applied sciences

MDPI

Article

Cluster-Based Approaches toward Developing a Customer Loyalty Program in a Private Security Company

Arthur de Sousa [1], Sérgio Moro [2,*] and Renato Pereira [3]

[1] Instituto Universitário de Lisboa (ISCTE-IUL), DCTI, 1649-026 Lisboa, Portugal; arthur_sousa@iscte-iul.pt
[2] Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, 1649-026 Lisboa, Portugal
[3] Instituto Universitário de Lisboa (ISCTE-IUL), BRU, 1649-026 Lisboa, Portugal; renato.pereira@iscte-iul.pt
* Correspondence: sergio.moro@iscte-iul.pt

Abstract: This study aimed to create a loyalty program for a private security company's most valuable customers using clustering techniques on a dataset from the company. K-means was employed as an unsupervised machine learning algorithm to segment customers. Performance evaluation metrics, including the silhouette coefficient, were utilized to compare various algorithmic approaches. As a distinctive feature of this study, in addition to the evaluation metric, strategic questionnaires were administered to business decision-makers to facilitate the integrated development of a loyalty program with key stakeholders invested in customer retention and profitability. The results show the existence of three customer clusters with an optimal silhouette coefficient for loyalty program development. Interestingly, the customer group to be targeted for the loyalty program did not exhibit the highest silhouette coefficient metric. Business leaders selected the group they perceived as most efficient for program implementation. Consequently, the study concludes that customer segmentation not only entails statistical analyses of individual user groups but also requires a comprehensive understanding of the business and collaboration with stakeholders. Furthermore, this study aligns with findings from other authors, demonstrating that private security companies can benefit from implementing a loyalty program, although avenues for further investigation remain.

Keywords: loyalty program; clustering; customer segmentation; k-means; private security companies

Citation: de Sousa, A.; Moro, S.; Pereira, R. Cluster-Based Approaches toward Developing a Customer Loyalty Program in a Private Security Company. *Appl. Sci.* **2024**, *14*, 78. https://doi.org/10.3390/app14010078

Academic Editor: Andrea Prati

Received: 6 November 2023
Revised: 4 December 2023
Accepted: 11 December 2023
Published: 21 December 2023

1. Introduction

Understanding customer behavior and demands is a crucial task in the corporate world and can contribute to customer satisfaction. Customer behavior can lead a company to either success or failure. Perceiving customer behaviors can be useful in developing a company's marketing and operation strategies for satisfying customer requirements [1]. Thus, identifying customer behavior plays a significant role in understanding the factors that lead a customer to make repeat purchases from the same company, engage in long-term recurring service payments, or simply discontinue their relationship.

Customer loyalty can be understood as favorable behavior toward a company, evidenced through a customer's likeliness to do business with a given retailer repeatedly, have a preference for a certain brand, and provide word-of-mouth advocacy [2]. According to Murugiah et al. [3], some of the consequences of customer loyalty include increased purchases and cross-selling, as well as recommendations, as these customers tend to provide more information about themselves. This enables institutions to be accurate in their proposals.

Loyalty programs are widely employed across various industrial sectors. Dogan et al. [4], in a case study conducted within the retail industry, asserted that loyalty programs are commonly utilized in the sports industry as a driver for customer retention. Another study conducted by Zavalishchin [5] which investigated optimal loyalty program management illustrated how the Russian railway transport sector employs loyalty programs in an endeavor to enhance the profitability of the companies responsible for this segment. Additionally, according to Mrkosová et al. [6], in the United States, 75% of business in the retail network manage at least one loyalty program with their respective customers.

Despite the current literature indicating that an increasing number of diverse industries are employing loyalty programs as catalysts in their respective businesses, no study has been found to demonstrate the creation of such programs within private security companies. Thus, the objective of this study is to develop a loyalty program for a private security company's most valuable customers based on customer segmentation techniques applied to a dataset from the company.

One of the most frequently employed techniques for identifying similar customers and grouping them into segments with similar characteristics is referred to as cluster analysis. Cluster analysis techniques represent the most suitable methods that enable companies to discern segments or groups of customers with the aim of reaching a potential user base [7]. Originating from the field of marketing, the RFM model is a technique used to summarize a customer's past behavior based on recency, frequency, and monetary value attributes [8]. As per Dogan et al. [4], this technique has been employed for over 50 years for customer segmentation purposes. Therefore, this study proposes an integrated conceptual model that includes constructs based on previous research in customer segmentation and loyalty program development.

1.1. Contribuition

The proposed model was empirically validated through a study conducted by Alkharat [9] within a telecommunications company. This study aimed to achieve customer segmentation by employing behavioral, demographic, and financial variables. As far as we know, this is the first study that proposes creating a loyalty program for the most valuable customers of a private security company. Nevertheless, the present model adopts a comprehensive set of variables encompassing personal, geographical, and behavioral characteristics from which hypotheses are formulated and tested. The key contributions of this research are as follows:

- Using a variable selection procedure to enhance the original model (used as a baseline), i.e., RFM;
- Comparing the proposed approach with previous research (used as a baseline) and analyzing the improvement achieved considering the relevance of the proposed model;
- Finally, demonstrating how the model utilizing customer historical information can benefit the private security business in terms of customer segmentation.

1.2. Organization

This paper is organized as follows: Section 2 presents a literature review of existing studies conducted in the area of the research topic. Section 3 provides a comprehensive overview of related work, along with various techniques employed throughout the study, followed by a description of the proposed framework for analyzing the identified behaviors. Section 4 reports the results of the proposed framework in a case study and analyzes the discovered patterns. Managerial implications are presented in Section 5. This paper is concluded in Section 6.

2. Related Research

2.1. Loyalty Programas

Customer segmentation is one of the most fundamental processes in marketing, involving the categorization of customers into representative groups with distinct consumption needs [7]. The customer segmentation process can assist companies in comprehending customer behavior and, consequently, in developing effective marketing strategies for different target customer groups. As stated by Mihova and Pavlov [10], customer segmentation is a complex process that demands expertise in financial product sales, market knowledge, and intuition.

This segmentation can be carried out using a variety of different customer characteristics. The most common variables include customer geographic regions, customer demographic data (e.g., age, gender, marital status, income), customer psychographics (e.g., values, interests, lifestyle, group affiliations), and purchasing behavior (e.g., past purchases, shipping preferences) [7]. Companies have employed several segmentation criteria and techniques to identify and better understand customer groups and offer preferred products and services. However, many marketing professionals struggle to identify the right customer segments for organizing marketing campaigns [11].

A loyalty program is more than a tool used to strengthen the loyalty of existing customers. Leenheer et al. [12] define a loyalty program as "an integrated system of marketing actions aimed at making customers more loyal". The customer should become a member and identify as such. Many companies develop loyalty programs to increase customer retention. Becoming a member of a loyalty program enables customers to enjoy special privileges, often bearing the designation of a "very important person" (VIP), thereby encouraging increased purchasing behavior [5].

Current state-of-the-art research demonstrates that an increasing number of different industries are employing loyalty programs as drivers in their respective businesses. Most loyalty programs are associated with everyday consumer goods and services, particularly a range of food products, pharmaceuticals, and apparel [6]. However, the same author argues that these programs are also valuable, particularly for companies with an extensive network of stores and/or a complex range of products.

The existing literature presents several significant findings regarding the adoption of loyalty programs by businesses. Customers who are members of loyalty programs exhibit greater loyalty and improved behavioral attitudes toward the retailer, satisfaction with the loyalty program, and an enhanced relationship with the company [13]. According to Melnyk et al. [14], customers who enroll in loyalty programs make more purchases than non-members. Furthermore, the results of an indirect analysis conducted by Koo et al. [15] suggest the presence of an affective commitment relationship between the perceived value of a loyalty program and brand loyalty itself.

2.2. Loyalty Programs and Data Science Approaches

Segmenting customers and implementing more successful loyalty programs has become easier and useful in recent years thanks to advances in data mining techniques [2]. One of the goals of segmentation is to ascertain the attitudes of individual customer groups toward a specific product or service [10].

