IJMSIT

International Journal of Multidisciplinary Studies and Innovative Technologies

e-ISSN: 2602-4888
dergipark.org.tr/en/pub/ijmsit
Research Article

2024, 8 (2), 109-117
DOI: 10.36287/ijmsit.8.2.10
Received: November 6, 2024; Accepted: December 9, 2024

A Sample Strategic Marketing Application: Patient Segmentation and Channel Analysis with The LRM Model

Mustafa Şehirli[1*], and Samet Aydın [2]

[1*]*SHMYO, Managemnt And Organization Department, University of Health Sciences, İstanbul, Türkiye (mustafasafrani@gmail.com) (ORCID: 0000-0002-4800-0283)*

[2]*International Trade and Logistics Management, Maltepe University, Istanbul, Turkiye (sametaydin@maltepe.edu.tr) (ORCID: 0000-0003-2275-4682)*

Abstract – This research aims to develop a new customer segmentation method and to propose strategies for acquiring new customers accordingly. To this end, data from 48,870 patients of a healthcare institution were segmented using the K-Means method. Patients were classified based on their longevity (L), recency (R), and monetary return (M) status and analyzed according to acquisition channels. The findings revealed that the total patients were divided into four distinct clusters. Two clusters containing 2,981 patients, representing 6% of the total, were identified as the ideal segments. While evaluating the clusters, a new indicator based on the Profitability Ratio per Patient was also utilized. The research concluded that the hospital primarily acquires patients through referral channels, with search engines and the website as the second most effective channel. At the same time, social media advertising had a comparatively lesser impact on patient acquisition. Furthermore, it was found that there were no significant differences among customer acquisition channels between the clusters. Recommendations for managers at the end of the study include maintaining more comprehensive customer data, developing profiles for cluster patients for similar sales activities, organizing "refer a friend" campaigns, and conducting evaluations between channel costs and profits.

Keywords *– segmentation, LRFM model, K-means method, clustering, strategic marketing*

Citation: Şehirli, M., Aydın, S. (2024). A Sample Strategic Marketing Application: Patient Segmentation and Channel Analysis with The LRM Model. International Journal of Multidisciplinary Studies and Innovative Technologies, 8(2): 109-117.

I. Introduction

According to Kotler & Armstrong [16], the art of retaining profitable customers in marketing only aims to acquire or retain some customers. According to the authors, a profitable customer is one whose total revenue over time exceeds the costs incurred while acquiring and serving them; in other words, a customer with a positive Customer Lifetime Value (CLV). In this case, it becomes the most rational approach for companies to identify and invest in customers with high CLV. On the other hand, identifying and distinguishing these customers is a highly challenging and important task. Therefore, during the research phase, which is the first step in the marketing management process, sufficient data must be collected, classified, analyzed, and customers grouped. This process is referred to as segmentation. In the strategic marketing phase, one or more of these segments are targeted, and efforts are made to establish long-term profitable relationships using all marketing mix instruments.

This research has been conducted based on existing customer data related to the segmentation phase. Therefore, this study is not a market analysis but a work that focuses on differentiating, understanding, and clustering existing customers, analyzing them according to acquisition channels, and providing recommendations for customer groups that companies should focus on, including potential customers. Similar studies exist in the literature. In this regard, it cannot be claimed that this study is highly original; however, since segmentation studies are, as noted below, highly subjective and often require posterior approaches, and because the results are produced according to the characteristics of the dataset, it can be said that this study has a degree of originality. For instance, the F variable, which shows the frequency of customer acquisition, was not used in the evaluation, and the study analyzed customers based on the channels through which they arrived. Viewed from this angle, this research serves as an example of the subjective segmentation studies that firms should conduct according to their customer profiles and industry characteristics.

Another significant difference in this research is its perspective, which includes potential customers likely to be acquired. The focus is on retaining existing customers and producing practical outputs that facilitate acquiring new customers. Thus, the research serves as an example for the players in the industry being studied and managers and researchers in any sector.

