growth rate
	2020	2021	2020	2021	
The best customers	8333,85	7920,45	32,96	33,05	0,95
Loyal customers	4548,992	4180,842	42,16	42,12	0,91
Low check loyal customers	1420,692	1165,075	14,25	10,99	0,82
Potentially loyal customers	2195,242	1958,092	6,29	8,89	0,89
Ordinary customers	2811,542	2576,292	2,40	2,5	0,92
At risk	969,3833	731,05	1,49	1,85	0,75
Lost customers	644,4833	464,7417	0,46	0,61	0,72

Source: [author research]

The most significant decrease in the average amount of purchases per month affected only two groups: at risk and lost customers. Nevertheless, the share of these groups in the total amount increased in 2021 compared to 2020.

The histogram below clearly shows the main changes in the indicators of the RF M analysis for the city C. (Figure 18)

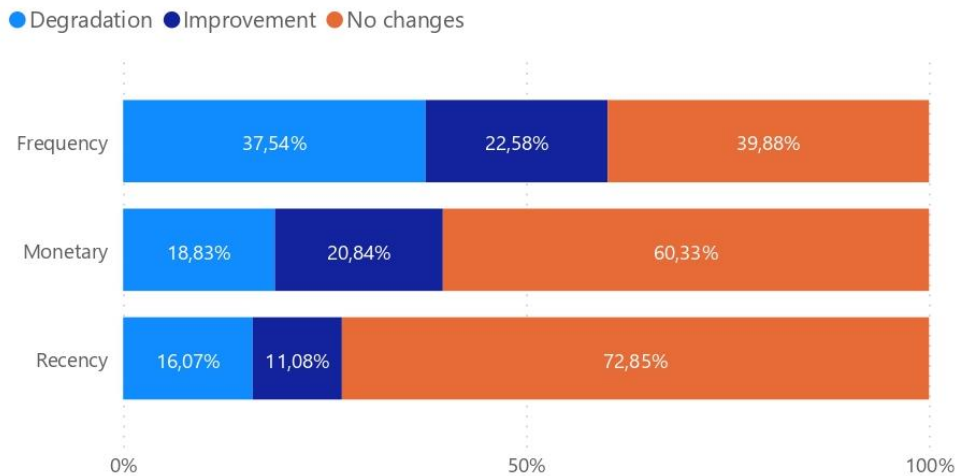


Figure 18. Dynamics of RFM indicators in City C

Source: [author research]

It is interesting that in most cases all the indicators of the RFM analysis have not changed, mostly it applies to recency indicator.

The graph below shows the change in the largest segments of 2020 and their distribution in 2021. (Figure 19)

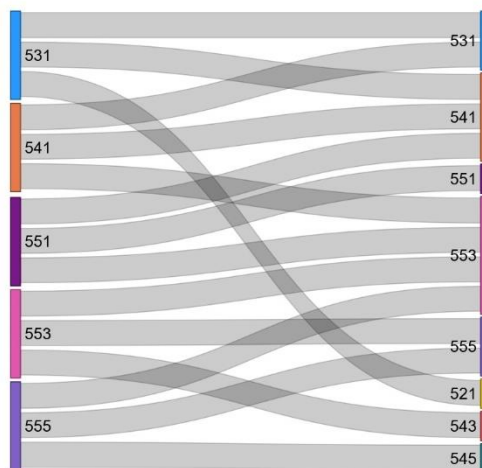


Figure 19. Dynamics of the main segments and groups of the city C

Source: [author research]

The changes occurred mainly within the following groups: the best customers, low check loyal customers and loyal customers.

During the two analyzed years, 26 147 unique customers were identified. Of this number, in 2021, 1675 people or 6% began to use supermarkets of this company. The outflow included 1210 people, or 4,6%. The excess of outflow over the number of new

customers is a positive result. However, it is also important to track the features of customer churn, which is shown in the pie chart below. (Figure 20)