Clustering techniques are the most suitable methods enabling companies to identify segments or groups of customers with the aim of reaching a potential user base [7]. Customer-related data typically contain thousands of observations. One of the most employed techniques for finding similar customers and grouping them into homogeneous segments is referred to as cluster analysis.

According to Dogan et al. [4], companies have generated more valuable data to enhance customer segmentation using the above-mentioned techniques, depending on their data analysis and interpretation capabilities. The RFM technique, which is based on recency, frequency, and monetary value characteristics, has been in use for over 50 years for customer segmentation purposes [4]. RFM can be utilized to calculate customer lifetime value (LTV) [8]. According to Moro et al. [16], LTV can be considered a significant factor in every decision aimed at improving customer relationships and profitability. The author further asserts that LTV is associated with predicting future customer behaviors.

The current literature presents numerous different approaches to customer segmentation and loyalty program creation. Mihova and Pavlov [10] applied the k-means algorithm in the banking sector to create three clusters of loyal borrowers: "platinum", "gold", and "silver". In this study, the authors employed three variables: the total Euro value of a loan requested, customer tenure in months, and the quantity of "non-payments" by customers in the past 12 months. The same k-means approach was utilized in the study conducted by

Sharma et al. [17]. This study identified four groups of influencing factors for selecting a specific service provider in the telecommunications industry. Personal data such as age and occupation; monetary data like recurring monthly expenditure; and behavioral data were used for segment creation.

The hierarchical clustering technique is evident in a study conducted by Güçdemir et al. [18] and in another study by Tanford et al. [19]. The former applied the mentioned method to identify commercial customers of an international original equipment manufacturer. Variables derived from the RFM model were used, such as “loyalty score”, “average annual demand”, “long-term relationship potential”, “average percentage change in annual sales revenue”, and “average percentage change in annual demand”. The latter study identified four loyalty segments indicative of customer behavior in the casino industry. To achieve this, the authors utilized behavioral and attitudinal dimensions.

2.3. Private Security Companies and Hypotheses

The presented studies indicate that loyalty programs are prevalent across various industries, with distinct approaches stemming from diverse variables. Managers should keep in mind that customer behaviors are in a constant state of flux, as are their needs [10]. A loyalty program should not be regarded as a one-time action but rather as an ongoing process. A cluster analysis can assist businesses in developing a robust loyalty program, and the private security sector is no exception.

Despite the current literature lacking studies that specifically address the development of a loyalty program or even customer segmentation within this industry segment, based on the results of the presented studies, the following hypothesis is proposed for the creation of loyalty programs in private security companies:

1. *H1.* Private security companies can establish loyalty programs for their customers similar to those in any other industry.

Khajvand et al. [20], in their study conducted in the health and beauty industry, proposed a model that grouped customers into segments based on RFM. As a result, grouping customers into different segments helped decision-makers identify market segments more clearly and develop more effective marketing and sales strategies for customer retention. In order to identify key customer clusters in the banking sector and create more efficient marketing strategies, Ansari and Riasi [21] employed cluster analyses based on RFM variables. However, Alkharat [9] used financial, behavioral, and geographical variables to create customer segmentation in a telecommunications company. Therefore, the following hypotheses were proposed:

2. *H2.* RFM variables are sufficient for the creation of a loyalty program.
3. *H3.* Behavioral and geographical variables, combined with variables for building an RFM model, yield better results in loyalty program creation.

According to Bell et al. [22], a loyalty program can reward customer purchasing behavior or select profitable customers and provide them with privileges and additional services. Leenheer et al. [12], in their study analyzing the implementation of loyalty programs in the retail industry and their effectiveness, argue that loyalty programs can help allocate marketing budgets more efficiently if customers differ substantially in terms of profitability. Consequently, the following hypothesis was formulated:

4. *H4.* It is possible to create marketing strategies for specific customer segments identified through cluster analysis techniques.

3. Materials and Methods

The proposed research methodology included three main steps. The first phase was related to the collection of data from a multinational private security company operating in Portugal (the market that was the target of the present analysis) and their related efforts regarding the cleaning, manipulation, processing and creation of variables for the later segmentation of the data. In the second step, from the creation of the new variables, the k-means clustering technique was applied to create customer clusters. Finally, a questionnaire was applied to the stakeholders to evaluate the results found in the study. The complete process of the steps of the methodology is presented in Figure 1.

Figure 1. Proposal for methodology development (source: the authors).

3.1. Data from a Private Security Company

The present study used real data from a private security company on their respective customers. As a private security company, the company offers products such as theft alarms, anti-intrusion motion sensors, and image detection cameras. Its main customers are named as individuals or businesses. The former refers to houses and apartments and the latter to establishments that provide some service, such as pharmacies, jewelers, and pastry shops.

The accuracy of customer segmentation mainly depends on the selection of variables and the techniques used [9]. A study conducted by the same author in a telecommunications company, with the aim of creating a customer segmentation, used behavioral variables (e.g., number of calls, duration of calls, reasons), demographic variables (e.g., location, age, number of products, types of products), and financial variables (e.g., total spent in the analyzed period, debts contained in the period). Dogan et al. [4] used the variables in the RFM model applied to the k-means technique to conduct customer segmentation in a retail industry.

The original data collected for this study consisted of different types of variables: behavioral (e.g., number of service expansions, number of times in collections, time in collections), geographical (e.g., location), and financial (e.g., monthly fee, installation fee, MVC). Due to confidentiality issues, details beyond those provided in this manuscript about the data sources from which the data were collected cannot be disclosed. Nevertheless, the sample included all active customers of the company, meaning customers who were paying a recurring monetary amount until the date of data collection. The data were collected on 12 July 2022 and cover the entire period of the organization's operation with its customers. For example, the oldest customer relationship dated to 1995, corresponding to 27 years of seniority as an active customer of the company. In addition, the data consisted of 88,390 observations that corresponded to the unique contracts of all customers. Of the total volume of customers, 15.5% had more than one active contract with the company, meaning a single customer may appear multiple times in the database. Furthermore, 18 characteristics were considered, described in Table 1, with the main objective of performing a customer segmentation process that best fits this dataset.

Table 1. Initial dataset for customer segmentation in a private security company (source: the authors).

Attributes	Description
contract id	Unique identification code for each customer contract.
customer id	Unique identification code for each customer.
date installed	Date when the service was installed at the customer's location.
Location	Region where the service contracted by the customer is located.
contract type	Contracts are defined as "private" or "business". The first refers to a residential contract, and the second is characterized by promoting a negotiation and/or service.

Table 1. *Cont.*

Attributes	Description
local description	Description of the type of contract (e.g., housing, apartment, pharmacy, restaurants).
payment type	Method by which the monthly fee payment is paid by customers.
invoice type	How invoices are sent for payment.
monthly fee	Amount paid monthly for the service contracted by the customer.
installation fee	Amount initially paid by the customer to install the service.
extensions amount	Number of extensions added to the initial contract, which indicates new services contracted by the customer.
extensions fee	Addition to the monthly fee initially contracted by the customer.
first extension date	Date of the first extension added to the service initially contracted by the customer.
last extension date	Date of the last extension added to the service initially contracted by the customer.
debt amount	Number of times the contract was managed by the debt department.
time in debt	Time in days that the contract was managed by the debt department.
MVC	Internal financial variable of the company which evaluates how much a customer is profitable for the company after removing all its costs.
debt	Whether or not the customer had debts up to the day of data collection.

3.1.1. Data Preprocessing

A data cleaning process was carried out, and missing and incorrect values were excluded from the analysis. Furthermore, only customers who did not have outstanding debt management participated in the study. In this way, 83,234 observations related to all contracts and active customers for the company until the date of data extraction were extracted and used in this study.