The literature section presents application examples rather than theoretical explanations of the segmentation concept, focusing on the most commonly used methods and providing more detailed information about them. The research is conducted using models and methods (LRFM and K-Means) commonly used in previous studies and with a considerable amount of actual data (approximately 48,870). The research approach is both a priori and posterior, that is, a hybrid approach. In other words, while segmenting customer data, the features in a widely accepted method (LRFM) have been revised, and different evaluations based on the dataset's characteristics have been made.

The research findings have been interpreted to serve as an example for industry managers and researchers.

II. LITERATURE REVIEW

A. Strategic Marketing and Customer Segmentation Concept

In today's increasing competition, companies must focus their interests and resources on specific customer groups to capture customers' attention and establish close, long-term relationships. This approach is one of the most strategic methods in marketing. Selecting one or more customer groups after segmentation is referred to as targeting, and the formulation of brand communication aimed at these target customer groups is known as positioning.

The three-phase strategic marketing process initiated from customer data is illustrated in Figure 1.

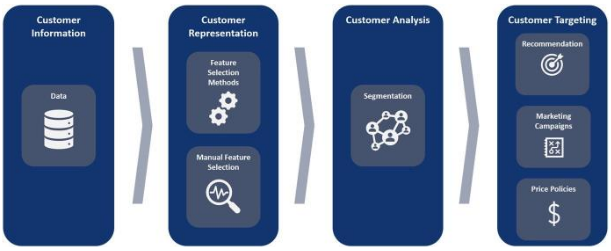

Fig. 1. Strategic Marketing Process Based on Customer Data [1]

This research focuses primarily on segmentation; therefore, the literature review includes information exclusively about this concept. Market segmentation can be defined as the process of dividing a market into customer groups with different needs, characteristics, or behaviours [16]. In other words, segmentation is a strategy for breaking the market into smaller, homogeneous parts to develop products and services tailored to groups with distinct wants and expectations [28].

Through segmentation, it is possible to develop different products for various groups and different marketing strategies and efforts [31]. Segmentation enables companies to focus on a specific subset of consumers that can provide the best service, thereby clarifying the marketing planning process by identifying the needs of particular customer groups for marketing programs [6].

B. Segmentation Criteria and Methods

Various approaches to market segmentation exist, each with its own advantages and disadvantages. The most suitable approach largely depends on the research objective and the type of data available [29]. Generally, analytical marketing strategies, such as data mining, are used to uncover and group customer segments. These strategies utilize demographic, behavioural, and psychographic data to acquire new customers while addressing existing customers' specific needs and wants, enhancing their loyalty [24].

As noted, data is essential for effective segmentation. This data can be categorized into explicit (open) and implicit data. Explicit data refers to customer information, such as demographic details, whereas implicit data pertains to behavioural information, like purchase histories, for accounting purposes. Collecting explicit data can be challenging and potentially misleading due to its constantly changing nature [3]. In contrast, implicit data is generally more accessible and accurate [1].

Despite the growing research and literature on segmentation models, many researchers rely solely on demographic variables to classify consumers, which hinders the discovery of unique patterns, relationships, and latent characteristics [19].

The next phase after obtaining customer data is deciding which customer information will represent the customers and, consequently, be used in segmentation. At this point, there are two approaches: A priori approaches and posterior approaches. The first approach divides the market based on predetermined criteria (demographics, purchasing behaviour, geography, and income). In contrast, the second analyzes the market more thoroughly based on data obtained from the market [11].

In the first approach, the most commonly used models can be summarized as follows [1]:

- RFM (Recency, Frequently., Monetart): Proposed by Hughes [12] , the RFM model is advantageous and widely used because it assesses customer characteristics based solely on three criteria: recency, frequency, and monetary value. It is a powerful and simple model for identifying profitable customers [14] , [27] , [32]. Further details on this model are provided below.
- PCA (Principal Component Analysis): PCA is a dimensionality reduction model that removes features with low information content from consideration.
- PT (Purchase Tree): In this method, customers' products are represented as the leaves of a tree, while product categories are represented as the branches.
- CHAID (Chi-square Automatic Interaction Detector): A technical model based on decision tree techniques.
- CLV (Customer Lifetime Value): A popular metric that focuses on the profit a customer will bring to the company over their lifetime if they remain loyal to the brand.
- DWT (Discrete Wavelet Transform): This method captures location and frequency information.
- GRAPH: Segmentation is based on customer interactions according to location and frequency information.
- MCA (Multiple Correspondence Analysis): MCA allows for a lower-dimensional representation of categorical features.