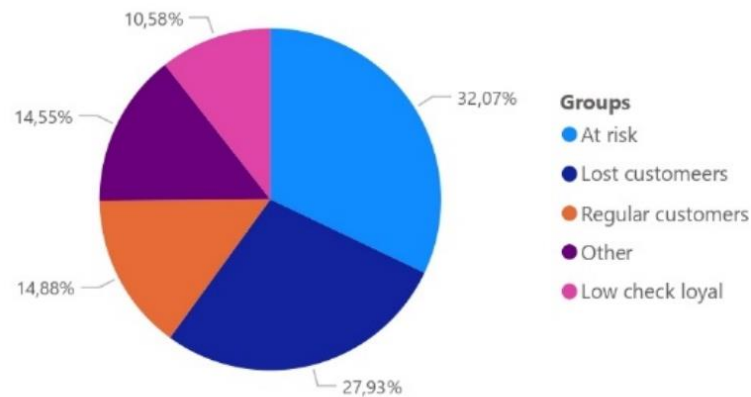


Figure 20. The shares of groups that entered the outflow in 2021 in the city C

Source: [author research]

In general, the situation is like other cities, except for the fact that the third largest outflow was a group of ordinary users.

3.2.4 Results for City D

The city D the smallest among all cities. In comparison with other cities, there are fewer supermarket chains in the city of D. However, there is still competition from independent retail stores

The primary analysis of the average amount of purchases in this city is shown in the graph below. (Figure 21)

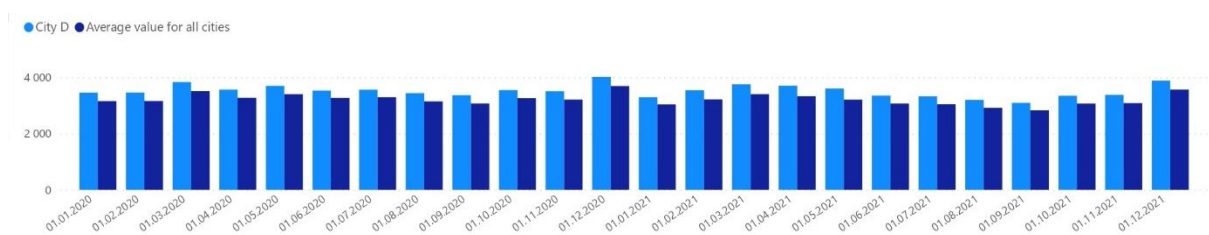


Figure 21. Average monthly sum of purchases in city D, rub

Source: [author research]

In general, the average amount of purchases in the city of D is 20% higher than the average.

Next, it is necessary to trace the dynamics of changes in customer groups in 2021 compared to 2020. (Figure 22)

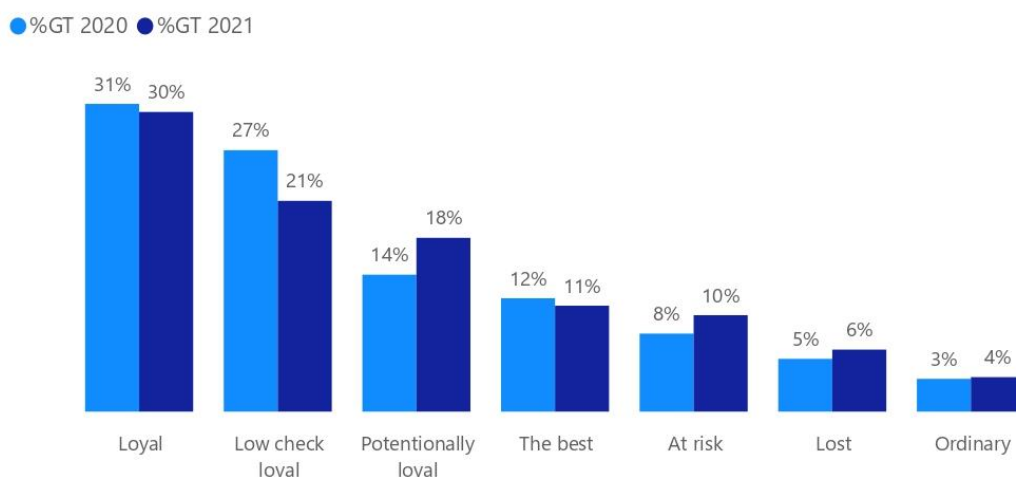


Figure 22. Dynamics of the segment share in City D

Source: [author research]

As in the case of other cities, there are certain trends in the city of D:

- Decrease in the share of loyal participants with an increase in the share of potentially loyal participants
- Increasing the share of such groups as at-risk customers and lost customers.