3.1.2. Feature Creation

The original raw data contain only variables that describe the history of the universe of contracts. As the goal of the study was to find more valuable customer groups, several transformations were necessary to obtain more meaningful variables. According to the company, the most valuable customers are those who have high gross revenue. In other words, a valuable customer pays a high amount throughout their lifetime experience after eliminating all their costs. For this study two types of variable groups were created to have the customer as the main subject and based on the original raw data.

For the creation of the first group of variables, this study utilized the RFM model adapted from the study by Dogan et al. [4], where R refers to the most recent purchase. The current study used the most recent extension of initially contracted services, referring to the R value. The variable F represents the total number of expansions made, and M represents total customer expenditure throughout their operational period with the company. The latter is calculated from the initiation of services with the company, that is, from their installation date to the data extraction date. The attributes chosen to create this first group of variables were the installation date, monthly fee, installation fee, extension fee, extension amount, first extension date, and last extension date.

The second group of variables was created from a questionnaire provided to the company's main managers to collect more relevant information based on the initially collected raw data. Table 2 presents the characterization of the respondents who were part of the research.

Table 2. Characterization of the interviewees (source: the authors).

Interviewee ID	Gender	Age	Department	Role	Year of Experience		Number of Institutions Worked	Field of Study	Post-Graduate
					Company of study	Other institutions			
1	Female	52	Customer service	Director	17	9	4	Law	People management
2	Male	47	Customer acquisition	Director	5	18	5	Psychology	Master of business administration
3	Female	38	Loyalty	Director	3	10	6	Management	Data for business
4	Female	43	Debt	Director	13	11	3	Management	Accounting and financial management
Total					38	48	18		
Average					9.5	12	4.5		

Figure 2 presents the methodology applied in the questionnaire, which has its premises defined in five steps: the presentation of the motivation and the identification of the problem, the definition of objectives for the solution, design and development, evaluation, and the communication of results to interviewees. The first three steps were presented to all respondents in an explanatory format, in a reading format, as shown in Table 3, before the questionnaire was applied. Table 4 describes a questionnaire that was applied for which respondents should answer a question as an implementation of step four: design and development.

Figure 2. Design of the methodology applied to the questionnaire (source: the authors).

Table 3. Presentation of the motivation and the identification of the problem, the definition of objectives for the solution, and design and development as an introduction to the application of the questionnaire (source: the authors).

Steps	Explanation
Motivation and problem identification	Creating a loyalty program for the most valuable customers based on cluster segmentation techniques.
Objective definition for the solution	From a set of 15 variables, identify the most important ones from a business perspective.
Design and development	Implementing a questionnaire in which only one question should be answered.

Table 4. Question included in the questionnaire for identifying the most important variables according to respondents (source: the authors).

Questionnaire Question
Please mark with an 'x' which of these variables you consider most important for creating a loyalty program for this company's most valuable customers.

The responses to the questionnaires were then tallied as shown in Table 5. For each x marking the variables judged as most important by the respondents, a point was counted, with the aim of quantifying the total votes. Then, the averages of the votes were calculated. As a result of the questionnaire, the variables included in this study were those that had an average vote greater than or equal to 0.75 or greater than or equal to 3 votes.

Table 5. Results of respondents' answers to the applied questionnaire to identify the most important variables (source: the authors).

Attributes	Interviewee ID				Total of Votes	Average
	1	**2**	**3**	**4**		
date installed	x	x	x	x	4	1
location	x	x	x	x	4	1
contract type	x	x	x	x	4	1
local description		x			1	0.25
payment type		x	x	x	3	0.75
invoice type					0	0
monthly fee		x			1	0.25
installation fee	x	x	x	x	4	1
extensions amount	x	x		x	3	0.75
extension fee	x	x	x		3	0.75
first extension date	x	x	x		3	0.75
last extension date	x				1	0.25
debt amount	x	x	x		3	0.75
time in debt				x	1	0.25
MVC	x	x	x	x	4	1

Table 6 shows the result of the variables that were considered by the study. Based on these variables, it was possible to create new variables with a focus on the customer. As a result of the questionnaire scores, the previously created M attribute (named gross revenue) was altered and calculated from the MVC variable, chosen as one of the most important variables by managers. This variable is an internal monetary indicator which, after eliminating all customer costs, indicates how much, in Euros, a particular customer is pays the company on a recurring basis.

Table 6. Variables with the most votes from the perspective of the main business managers (source: the authors).

Attributes	Total of Votes	Average
date installed	4	1
location	4	1
contract type	4	1
invoice type	3	0.75
installation fee	4	1
extension amount	3	0.75
extension fee	3	0.75
first extension date	3	0.75
debt amount	3	0.75
MVC	4	1

The resulting variables created and included in this study, as well as their descriptions and types, are presented in Table 7. Following feature creation, the k-means clustering technique was initiated, which will be discussed in the subsequent subsection of this paper.

Table 7. Features created based on the voting results of the most important characteristics applied in the questionnaire with the main business managers (source: the authors).

Attributes	Type	Description
Customer ID	Categorical	Unique identification of the customer.
Gross revenue	Numeric	Monetary value spent by the customer since the start of their active life in the company.
Frequency	Integer	Number of times the customer has expanded the services initially contracted.
Recency	Integer	Time in days since the last expansion made by the customer.
Business	Integer	Binary identification of the contractual segmentation of the customer, where 1 represents service establishments and 0 private establishments.
Direct debit	Integer	Binary identification of the segmentation of the method of payment made by the customer, where 1 represents customers who make payment by direct debit and 0 other forms of payment.
Time to first extension	Integer	Time in days from the installation date to the first expansion of initially contracted services.
Quantity of debt	Integer	Number of times the customer was managed by the debt department.
North region	Integer	Binary identification of the region in which customers are located, where 1 represents customers who have an installation in the northern region and 0 represents customers in the southern region of the country.

3.2. K-Means Clustering Technique

Clustering is the process of grouping a set of observations into groups of similar behaviors [23]. A cluster is a set of observations that are similar to each other within the same cluster and different from observations in another cluster. K-means clustering is a simple and popular approach to segmenting a dataset into K distinct clusters, where K is the number of clusters. To perform k-means clustering, the desired number of clusters must be specified, and each observation must be assigned to exactly one of the clusters. The calculation process of the k-means clustering technique is detailed succinctly below.

- **Step 1:** Aiming to find good centroids that represent unique information about the characteristics reported in the dataset, the algorithm starts with an initial estimate of the centroids and alternately determines clusters from the centroid and the centroids from the means of the clusters.
- **Step 2:** Iterate until the cluster assignments stop changing.

5. For each of the K clusters, calculate the centroids of the cluster.
6. Assign each observation to the cluster whose centroid is closest. In this study, the closest is defined via Euclidean distance.

J. du Toit et al. [24] describe the calculation of the k-means algorithm in their study aiming to create customer segments in the energy industry. Razieh et al. [25] not only corroborate the calculation of the same clustering technique but also explain how to determine customer loyalty based on classification after clustering.

3.3. Proposal and Evaluation

The loyalty program is designed to identify the segment of customers that are most valuable to the company. As previously defined and explained, the most valuable customers are those with the highest gross revenue. Therefore, the program aims to benefit the most valuable customers with rewards and privileges to be determined by the company's managers.

Determining the number of clusters is crucial for managing these groups, and increasing this number makes the task more challenging [18]. According to the same authors, the number of customers is smaller in business markets compared to consumer markets, and their behaviors vary less than those of consumer market customers. Therefore, to find the most valuable customer groups that can be better managed by the different departments involved in the program, the authors of this study proposed creating three types of segments of the most valuable customer groups.

The proposed clusters were based on the results of the k-means clustering technique with different numbers of K. According to Abdulhafedh [7], validation can be based on external criteria (evaluating the results in relation to a pre-specified structure) or internal criteria (evaluating the results in relation to information related to the data alone). The current study used both criteria, using (a) the silhouette coefficient as an internal criterion for evaluating the optimal number of K, and using (b) a questionnaire previously applied to respondents with the aim of finding which was the most valuable customer cluster and which was best managed by the departments involved as an external criterion. The latter was applied after the results were found using the silhouette coefficient metric and will be analyzed and discussed in the subsequent section.