Alves Gomes and Meisen [1] identified 105 publications analyzing customer behaviour through segmentation methods between 2000 and 2022 and presented a comprehensive study on segmentation techniques. Their research evaluated four stages: data collection, customer representation, customer analysis through segmentation, and customer targeting. According to their findings, customer representation is generally conducted through manual feature selection or RFM analysis. The most commonly used segmentation method is the k-means method, as noted below.

According to Alves Gomes and Meisen [1], RFM has been the most widely used method in the last 20 years (41.9%). There is also a significant amount of research selecting customer data through posterior approaches (47.6%). Figure 2 presents information showing the distribution of methods in this research.

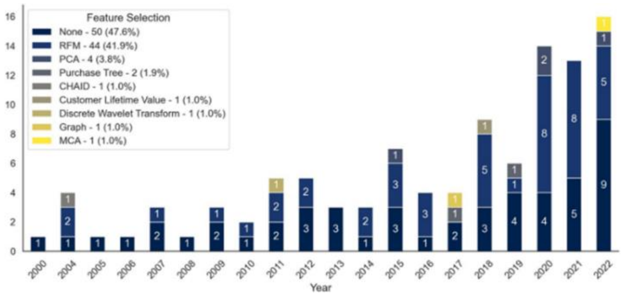

Figure 2: Selection Methods for Customer Representing Features [1]

C. RFM and Its Variants

RFM, as previously defined, is a targeting method that segments customers based on the assumption that newer, more frequent, and higher-spending customers have greater potential. Here, recency refers to the number of days or months since the last purchase; frequency indicates the number of purchases made within a specific period; and monetary value represents the total amount spent [36]. However, RFM has yet to undergo significant improvements over decades. In an actual segment, customers within each segment should respond differently. For example, a customer who has made four purchases may be more likely to make future purchases than a customer with five, which explains why RFM is often considered lower quality compared to methods like CHAID or regression analysis. While RFM is a coarse model based on heuristic perception, advanced methods benefit from statistical frameworks [37].

Indeed, RFM tends to focus excessively on transactional data while overlooking significant differences among customers, such as their values and lifestyles [20]. This can lead to inaccuracies in predicting customer behaviour. Consequently, different parameters have been added to RFM analysis, especially in dynamic sectors like healthcare. For instance, Chang and Tsay [2] introduced length (L), while Yeh et al. [38] added first purchase (T) and customer churn risk (C).

The LRFM model developed by Chang and Tsay [2] segments customers based on the length of their relationship with the company, as illustrated in Figure 3.

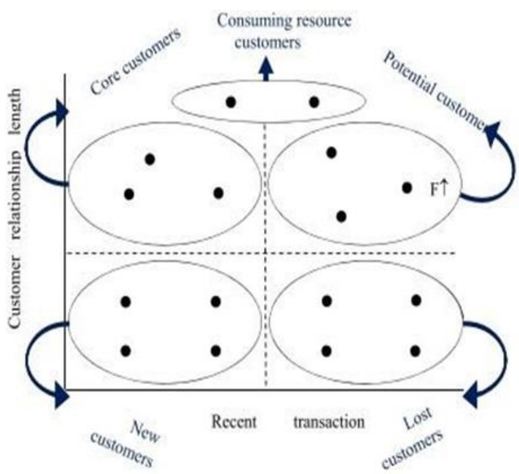

Figure 3: Chang & Tsay [2] LFRM Customer Segmentation Model [18]

As shown in Figure 3, L (length) represents the number of periods from the first purchase to the most recent purchase [36]. Moghaddam et al. [21] introduced the variety (Variety-V) parameter to the model, calling it RFMV, emphasizing the importance of considering product diversity [26].

RFM evaluation is generally conducted using twenty percent segments. For example, the recency (R) score reflects how current a customer's shopping history is; the more recent the purchases, the higher the R-value. Customers in the most current 20% receive a score of 5 points, while those with the oldest purchase dates receive 1 point. Similarly, the frequency (F) and monetary (M) values are also scored based on these twenty percent thresholds [34].