The table below shows the dynamics of changes in the average cost of purchases per month, as well as the shares of each group in total revenue. (Table 15)

Table 15. Average monthly purchase amount in City D

Group	Average monthly purchase amount, rub		Share in the total amount, %		Growth rate
	2020	2021	2020	2021	
The best customers	7431,7	7138,5	31,37	31,22	0,78
Loyal customers	3772,6	3513,0	43,23	43,47	0,74
Low check loyal customers	1072,9	904,6	10,45	7,87	0,84
Potentially loyal customers	1631,8	1461,4	8,32	10,49	0,93
Ordinary customers	2432,4	2218,1	8,32	3,16	0,91
At risk	880,9	682,8	2,56	1,07	0,89
Lost customers	559,2	416,0	1,10	2,72	0,96

Source: [author research]

Perhaps the most serious change is the decrease in the average monthly cost in such key groups as the best customers and loyal customers.

To identify the features of the client's behaviour, a histogram was constructed, which indicates the structure of what happened to the RFM indicators in 2021. (Figure 23)

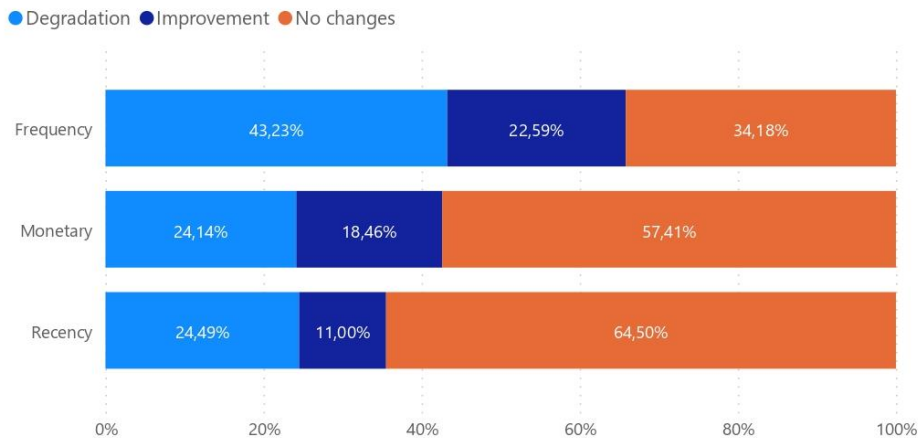


Figure 23. Dynamics of RFM indicators in City D

Source: [author research]

It can be concluded that the frequency of purchases is the most significant indicator. In general, the situation with other indicators has changed a lot. However, the percentage of improvements is also low.

The graph below shows the dynamics of the transition of the main RFM segments in 2020 to the largest groups in 2021. (Figure 24)

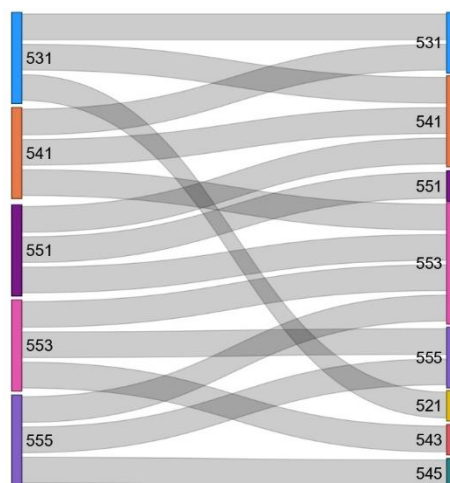


Figure 24. Dynamics of the main segments and groups of the city D

Source: [author research]

In general, the transition of the main number of customers occurred within the following groups: the best customers, loyal customers, and low check loyal check customers.

In the city of D, about 176 thousand unique users were received in two years. Of this number, 10925 users appeared in 2021, or 6%. The outflow amounted to 18301 customers or 10,3%. The outflow structure is as follows. (Figure 25)

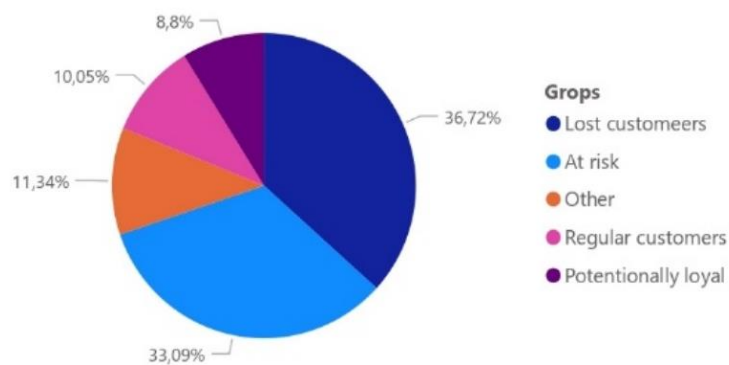


Figure 25. The shares of groups that entered the outflow in 2021 in the city D

Source: [author research]

In the city D, most of the outflow was made up of lost customers and customers at risk. Therefore, an increase in the shares of these groups makes it possible to predict an increase in outflow next year, unless several changes in the loyalty program occur.