There are different metrics for the internal evaluation of clusters. Abdulhafedh [7] used three different metrics to evaluate the results of the clustering techniques applied in his study: the Davies–Bouldin Index, the silhouette coefficient, and the Dunn Index. Dogan et al. [4], on the other hand, used Schwarz's Bayesian Criterion (BIC) as a criterion for evaluating the clusters found. As previously mentioned, the present study used the silhouette coefficient metric for evaluating the clusters.

The silhouette coefficient ranges between −1 (indicative of poor clustering) and +1 (indicative of good clustering), with the goal being maximizing the values. The following equation shows how the silhouette coefficient is measured:

$$S(i) = \frac{b(i) \; - \; a(i)}{max\{a(i),\; b(i)\}} \tag{1}$$

where a(i) is the average distance of point i from all other points in its cluster, and b(i) is the smallest average distance of i from all points in any other cluster. To clarify, b(i) is found by measuring the average distance of i from each point in cluster A, finding the average distance of i from each point in cluster B, and taking the resulting smallest value.

Furthermore, this paper makes a comparison of the results obtained for the RFM variables as part of the process of evaluating the results found in the implementation of k-means clustering. The methodology and application were carried out for the dataset created from the variables indicated as main variables, listed in Table 7, by the interviewees. This comparison will be presented in the Section 4 and discussed later.

4. Results

This section presents the actual implementation of the proposed methodology described in the previous section, using data collected from a private security company. The entire experimental component was carried out using Python 3.8 and several libraries. The pandas and numpy libraries were used for data manipulation and processing; seaborn and matplotlib assisted in data analysis visualization; sklearn was applied for creating data clustering algorithms; and the visualizations and evaluations of the clusters found were performed through the yellowbrick library. By applying the above-described methodology, it is understood that data related to customer behavior are very sensitive and analyzing them using machine learning techniques is vital to understand this behavior. Furthermore, this section is subdivided into three other subsections.

The first subsection is related to the application of k-means clustering to the collected dataset, which was processed and had its respective variables selected as shown in Table 7. The aim of this subsection is to create clusters of the most valuable customers. The second subsection describes the clusters found in the previous subsection. Finally, the last subsection involves the implementation of a questionnaire provided to the interviewees, as presented in Table 2, with the results found in the first subsection. The purpose of the last subsection was to provide decision-making alternatives for the main business managers in the process of managing this group of customers.

4.1. Application of k-Means

In this subsection, the data were first normalized using min–max normalization, as suggested by Visalakshi and Thangavel [26]. The number of clusters was not fixed with the aim of evaluating them, which means that the number of clusters was determined automatically based on the result of the silhouette coefficient.

The cluster results were derived from the k-means algorithm applied to variables generated from the managers' questionnaire responses, as shown in Table 7. The k-means application facilitates the identification of customer groups with similar behaviors. According to Abbasimehret et al. [27], the choice of distance metric between points is a critical decision for achieving high-quality clustering results in this technique.

This study employed the Euclidean distance metric. The number of clusters, K, was selected based on the results obtained through the silhouette coefficient combined with business experience. The result of the silhouette coefficient is shown in Figure 3.

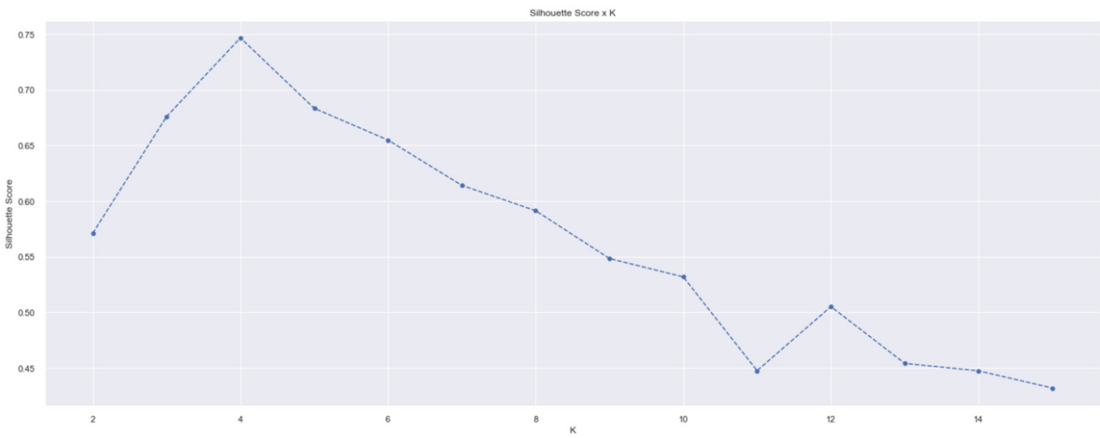

Figure 3. Result of the silhouette coefficient obtained through the implementation of k-means clustering (source: the authors).

According to the result shown in Figure 3, the optimal number of K is 4, corresponding to a result of 0.757 for its silhouette coefficient. In addition to finding the optimal value of K, the current study also aimed to propose new clusters that can be managed more efficiently by the business. Table 8 presents the silhouette coefficient results for the largest K found.

Table 8. Silhouette coefficient values for the 4 largest k values found (source: the authors).

K Value	Silhouette Coefficient
3	0.676
4	0.757
5	0.684
6	0.655

Also, the highest average scores of the silhouette coefficient were found for higher K values, such as K = 5 and K = 6. Although K = 3 showed a high value in the result of its silhouette coefficient, it divides into only three groups. As the objective of k-means clustering is also to propose scenarios of more valuable customer clusters which are manageable by the business, dividing the total customer base into only three groups can make this task more difficult. Therefore, K = 3 was not described and proposed to the business team.

Figure 4 presents the distribution of the cluster results found for each K. As previously mentioned and corroborated with the result shown in the same figure, K = 3 has good clustering, but it will not be relevant to the business. The same happens with K = 4, K = 5, and K = 6, for which their clusters are well divided and defined within each of their groups.