In summary, RFM has advantages such as being simple and comprehensible, offering flexible coding options, and providing insights for predicting future customer behaviour [37].

D. Segmentation Methods

According to the study by Alves Gomes and Meisen [1] , the most commonly used method for segmentation is K-Means, accounting for 39% of the approaches. The following section provides more detailed information about the K-Means method. The distribution of other methods can be found in Figure 4.

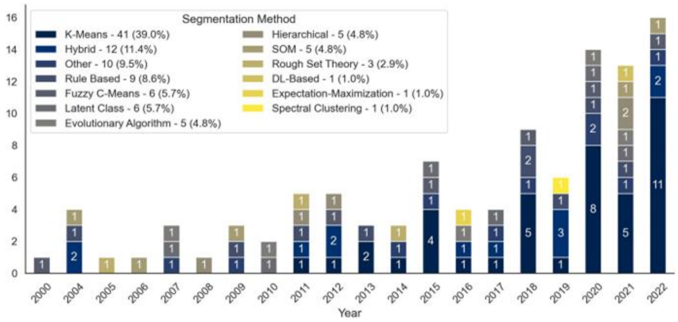

Figure 4: Usage Rates of Segmentation Methods [1]

Outside of the K-Means method, the usage rates for other segmentation approaches—such as time series segmentation, rule-based clustering, Fuzzy C-Means (FCM), Latent Class Analysis (LCA), and Evolutionary Algorithms (EAs)—are quite low, with a total of ten different algorithms showing minimal application. Therefore, it is sufficient to focus solely on the K-Means method:

D.1. K-Means Method

The purpose of the K-Means algorithm is to partition data points into kkk segments, minimizing the distance between the points within each segment. In other words, K-Means clustering is a method aimed at dividing observations into kkk clusters, where each observation belongs to the cluster with the nearest mean. Initially used by MacQueen [17] , the K-Means clustering algorithm is widely applied in various fields such as data mining, statistical data analysis, and other business applications, clustering each observation according to its nearest mean [4]. The Euclidean distance is commonly used in the analysis. However, approximate algorithms such as K-medoids may also be employed due to the NP-hard nature of the underlying optimization problem [17].

The study by Christy et al. [5] observed that the K-Means clustering algorithm consumed less time and reduced the number of iterations compared to Fuzzy C-Means and modified K-Means algorithms. Similarly, in Kanca et al.'s [13] research on 1.9 million unique customer records in the textile sector, the clustering group defined by the fuzzy C-means algorithm (five clusters) provided less in-depth analysis compared to the clustering group produced by the K-Means algorithm (eight clusters).

E. Segmentation in the Healthcare Sector

According to Nnaji et al. [23], data analytics is revolutionizing the healthcare industry and enhancing customer experience and market penetration. Through data analytics, healthcare providers gain deep insights into patient preferences and behaviours, allowing them to develop effective communication strategies [23]. As providers continue adopting data-driven strategies, the healthcare industry is poised for transformative changes that will benefit providers and patients.

According to Torkzadeh et al. [29], while many studies have addressed market segmentation in the healthcare sector using various methods, there has yet to be a consensus on the best approach. Additionally, no technique has been conducted to compare these methods. Therefore, the authors conducted an extensive study, evaluating 22 articles selected from 239 examined to identify the best techniques and criteria. The authors' findings are presented in Figure 5.

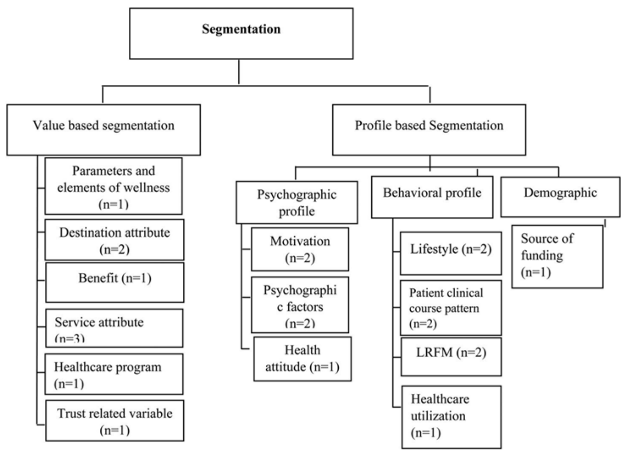

Figure 5: Academic Research on Segmentation in Health Tourism [29]