3.2.5 Comparison of segmentation results for all cities

When comparing the results of all cities, first it is necessary to track the dynamics of the average monthly amount, that is shown in graph below. (Figure 26)

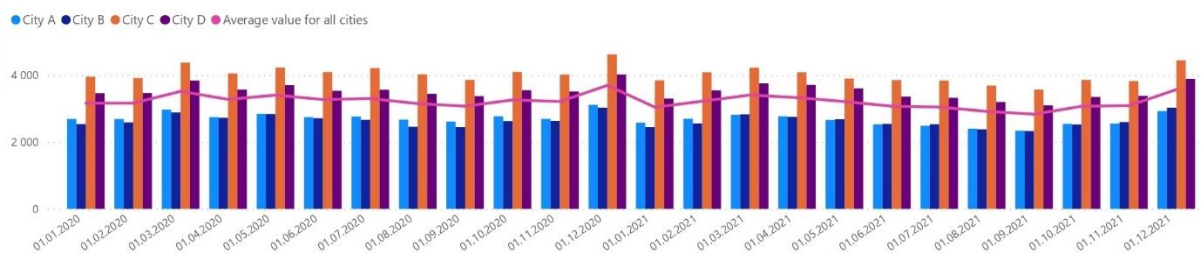


Figure 26. Dynamics of the average monthly cost of purchases in all cities

Source: [author research]

In general, the results of the comparison are not surprising. Despite the high competition from federal supermarket chains, the regional centre still has the highest average monthly number of purchases. This may be due to relatively higher wages relative to other cities. In the city of D, the high average check can rather be explained by the lack of high competition.

The table below shows the average values more clearly. (Table 16)

Table 16. Average monthly purchase amount in all cities

Group	Average monthly purchase amount, rub		Growth rate
	2020	2021	
The best customers	6976,7	6792,5	0,97
Loyal customers	3322,1	3160,9	0,95
Low check loyal customers	917,4	738,2	0,80
Potentially loyal customers	1455,2	1281,2	0,88
Ordinary customers	1989,2	1953,6	0,98
At risk	634,1	498,6	0,79
Lost customers	443,1	336,1	0,76

Source: [author research]

The table shows that such groups of customers as the best, loyal and ordinary have undergone the least changes in all cities. A serious decrease in the average cost of groups combined with low rates of frequency and recency makes the return of such customers more problematic. Also, an important problem is the reduction of the average check of a group of loyal customers with a low check. Given that such customers are regular customers, this group should be stimulated with information about promotions and special offers.

In general, for proposals to improve the loyalty program of the supermarket chain, an average value was obtained for the size of customer groups, which is shown in the histogram below. (Figure 27)

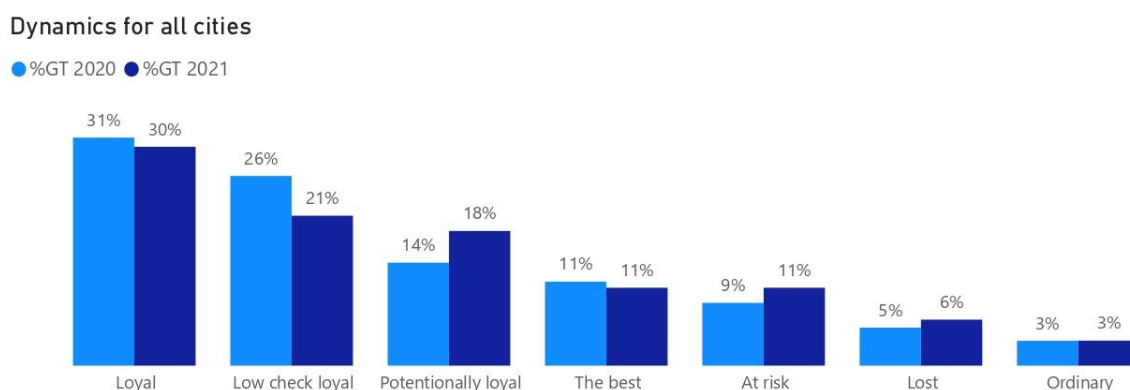


Figure 27. Dynamics of segments for all cities

Source: [author research]

The histogram shows that the main problems are the decrease in the frequency of visits and the growth of the group of potential customers, respectively.