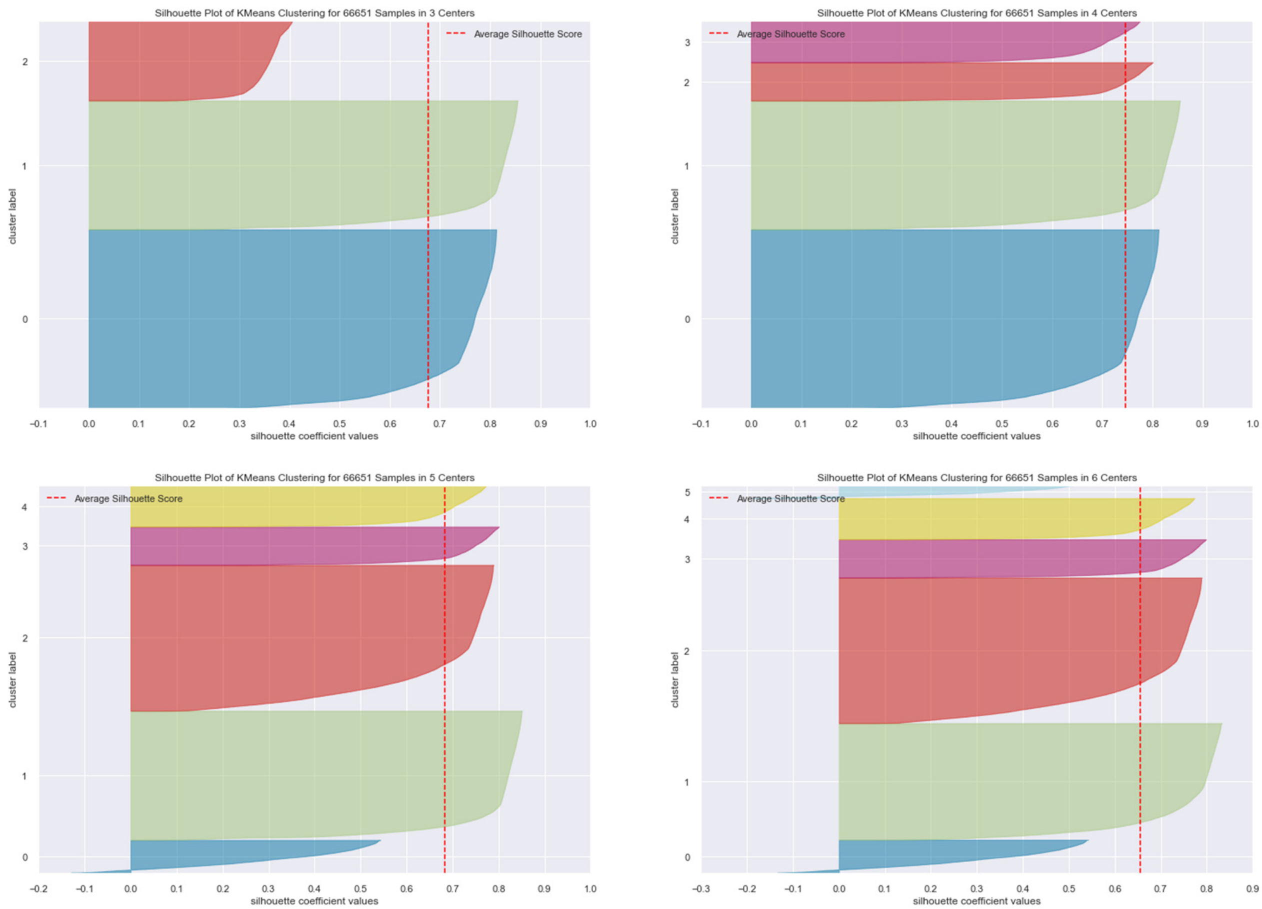

Figure 4. Average silhouette coefficient for each k found and their distributions in each group. (each color represent a different cluster).

4.2. Cluster Descriptions

4.2.1. Clusters for K = 4

This study focused on determining which cluster had the highest average gross revenue to identify the most valuable customer groups. Table 9 displays the results of the clusters using the k-means clustering technique applied to K = 4. Furthermore, the table is sorted in descending order by the gross revenue variable to indicate the cluster with the highest value for this attribute. The gross revenue value is represented by the average amount spent in Euros for each cluster.

Table 9. Number of cases, corresponding percentage of the total customer base, and average gross revenue value in each cluster for k = 4 (source: the authors).

Cluster	Number of Cases Customer Quantity	Percentage of Total Customer Base (%)	Gross Revenue (Euros)
3	6589	9.886%	4540.79
2	22,203	33.312%	4202.22
1	30,775	46.173%	3543.79
4	7084	10.628%	3467.47

In addition to having the highest average gross revenue among the others, Cluster 3 also has the smallest number of customers within the clusters. Furthermore, Cluster 3 has the highest frequency and the second highest recency days (how recent the last expansion was) compared to the other clusters. Table 10 presents all results related to Cluster 3.

Table 10. Results corresponding to Cluster 3 in the implementation of k-means with k = 4 (source: the authors).

Attributes	Aggregation Type	Result
Customer quantity	Absolute Quantity	6589
Percentage of the total customer base	Percentage (%)	9.886%
Gross revenue	Average (Euros)	4540.79
Frequency	Average	2.562
Recency days	Average (days)	1795.81
Time to first extension	Average (days)	1862.01
Number of times in debt	Average	16.24
Percentage of the total customer base that have a business	Percentage (%)	22.88%
Percentage of the total customer base that have direct debit	Percentage (%)	0.00%
Percentage of the total customer base located in the north region	Percentage (%)	11.99%

Some interesting results from this cluster are that despite having the highest value for gross revenue, it contains zero customers with direct debit. This cluster also holds approximately 12% of the total customers located in the northern region and almost 23% of the entire customer base that has a business.

4.2.2. Clusters for K = 5

The clusters obtained through the implementation of k-means clustering for K = 5 are described in Table 11. Unlike the result of the cluster exposed earlier, the most valuable customer group represents 8.421% of the total customer population, and its average gross revenue value is 1.6 times higher than the second most valuable customer group.

Table 11. Number of cases, corresponding percentage of the total customer base, and average gross revenue value in each cluster for k = 5 (source: the authors).

Cluster	Number of Cases Customer Quantity	Percentage of Total Customer Base (%)	Gross Revenue (Euros)
1	5613	8.42%	7153.40
4	6589	9.886%	4540.79
2	22,203	22.312%	4202.22
5	7084	10.628%	3467.47
3	25,162	37.752%	2738.58

Other significant results from this cluster are presented in Table 12. Unlike the most valuable cluster for K = 4, this cluster does not have business-type customers in its composition. However, it represents approximately 11% of the total customers who pay via direct debit and almost 5% of the total customer base located in the northern region. Furthermore, customers in this group have been to debt management on average five times throughout their lifetimes as customers.

Table 12. Results corresponding to Cluster 1 in the implementation of k-means clustering with k = 5 (source: the authors).

Attributes	Aggregation Type	Result
Customer quantity	Absolute Quantity	5613
Percentage of the total customer base	Percentage (%)	8.421%
Gross revenue	Average (euros)	7153.40
Frequency	Average	1.571
Recency days	Average (days)	4157.87
Time to first extension	Average (days)	1862.01
Number of times in debt	Average	5.30
Percentage of the total customer base that have business	Percentage (%)	0.00%
Percentage of the total customer base that have direct debit	Percentage (%)	10.59%
Percentage of the total customer base located in the north region	Percentage (%)	4.94%

4.2.3. Clusters for K = 6

Cluster 6, for k-means of K = 6, is the smallest of all the most valuable customer groups explained earlier. This cluster represents only 3.256% of the entire customer population but has the highest average gross revenue compared to the others already discussed. The Table 13 shows the results.

Table 13. Number of cases, corresponding percentage of the total customer base, and average gross revenue value in each cluster for k = 6 (source: the authors).

Cluster	Number of Cases Customer Quantity	Percentage of Total Customer Base (%)	Gross Revenue (Euros)
6	2170	3.256%	8683.11
1	5657	8.488%	7136.71
4	6589	9.886%	4540.79
2	20,033	30.057%	3716.83
5	7084	10.628%	3467.47
3	25,118	37.686%	2734.61

However, the results presented for the attributes of "recency days" and "time until the first expansion" are higher than those of the other clusters found to be the most valuable. Furthermore, this customer group represents almost 3% of the total customer base located in the northern region, 4.1% of the total customers who pay via direct debit, and approximately 8% of the entire customer base that has a business. Table 14 shows the results corresponding to Cluster 6.

Table 14. Results corresponding to Cluster 6 in the implementation of k-means clustering with k = 6 (source: the authors).

Attributes	Aggregation Type	Result
Customer quantity	Absolute Quantity	2170
Percentage of the total customer base	Percentage (%)	3.256%
Gross revenue	Average (euros)	8683.11
Frequency	Average	1.93
Recency days	Average (days)	3879.90
Time to first extension	Average (days)	5441.38
Number of times in debt	Average	6.29
Percentage of the total customer base that have a business	Percentage (%)	7.54%
Percentage of the total customer base that have direct debit	Percentage (%)	4.10%
Percentage of the total customer base located in the north region	Percentage (%)	2.86%

4.3. Evaluation

In addition to creating a loyalty program for the company's most valuable customers, the main objective of this study was also to provide decision-making alternatives to the main business managers in the process of managing this group of customers. In this way, the company can develop strategies to motivate these customers to stay longer with the company.

A questionnaire was developed for the previously interviewed respondents, as presented in Table 2, t hus incorporating the main results from the k-means implementation for K = 4, K = 5, and K = 6. The purpose of this questionnaire was to engage key stakeholders in the creation of a loyalty program and facilitate decision making in the program's design. Before presenting the results of the identified groups to all respondents, a brief explanation of the silhouette coefficient was provided as a metric for evaluating the most valuable groups found. Table 15 displays the explanation silhouette coefficient concept to interviewees.