E.1. Similar Studies

Wei et al. [33] conducted a segmentation study at a dental clinic in Taiwan using the LRFM technique, dividing 2,258 patients into 12 clusters. They developed various suggestions for each cluster, such as offering referral discounts, waiving registration fees, free parking and medical consultations, and giving small gifts alongside medical services. However, they did not present a financial analysis.

In a similar study, Wuu et al. [36] segmented 1,462 patients at a dental clinic into 12 groups. They classified patients using two matrices: customer value and customer relationship. The customer value matrix included treatment frequency and payment amounts, while the customer relationship matrix featured innovation and length variables. Based on these matrices, the authors categorized patients as loyal, those to retain, and those to discard.

Hallab et al. [10] segmented 520 patients in the health sector based on health lifestyle attitudes, which are considered a psychographic variable. They identified this group as significant for health tourism and offered practical recommendations, such as creating smoke-free areas and providing exercise facilities.

Dryglas & Salamaga [8] examined 2,050 tourists at a spa centre in Poland, categorizing them by destination choice motivations (wellness and treatment seekers, treatment seekers, tourism, treatment and wellness seekers). They found that segments differed in socio-demographic, behavioural, and psychographic factors and provided practical suggestions for improving service quality and offering discounts for family packages or during low seasons.

Chen et al. [4] developed a hybrid method using fuzzy logic to support the K-Means algorithm for identifying a hospital's Target Customer Segment (TCS) based on a dataset of 183,947 records.

The results focused on a patient group of 75 individuals and identified 10 key characteristics (Age, Gender, Maximum Monetary Value, Novelty, Frequency, Surgical Chronic Illness, Critical Condition, CT and MRI, and Department) that the hospital should concentrate on. Nakano and Kondo [22] segmented customers based on their entry channel (either online or offline), highlighting differences between the two groups and providing recommendations for managers. Similarly, Konuş et al. [15] proposed a segmentation scheme using Latent-Class Cluster Analysis that considered the communication channel with the company, the stages of information search and purchase, as well as psychographic and demographic individual differences, resulting in three customer segments: multichannel enthusiasts, indifferent shoppers, and store-oriented consumers.

Özkan & Kocakoç [25] segmented patients visiting a children's hospital department using the LRFM method, dividing them into four clusters and offering recommendations for hospital management. Although the fact that the patients were children increased the constraints of the research, practical results were still achieved. Witschel et al. [35] identified that a learned decision tree model was the best descriptive method to capture the essence of clusters following segmentation efforts. Sarıoğlu and İnel [26] conducted three different segmentation analyses—RFM, LRFM, and RFMV—on 228 customers of a biotechnology firm operating under a B2B business model, demonstrating that the RFMV model, which includes the variety (V) parameter, could significantly enhance the firm's customer-centricity and customization capabilities. The authors recommended that different parameters be considered in segmentation studies based on the characteristics of the respective sectors.

III. MATERIALS AND METHOD

This study used data from a dental hospital operating 15 branches in Turkey, covering the period from January 2023 to September 2024. Permissions were obtained by the KVKK (Personal Data Protection Law) to use data from all domestic and international patients served during this period. The data included patient IDs, the relevant branch, the previously explained L, R, and M information, the channel through which the patient arrived at the hospital, and the procedures performed. A total of 48,870 patient records were analyzed. No personal or sensitive data that would create a personal data right was analyzed; patient names and all personal identifiers were coded and stored. Thus, the entire hospital patient profile was segmented..

F. Data Standardization

Prior to analysis, standardization was performed on each dataset. Categorical variables, such as the last visit date of the patient (R - Recency), were converted into numeric values. The first day included in the analysis was assigned a value of

1, and subsequent days were numerically counted in reverse order (with the date one year later coded as 365). This method ensured that patients with the highest R values were those who had visited the hospital most recently.