Summing up all the above, it is possible to make a conclusion, that in all cities of the analyzed supermarket chain, the main problem in 2021 was a decrease in the frequency of visits, which led to a decrease in the total amount of purchases. Therefore, when updating a new loyalty program, it is important first to encourage customers to visit stores as often as possible.

3.3 Development of a loyalty programme based on the received customer segmentation

Based on the results obtained, it can be concluded that it is necessary to pay more attention to the four main groups:

- The best customers
- Loyal customers
- Low check loyal customers
- Potentially loyal customers.

Despite the good indicator monetary, the group of ordinary consumers was not selected for special events for the following reasons:

- low recency, which is, as already mentioned earlier, the most important indicator of the RFM analysis
- a small number of people in the group (for all cities, the share of this group is 3%).

Nevertheless, with the improvement of the loyalty program, the customers of this group can significantly improve their performance.

The total summarized information about these groups is presented in the table below. (Table 17)

Table 17. Characteristics of main customer groups

Group	Number of participants	Share of the total number of participants, %	Average monthly purchase amount	Share of the total amount, %
The best customers	28 956	11	6792,5	46
Loyal customers	78 972	30	3160,9	21
Low check loyal customers	55 280	21	738,2	5
Potentially loyal customers	47 383	18	1281,2	8

Source: [author research]

These groups were selected because of their largest shares of participants and share of the amount of purchases of group members in total revenue. The remaining groups have a high probability of outflow, as well as a relatively low average monthly number of purchases. So, there is a risk that special events specifically for these groups may be irrational and lead to high costs with low efficiency. Nevertheless, such groups may be positively affected by general changes in the loyalty program.

Such groups as low check loyal customers and potentially loyal customers, have certain features, and both similarities and differences, as can be seen in the table. (Table 18)

Table 18. Comparison of customer groups

Group name	Similarity	Difference
Low check loyal customers	The recency indicator for these groups has the highest score - 5. That means, that customers of these groups made purchases relatively recently, somewhere up to 2 months	- Customers visited supermarkets quite often - Customers made purchases for an average or above average amount
Potentially loyal customers		- Customers did not visit supermarkets so often - Customers made purchases for an average or above average amount

Source: [author research]

Based on the table, it can be concluded that group of low check loyal customers has the advantage of the frequency of visits, while group of potentially loyal customers is important because although customers do not come so often, they are ready to spend quite a lot on purchases.

Based on this, even without significant changes to the current loyalty program, the analyzed supermarket chain can already affect the situation for example by email and SMS mailing.

These methods are among the most effective methods of reminding store customers and stimulating purchases. At the same time, SMS mailing is still the most effective (provided that the clients of these groups have agreed to receive SMS messages), since very often the participant either does not have time to read all the mailings, or the message may simply get lost among other messages. When consumer receive SMS messages, a reminder is displayed on the phone screen and a short message with information about current promotions or the number of accumulated points may interest the consumer.

With the help of these methods, the potentially loyal customers will receive a reminder about the store and with the help of promotions they may be interested in visiting

the store more often, and the low check loyal customers will receive more information about the current supermarket offers, which may motivate them to buy more products of different categories.

Such groups as the best customers and loyal customers do not need special incentives, so the main task in this case is to continue to motivate customers to buy, which can be achieved by changes in the loyalty program. The study of the types of loyalty programs has shown that in this case, an effective way is to transform the loyalty program depending on the segmentation indicators.

As mentioned earlier, now the most popular and effective for supermarket chains is the bonus system. However, it is possible to achieve the greatest efficiency with the help of multi-level loyalty programs. For this case, the following possible options were selected for changing the existing loyalty program. Among such loyalty programs, there may also be different types, each of which has both advantages and disadvantages, which is presented in the table. (Table 19)

Table 19. Comparison of proposed types of loyalty programs

Type of loyalty program	Advantages	Disadvantages
Accrual of bonuses depending on the number of purchases	The ability to buy and write off bonuses faster can motivate customers to make more purchases to reach the maximum level	There may be a risk of reducing the frequency of visits when a certain amount is reached
Loyalty bonus program with burnable bonuses after a certain period	To write off the accumulated bonuses, customers will visit supermarket chain stores more often	Reduced loyalty among customers who prefer to accumulate points

Source: [author research]

Based on this, as an improvement of the existing loyalty program, it is proposed to introduce additional levels of receiving bonuses depending on the purchases spent per month, with maintaining the conditions for calculating bonuses for certain groups of goods (own-brand goods, promotional goods), as well as maintaining discounts during special offers for cardholders. Also, for the new loyalty program, the condition remains that it is possible to pay up to 99% of the purchase price.