Table 15. Silhouette Coefficient metric explanation to the respondents. (source: the authors).

Silhouette Coefficient Concept
The silhouette coefficient is a metric used to evaluate the clusters of customers identified. The coefficient ranges between −1 (indicative of poor clustering) and +1 (indicative of good clustering). The aim is to maximize its value.

Next, the results of the clustering technique were presented to the organization's professionals/decision-makers, hiding the methodology, which was the value of K used. Only a corresponding alias was presented, as shown in Table 16.

Table 16. Summary of the results found in the implementation of k-means clustering for presentation to the interviewees (source: the authors).

K Value	Alias	Customer Quantity	Percentage of Total Customer Base	Gross Revenue (Euros)	Silhouette Coefficient
4	Program 1	6589	9.886%	4540.79	0.757
5	Program 2	5613	8.421%	7153.40	0.684
6	Program 3	2170	3.256%	8683.11	0.655

Along with the results presented in Table 16, the interviewees were encouraged to answer two questions as part of the evaluation process for the created loyalty program. A questionnaire for identifying the best loyalty program is shown in Table 17.

Table 17. Questionnaire for identifying the best loyalty program (source: the authors).

Question Number	Question
1	Based on the results presented, mark with an 'x' which of these programs you would implement as a loyalty program for the company's most valuable customers.
2	Justify your choice to the previous question.

Table 18 presents the result of question 1 of the questionnaire proposed to the group of respondents, as shown in Table 17. Program 3 received the highest score for the best loyalty program, with three votes out of a total of four. The justifications for each respondent's choices are presented in Table 19.

Table 18. Results of the interviewee' answers to the applied questionnaire to identify the best loyalty program for the company's most valuable customers (source: the authors).

Attributes	Interviewee ID				Total of Votes	Average
	1	2	3	4		
Program 1			x		1	0.25
Program 2					0	0
Program 3	x	x		x	3	0.75

Table 19. Justifications for choosing the best loyalty program—question 2 (source: the authors).

Interviewee ID	Summary of Justification
1	"Despite having the lowest evaluation coefficient, but not so far from the best, program 3 presents a smaller customer base and with higher gross revenue. Because it is smaller, we can manage these customers more efficiently".
2	"Program 3 is better because it has the smallest group of customers, it is easier to have excellent management and, not least important, the cost of the program is lower".
3	"Program 1 presents the best silhouette, almost 10 points ahead of the second best. Also, this group holds almost 10% of the total customers. Excellent quantity for a loyalty program".
4	"The cost for program 3 is lower and it is simpler to be managed. Even with a lower coefficient, this is the best program".

Upon analyzing the justifications, some common points were highlighted by most of the individuals surveyed regarding the choice of the best loyalty program. The first was that program 3 has a smaller customer group, which could lead to greater efficiency in managing the customers included in the program. This can be explained by the fact that with fewer customers in the program, a higher level of service can be provided. The second common point was related to the cost of the program, which was mentioned by two of the directors. A loyalty program with a smaller customer group is less financially burdensome for the company.

Finally, after explaining the results and choosing the best loyalty program to be implemented by the company, the respondents were asked to answer a questionnaire to evaluate the usefulness and positive and negative aspects of creating this program. The corresponding questions that guided the interviewees' answers are presented in Table 20.

Table 20. Questionnaire for evaluating the creation of the loyalty program (source: the authors).

Question Number	Question
1	Was the creation of this program useful?
2	Were you already aware of this type of loyalty program for more valuable customers?
3	What are the positive aspects of creating this program?
4	What are the negative aspects of creating this program?

Table 21 shows the answers to questions 1 and 2 presented in Table 20. Despite their experience, some of the interviewees did not know about this type of program for more valuable customers, although they were aware of the concept of a loyalty program.

Table 21. Results of the interviewee' answers to the applied questionnaire to identify the utility and existence of the program (source: the authors).

Interviewee ID	Question 1		Question 2	
	Yes	No	Yes	No
1	x			x
2	x		x	
3	x		x	
4	x			x
Total	4	0	2	2
Average	1	0	0.5	0.5

A synthesis of the answers given to questions 3 and 4 is shown in Table 22. The positive aspects are related to the identification of the most valuable customer group for the company, which had never been carried out before. Another positive point raised was the possibility of creating a marketing plan for this group of customers and benefiting them through the program. Furthermore, as a positive aspect, the possibility of creating a strategic marketing plan targeted at these customers was raised. The only negative aspect was related to the creation and presentation of a second group of more valuable customers, as this would make it possible to create a broader strategic marketing plan.

Table 22. Evaluation of the creation of the loyalty program (source: the authors).

Question Number	Response ID	Summary of Responses
3	3.1	"The program is extremely useful and objective. It is the first time that we will be able to identify and retain customers who bring more value to the business".
	3.2	"Positive aspects of the program: identification of the best customers; loyalty and creation of targeted Marketing".
	3.3	"Identification of the most valuable customers and creation of a specific and excellent Marketing for these customers".
	3.4	"Knowing who are the most valuable and creating benefits for them".
4	4.1	"The second group of most valuable customers should be presented, because in this way we could do Marketing for this group to become the most valuable too".

5. Discussion

As the current literature does not present studies that contemplate the creation of a loyalty program or even customer segmentation in this industry segment, H1 was created. Based on the results presented in this study, Table 16 fully confirms H1. Thus, this study expands recent findings on loyalty programs. The results found are aligned with previous studies which emphasized the creation of a loyalty program in various industry segments [4,10,17–19].

Regarding H2, which is related to the creation of a loyalty program based solely on RFM variables, Figures 5 and 6 present the confirmation of the corresponding hypothesis raised. A study conducted by Dogan et al. [4] in which the variables of the RFM model were applied to the k-means technique to create customer segmentation in a retail industry also validates this hypothesis. However, Figures 5 and 6 also show how unsatisfactory the segmentation of the clusters found is and their respective silhouette coefficient results with the use of only RFM variables.

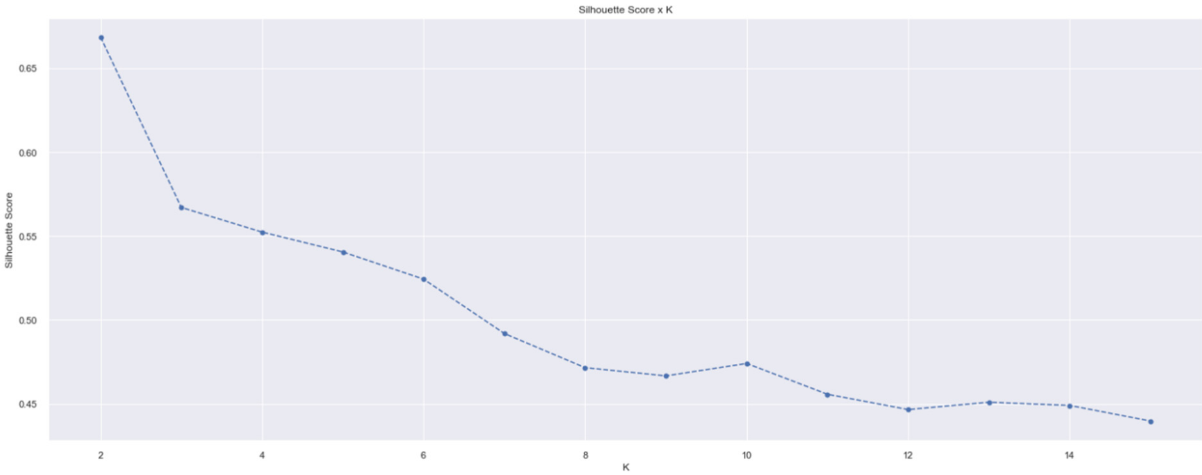

Figure 5. Results of the silhouette coefficient obtained through the implementation of k-means clustering for RFM variables (source: the authors).