The process of the proposed loyalty program is shown in the diagram below. (Figure 28)

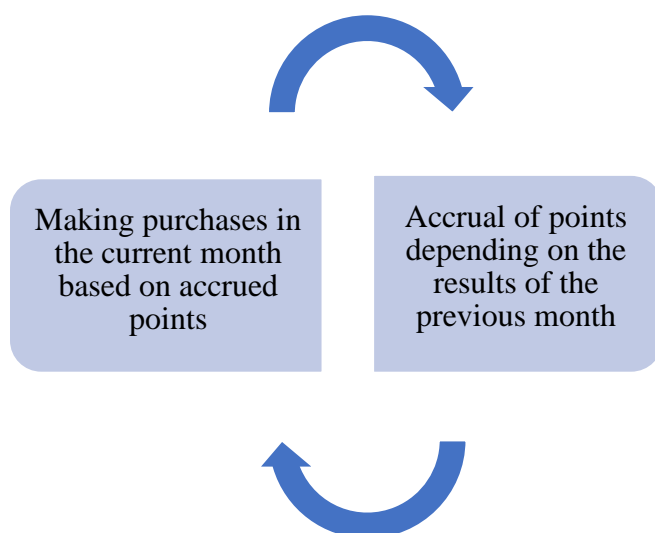


Figure 28. The process of accrual of points under the proposed loyalty program

Source: [author research]

The table below shows the dependence of the amount of percentage accrual on the number of purchases per month. (Table 20)

Table 20. Conditions of the proposed loyalty program

Percentage of bonus accrual	Purchase amount per month
1%	<999 rub
2%	>1000 rub < 1999 rub
3%	> 2000 rub < 4999 rubs
4%	> 5000 rubs < 7999 rubs
5%	> 8000 rubs

Source: [author research]

To develop a new bonus accrual system, the results of average checks for the received consumer groups were used. In particular, the average maximum number of purchases per month in all cities is about 7000 rubbles, but in the city C in 2020 this value exceeded 8000 rubbles, and in the city of D it also fluctuates around this amount. Thus, the maximum value can significantly increase the average check, as well as return the indicators of the regional centre to at least the level of 2020.

In all cities, even for the regional centre with the highest average amount, the average amount of the check for groups «at risk customers» and «lost customers» does not exceed 1000 rubbles. Nevertheless, such consumers still should not be underestimated, as there remains a chance to transfer this type to other more loyal categories. Therefore, for such clients, it was decided to keep the initial bid in order not to lose them. For a regional centre where the average amount of purchases was close to 1000 rubbles in 2020, there is a very big chance to increase the average check. Such value can also motivate the group of low check loyal customers to increase the amount of the check, which is especially true for cities A and B where the average amount is below 1000 both in 2020 and 2021 years.

For ordinary customers who do not visit stores so often, but often have a good average check (on average in all cities about 2000 rubbles), the introduction of new loyalty program with can become a motivation to visit stores more often to be able to save and spend bonuses faster.

For loyal customers, whose share in all cities has not significantly decreased in 2021, the new loyalty program will simply be another reason to visit the stores of this supermarket chain rather than the stores of competitors. The problem of preference of competitors' supermarkets chains is especially relevant to the city A, B and C, as there is an active development of federal supermarket chains.

The analysis showed that about 26% of customers made purchases in the area of the threshold amount, so based on this, it can be expected that such changes will inspire such customers to increase the average check in order to be able to receive more bonuses next month.

Before launching, it is advisable to test the program on a small number of the most loyal consumers. This will allow to identify the main problems and fix them. Many companies, when launching, give customers a part of the cards or sell them on special terms. This allows you to recruit the necessary minimum of participants.

To check the effectiveness of the proposed loyalty program, for tracking changes it can be used such indicator as an increase in the average amount of purchases per month, as well as an increase in the number of visits.

For the most objective assessment of the new loyalty program, in theory, a large amount of time should pass, that is, at least a year, especially for comparison with the results in 2021 and 2020. This is necessary since there is a certain seasonality in behaviour of the supermarket chain consumers. In particular, the traditional decline in sales in summer and in September.

3.4 Conclusions on chapter 3

Based on the analysis, it can be concluded that, in general, the problems of frequency reduction affected all cities.

However, City C showed the most positive results among all cities:

- the excess of the influx of new customers over the outflow
- a high average amount of purchases
- a high proportion of loyal user groups.

At the same time, since this city is a regional centre with higher competition, there were expectations that this fact would significantly worsen the indicators.

In the city of D, despite the small number of competitors, the outflow of customers in 2021 was the highest. This could be influenced by high prices for products relative to competitors, which makes it unprofitable for city residents to purchase in a store.

To increase customer loyalty and increase their average check, as well as the frequency of visits, it was decided to update the existing loyalty program. Now the amount of points accrual will depend directly on the amount spent for the month. Such changes can significantly improve the situation in the analyzed supermarket chain.

Conclusion

The main goal was to improve the existing loyalty programme through customer segmentation to achieve the goal, the following tasks should be completed: data cleaning and preparation, conducting customer segmentation and identifying key changes, data visualization, problem identification and loyalty program development.

In the process of research, the following theoretical and practical results were obtained:

1. Now, retail companies mainly use the loyalty bonus program since it is more profitable for the company.

2. To determine outliers in data on customer purchase amounts, it is rational to use the interquartile method, as well as a box plot for visualisation.

3. For all cities except the regional centre, the frequency index decreased the most. In the regional centre, most indicators have mostly not changed.

4. In the regional centre, despite the high competition from other supermarkets, the situation is the most favourable, since the number of new users exceeded the number of those who left.

5. In general, the main problem throughout the network is the reduction in the frequency of visits.

6. To solve the problem of the frequency of visits, as well as to increase the average amount of purchases, a new multi-level loyalty bonus program was proposed with a dependence of the percentage of bonus accrual on the number of purchases for the previous month.

Overall, the research goal was achieved - all tasks and objectives were completed successfully.

Recommendations for future work

The research of the characteristics of customer behaviour can be improved and deepened with the help of a forecast of customer behavior for the next year by months, in particular, the forecast of customers who may be in the outflow group. Also, further work involves forecasting, taking into account the proposed recommendations and comparing these indicators. In general, subsequent research should be directed to the practical implementation and verification of proposals.

References

1. Ballestar M.T., Grau-Carles P.G., Sainz J. “Customer segmentation in e-commerce: Applications to the cashback business model.” In *Journal of Business Research*, pp. 407-414. 2018.
2. Beck J.T., Chapman K., Palmatier R.W. “Understanding Relationship Marketing and Loyalty Program Effectiveness in Global Markets.” In *Journal of International Marketing*, pp. 1–21. 2015.
3. Butcher S. “Loyalty programs and regular customer clubs” 2004
4. Bombaij N.J., Dekimpe M.G. “When do loyalty programs work? The moderating role of design, retailer-strategy, and country characteristics.” In *International Journal of Research in Marketing*, pp. 175-195. 2020.
5. Chaabane A.M., Pez V. “Make me feel special: Are hierarchical loyalty programs a panacea for all brands? The role of brand concept.” In *Journal of Retailing and Consumer Services*, pp. 108-117. 2017.
6. Chang H.H, Lin Y., Lu Y., Chang R.C. “Understanding the Effects of Loyalty Program Relational Benefits, Emotion Elicitation, and Self-disclosure on Long-term Customer Relationship.” In *International Journal of Information and Management Sciences*, pp. 237–260. 2021.
7. Chaudhuri M., Voorhees C.M., Beck J.M. “The effects of loyalty program introduction and design on short-and long-term sales and gross profits.” In *Journal of the Academy of Marketing Science*, pp. 640-658. 2019.
8. Cheng C.H., Chen Y.S. “Classifying the segmentation of customer value via RFM model and RS theory.” In *Expert systems with applications*, pp. 4176 – 4184. 2009.
9. Chernysheva A.M. and Yakubova T.N. *Branding*. [In Russian.] Moscow: Urait. 2020
10. Cho Y., Moon S.C. “Weighted mining frequent pattern-based customer's RFM score for personalized u-commerce recommendation system.” In *J. Converg.*, pp. 36-40. 2013.
11. Christya A. J., Umamakeswaria A., Priyatharsinib L., Neyaab A. “RFM ranking – An effective approach to customer segmentation.” In *Journal of King Saud University - Computer and Information Sciences*, pp. 1251-1257. 2021.
12. Coussementa K., Bosschebc F, Bock K. “Data accuracy's impact on segmentation performance: Benchmarking RFM analysis, logistic regression, and decision trees.” In *Journal of Business Research*, pp. 2751-2758. 2014.
13. Fang Z., Huang L., Wiermanc A. “Loyalty programs in the sharing economy: Optimality and competition.” In *Performance Evaluation*, pp. 1-29. 2020.