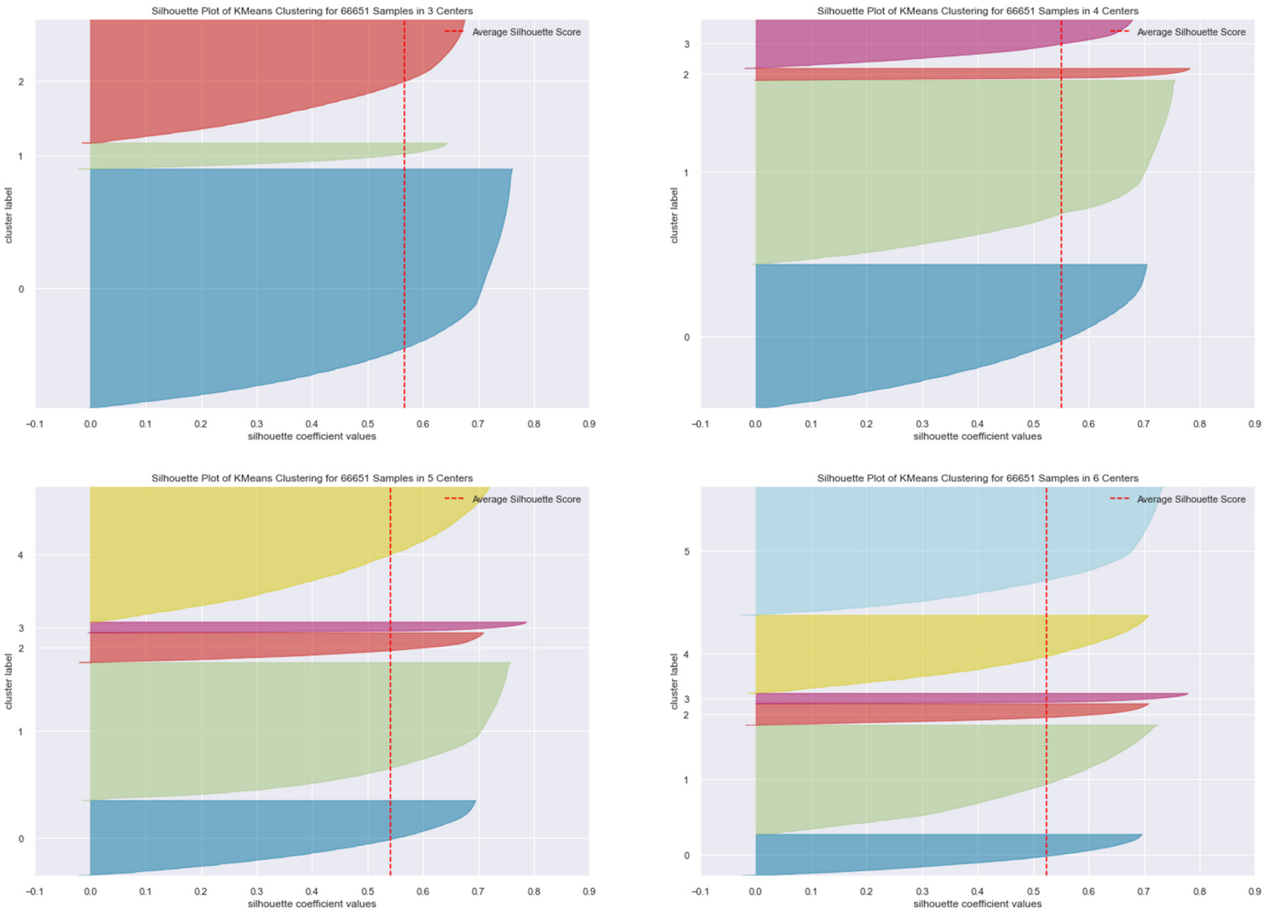

Figure 6. Average silhouette coefficient for each k found and their distributions in each group for RFM variables (source: the authors). (each color represent a different cluster).

The result presented in the previous hypothesis refers to H3, which raises the assumption that variables associated with RFM attributes present better results in the creation of a loyalty program. Table 23 confirms and validates this hypothesis and presents its respective silhouette coefficient values, which are higher than the results found using only RFM variables. Therefore, this study joins the findings of Alkharat et al. [9], in which behavioral, demographic and financial variables were used with the aim of carrying out customer segmentation in the telecommunications industry. Furthermore, the use of more variables in the current clustering model presented an average improvement of 27% in the silhouette coefficient results when compared to the model with only RFM variables.

Table 23. Silhouette coefficient values for the 4 largest k values found in the implementation of k-means clustering with multiple variables and only with RFM variables (source: the authors).

K Value	Silhouette Coefficient with Multiple Variables	Silhouette Coefficient with RFM	Percentage Comparison between the Silhouette Coefficient Multiple Variables and RFM (%)
3	0.676	0.567	19.2%
4	0.757	0.551	37.4%
5	0.684	0.541	26.4%
6	0.655	0.524	25.0%
Total	2.772	2.183	
Average	0.676	0.567	27.2%

However, to the best of our knowledge, this is the first study that proposes to create a loyalty program for the most valuable customers in a private security company. In view of this hypothesis, H4 was suggested, which relates to the creation of marketing strategies for the segmentation of customers found. Table 22 confirms the last-mentioned hypothesis in which the creation of the loyalty program is evaluated. All the positive aspects mentioned refer to the creation of a strategic marketing plan aimed only at these customers. As suggested by Dogan et al. [4], such loyalty programs should carry out some strategic activities for these customers, such as price regulation, promotions, and gift cards, which means increasing the number of contact points with customers.

6. Conclusions

A loyalty program aims to attract and stimulate loyal customers, assuming that the program aggregates the most lucrative customer segment for the company. It is crucial for a company to discern which customers yield the highest profits, ascertain whether lucrative customers exhibit behaviors indicative of loyalty, and thereby determine the appropriateness of implementing a loyalty program as a relational marketing tool.

This paper proposed the creation of a loyalty program for the most valuable customers to the company using machine learning techniques and historical customer information. In addition to quantitative resources, such as the selection of variables for the creation of a base RFM model, this study used qualitative resources; through providing a questionnaire to the company's main business managers, it was possible to add variables to the base model that were judged the most important. Furthermore, different segmentations of the most valuable customers were provided as an alternative to decision-making for the main business managers in the process of managing this group of customers.

Additionally, the proposed model, in conjunction with the selection of variables determined to be the most important for the business, proved to be accurate in identifying the most valuable cluster. In this way, the company can develop strategies to motivate these customers to stay longer with the company. For instance, a company could implement more personalized promotions, prioritizing its most loyal customers as the primary focus for retention efforts. According to Liu [28], owners of loyalty cards spend much money than people without them. This could be a marketing strategy for the company to adopt.

This study contributed in several ways to the existing body of knowledge about the creation of a loyalty program. First, the relevance of the research design to building a loyalty program in conjunction with main stakeholders made a difference because it combined business experience with the proposed methodology. Secondly, it showed that private security companies can benefit from the creation of a loyalty program, as already clarified in the study of the state of the art of this same research area. Specifically, the results show that typical data-based marketing strategies to deliver value from organizational data, such as clustering to segment customers using historic transactional data from customers to build valuable attributes such as RFM, are valid within the specific setting of the private security sector. This is a key contribution stemming from our study given the specificity of the products (e.g., alarms) and services being commercialized by this sector. Another important contribution was to show that despite presenting the best silhouette coefficient metric, the most valuable customer group was not chosen as the ideal cluster for the creation of the program. This finding affirms that customer segmentation not only involves statistical analyses of individual user groups but also requires business understanding and collaboration from stakeholders.

This study has insightful findings and advantages. However, it also has some limitations. Firstly, only one clustering algorithm and one similarity metric were implemented, potentially restricting the generalizability of the conclusions despite the numerous studies associated with the applied algorithm. Secondly, only four interviews were conducted, implying that additional findings may emerge in the future with a larger sample of interviewees. However, it is noteworthy that these four interviews were conducted with professionals who had a genuine interest in the development of this program, which also may come at the cost of some bias. Thirdly, the exploration of the internal variable justifying customer value was restricted due to the need for access to private data and concepts.