14. Gorlier T., Michel G. "How special rewards in loyalty programs enrich consumer-brand relationships: The role of self-expansion." In *Psychology & Marketing*, pp. 588-603. 2020.
15. Han J., Kamber M., Pei J. "Outlier Detection." In *Data Mining (Third Edition)*, pp. 543 – 584. 2012.
16. Ho Y.C., Ho Y.J., Tan Y. "Online cashback shopping: Implications for consumers and e-businesses." In *Information Systems Research*, pp. 250-264. 2017.
17. Hughes A.M. *Strategic database marketing*. Chicago: Probus Publishing Company. 1994
18. Tsipstsis K., Chorianopoulos A. *Data mining techniques in CRM: inside customer segmentation*. New York: John Wiley & Sons. 2011.
19. Kim J., Steinhoff L., Palmatier R.W. "An emerging theory of loyalty program dynamics." In *Journal of the Academy of Marketing Science*, pp.71–95. 2021.
20. Martinez R. G., Carrasco R. A., Garcia-Madariaga J., Gallego C. P., Herrera-Viedma E. "A comparison between Fuzzy Linguistic RFM Model and traditional RFM model applied to Campaign Management. Case study of retail business" In *Procedia Computer Science*, pp. 281-289. 2019.
21. Mikaelyan I. "The largest retailers are changing the strategy of working with customers. What does it mean?". RB.RU, May 17, 2017. <https://rb.ru/opinion/sale-vs-bonus/> (accessed April 25, 2022).
22. Nastasoiu A., Vandenbosch M. "Competing with loyalty: How to design successful customer loyalty reward programs." In *Business Horizons*, pp. 207-214. 2019.
23. Nikishkin V.V., Tverdokhlebova M.D. "Improving the efficiency of retail services in the bookstores on the basis of successful loyalty programs." In *Journal of Creative Economy*, pp. 2490-2496. 2019.
24. Pupentsova S.V., Nekrasova T.P., Pavlenko I.A. "Loyalty program management in retail trading networks." In *Journal of Creative Economy*, pp. 195 – 212. 2019.
25. Rahim M. A., Mushafiq M., Khan S., Araind Z. "RFM-based repurchase behavior for customer classification and segmentation." In *Journal of Retailing and Consumer Services*, pp. 1-9. 2021.
26. Ramaseshan B., Ouschan R. "Investigating status demotion in hierarchical loyalty programs." In *Journal of Services Marketing*, pp. 650-661. 2017.
27. Sheshasaayee, A., Logeshwari, L. "An efficiency analysis on the TPA clustering methods for intelligent customer segmentation." In *2017 International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, pp. 784–788. 2017.

28. Siti M., Putri N., Rice N. "Analysis for Customer Lifetime Value Categorization with RFM Model." In *Procedia Computer Science*, pp. 834 – 840. 2019.
29. Vieira V.A., Agnihotri R., Almeida A., Lopes E. "How cashback strategies yield financial benefits for retailers: The mediating role of consumers' program loyalty." In *Journal of Business Research*, pp. 200 - 212. 2022.
30. Yang M.X., Chan H., Yu I., Fock H. "Consumer motivation for reward pursuit: A culture-based and progress-based model of loyalty program effectiveness." In *Journal of Global Marketing*, pp. 255-268. 2019.
31. Zahrotun, L. "Implementation of data mining technique for customer relationship management (CRM) on online shop tokodiapers.com with fuzzy c-means clustering." In *2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, pp. 299–303. 2017.
32. Zeithaml V.A., Berry L.L., Parasuraman A. "The behavioral consequences of service quality." In *J. Market.*, pp. 31-46. 1996.
33. Zenga K. J., Yub I.Y., Chana H. "Communication strategies for multi-tier loyalty programs: The role of progress framing." In *Tourism Management*, pp. 1-13. 2022.