In the future, in addition to finding the best customer segmentation approach for creating a loyalty program, the proposed method can be extended to include various clustering algorithms, despite several studies associated with the applied algorithm. Secondly, only four interviews were conducted, which means that more findings can be presented in the future with more interviews. However, these four interviews were conducted with professionals who were genuinely interested in the creation of this program Moreover, a future research direction could be to consider different similarity metrics. This will allow for the most homogeneous clusters possible to consequently draw conclusions that can generalize this industry segment.

Author Contributions: Conceptualization, A.d.S. and S.M.; methodology, A.d.S. and S.M.; software, A.d.S.; validation, A.d.S., S.M. and R.P.; formal analysis, A.d.S.; investigation, A.d.S.; resources, A.d.S.; data curation, A.d.S.; writing—original draft preparation, A.d.S.; writing—review and editing, A.d.S., S.M. and R.P.; visualization, A.d.S.; supervision, S.M. and R.P.; project administration, A.d.S.; funding acquisition, S.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in article.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

RFM	Recency, frequency, and monetary
LTV	Lifetime value
VIP	Very important person
BIC	Bayesian information criterion
MCV	Value cost margin

References

1. Apichottanakul, A.; Goto, M.; Piewthongngam, K.; Pathumnakul, S. Customer Behaviour Analysis Based on Buying-Data Sparsity for Multi-Category Products in Pork Industry: A Hybrid Approach. *Cogent Eng.* **2021**, *8*, 1865598. [CrossRef]
2. Zeithaml, V.A.; Berry, L.L.; Parasuraman, A. The Behavioral Consequences of Service Quality. *J. Mark.* **1996**, *60*, 31–46. [CrossRef]

3. Murugiah, L.; Akgam, H.A. Study of Customer Satisfaction in the Banking Sector in Libya. *J. Econ. Bus. Manag.* **2015**, *3*, 674–677. [CrossRef]
4. Doğan, O.; Ayçin, E.; Bulut, Z.A. Customer Segmentation by Using Rfm Model and Clustering Methods: A Case Study in Retail Industry. *Int. J. Contemp. Econ. Adm. Sci.* **2018**, *8*, 1–19.
5. Zavalishchin, D. *Optimal Loyalty Program Management*; AIP Publishing: Crete, Greece, 2021; p. 040002. [CrossRef]
6. Mrkosová, K.; Dufek, O.; Majer, L. Loyalty Programs as a Part of Company's Marketing Strategy. *Acta Univ. Agric. Silvic. Mendel. Brun.* **2014**, *59*, 199–204. [CrossRef]
7. Abdulhafedh, A. Incorporating K-Means, Hierarchical Clustering and PCA in Customer Segmentation. *J. City Dev.* **2021**, *3*, 12–30. [CrossRef]
8. Fader, P.S.; Hardie, B.G.S.; Lee, K.L. RFM and CLV: Using Iso-Value Curves for Customer Base Analysis. *J. Mark. Res.* **2005**, *42*, 415–430. [CrossRef]
9. Alkhayrat, M.; Aljnidi, M.; Aljoumaa, K. A Comparative Dimensionality Reduction Study in Telecom Customer Segmentation Using Deep Learning and PCA. *J. Big Data* **2020**, *7*, 9. [CrossRef]
10. Mihova, V.; Pavlov, V. *A Customer Segmentation Approach in Commercial Banks*; AIP Publishing: Albena, Bulgaria, 2018; p. 030003. [CrossRef]
11. Mohammadian, M.; Makhani, I. RFM-Based Customer Segmentation as an Elaborative Analytical Tool for Enriching the Creation of Sales and Trade Marketing Strategies. *Int. Acad. J. Account. Financ. Manag.* **2019**, *6*, 102–116. [CrossRef]
12. Leenheer, J.; Bijmolt, T.H.A. Which Retailers Adopt a Loyalty Program? An Empirical Study. *J. Retail. Consum. Serv.* **2008**, *15*, 429–442. [CrossRef]
13. Ha, S.; Stoel, L. Designing Loyalty Programs That Matter to Customers. *Serv. Ind. J.* **2014**, *34*, 495–514. [CrossRef]
14. Melnyk, V.; Bijmolt, T. The Effects of Introducing and Terminating Loyalty Programs. *Eur. J. Mark.* **2015**, *49*, 398–419. [CrossRef]
15. Koo, B.; Yu, J.; Han, H. The Role of Loyalty Programs in Boosting Hotel Guest Loyalty: Impact of Switching Barriers. *Int. J. Hosp. Manag.* **2020**, *84*, 102328. [CrossRef]
16. Moro, S.; Cortez, P.; Rita, P. Using Customer Lifetime Value and Neural Networks to Improve the Prediction of Bank Deposit Subscription in Telemarketing Campaigns. *Neural Comput. Appl.* **2015**, *26*, 131–139. [CrossRef]
17. Sharma, A.; Medudula, M.K.; Patro, S. Marketing Flexibility Interaction Matrix and Consumer Clusters Preference Criteria in Telecommunication Sector. *Glob. J. Flex. Syst. Manag.* **2015**, *16*, 295–307. [CrossRef]
18. Güçdemir, H.; Selim, H. Integrating Multi-Criteria Decision Making and Clustering for Business Customer Segmentation. *Ind. Manag. Data Syst.* **2015**, *115*, 1022–1040. [CrossRef]
19. Tanford, S.; Baloglu, S. Applying the Loyalty Matrix to Evaluate Casino Loyalty Programs. *Cornell Hosp. Q.* **2013**, *54*, 333–346. [CrossRef]
20. Khajvand, M.; Zolfaghar, K.; Ashoori, S.; Alizadeh, S. Estimating Customer Lifetime Value Based on RFM Analysis of Customer Purchase Behavior: Case Study. *Procedia Comput. Sci.* **2011**, *3*, 57–63. [CrossRef]
21. Ansari, A.; Riasi, A. Taxonomy of Marketing Strategies Using Bank Customers' Clustering. *Int. J. Bus. Manag.* **2016**, *11*, 106. [CrossRef]
22. Bell, D.; Deighton, J.; Reinartz, W.J.; Rust, R.T.; Swartz, G. Seven Barriers to Customer Equity Management. *J. Serv. Res.* **2002**, *5*, 77–85. [CrossRef]
23. Shen, B. E-Commerce Customer Segmentation via Unsupervised Machine Learning. In *the 2nd International Conference on Computing and Data Science*; ACM: Stanford, CA, USA, 2021; pp. 1–7. [CrossRef]
24. Toit, J.D.; Davimes, R.; Mohamed, A.; Patel, K.; Nye, J.M. Customer Segmentation Using Unsupervised Learning on Daily Energy Load Profiles. *J. Adv. Inf. Technol.* **2016**, *7*, 69–75. [CrossRef]
25. Qiasi, R.; Baqeri-dehnavi, M.; Minaei-bidgoli, B.; Amooee, G. Developing A Model for Measuring Customers Loyalty And Value With Rfm Technique And Clustering Algorithms. *J. Math. Comput. Sci.* **2012**, *4*, 172–181. [CrossRef]
26. Visalakshi, N.; Thangavel, K. Impact of Normalization in Distributed K-Means Clustering. *Int. J. Soft Comput.* **2009**, *4*, 168–172.
27. Abbasimehr, H.; Bahrini, A. An Analytical Framework Based on the Recency, Frequency, and Monetary Model and Time Series Clustering Techniques for Dynamic Segmentation. *Expert Syst. Appl.* **2022**, *192*, 116373. [CrossRef]
28. Liu, Y. The Long-Term Impact of Loyalty Programs on Consumer Purchase Behavior and Loyalty. *J. Mark.* **2007**, *71*, 19–35. [CrossRef]