The Length (L) variable, which shows how long the patient has been receiving services from the hospital, was calculated by subtracting the first visit date from the last visit date. The highest value indicated the most loyal patient. The Monetary (M) variable was standardized to present value due to inflation. To achieve this, the four sales price increase rates (%25, %30, %20, and %20) implemented by the company within two years from the initial registration date were applied backwards to the customer payments using compound interest calculations. Thus, payments made in January 2023 were aligned with those made in September 2024.

G. Data Recording and Analysis

All categorical variables were standardized and recorded as separate variables (Zlength, ZMonetary, Zrecency). However, the Frequency (F) data indicates how often patients visited and was not calculated and excluded from the analysis. The F data could have been more meaningful for the dental hospital context and related more to treatment content outside the patient's control. Therefore, it was left out of the evaluation.

The standardized data were subjected to K-means clustering using the SPSS software. The resulting clusters were analyzed and interpreted based on the information obtained from the hospital, focusing on the channels through which patients arrived and the treatments they received. After the K-means analysis, comments were made using the Pivot Table tool in MS Excel.

IV. RESULTS

During the iterations, the number of individuals in each cluster was monitored. When two clusters were formed, one contained 47,597 individuals, and the other had 1,273. With three clusters, the counts were 46,530 in the first cluster, 2,129 in the second, and 211 in the third. When four clusters were established, the distribution of individuals is shown in Table 1. However, single-member groups emerged when attempting to create five or more clusters, indicating that having more than four would be meaningful.

Table 1. Number of Cases in Each Cluster

Cluster	1	2630.000
	2	22.000
	3	45867.000
	4	351.000
Valid		48870.000
Missing		.000

After conducting 10 iterations for four clusters, the central values for each cluster were obtained, as shown in Table 2.

Table 2. Final Cluster Centers

	Cluster			
	1	2	3	4
L	5	0	3	7
R	314	244	324	314
M	46488	304057	4176	140540

As indicated in Table 3, an ANOVA analysis was conducted to test the significance of the obtained data.

Table 3. Anova Test

	Cluster					
	Mean Square	df	Mean Square	df	F	Sig.
L	6308.230	3	501.817	48866	12.571	.000
R	137646.530	3	35003.238	48866	3.932	.008
M	42172498…	3	49364084.384	48866	85431.420	.000

The results are considered statistically significant since all p-values in Table 3 are less than 0.05. The statistical data for each cluster are presented in Table 4.

Tablo 4. Cluster Data

Cls	**Number Patients**		**L**	**R**	**M (₺)**
1	2630 (%5) Total M: ₺122.263 (%33)	Min	-11	28	26.034
		Max	617	665	98.172
		Av	5	314	46.488
		SD	29	184	17.760
2	22 (%0) Total M: ₺6.689 (%2)	Min	0	31	246.435
		Max	6	630	575.100
		Avg	0	244	304.057
		SD	1	171	69.755
3	45867 (%94) Total M ₺191.540 (%52)	Mi	-16	28	
		Max	833	665	26.023
		Avg	3	324	4.176
		SD	22	187	4.690
4	351 (%1) Total M ₺49.329 (%13)	Min	0	28	98.360
		Max	206	662	240.000
		Avg	7	314	140.540
		SD	29	192	36.470
Av			2,86	323	7.568

When evaluating the clusters, the variable representing the channels through which patients arrived at the hospital (i.e., how patients were acquired) was also considered. For this purpose, Table 5 lists the channels patients came to the hospital.

Table 5. Analysis of All Patients by Arrival Channel

Arrival Channels	**Num**	**%**	**Average M**
Recommendation	21198	43%	₺ 8.050
Google- Website	15654	32%	₺ 6.606
Walk-in Patients	9343	19%	₺ 7.151
WhatsApp	1359	3%	₺ 12.948
Unknown	799	2%	₺ 6.460
Social Med.	306	1%	₺ 10.439
Adv.TV/Mag./N.P	108	0%	₺ 11.442
Earthquake Victim	56	0%	₺ 9.334
Fair	20	0%	₺ 40.673
Overseas Branch	19	0%	₺ 17.299
Commission	8	0%	₺ 13.899
Total/Average	48870	100%	₺ 13.118

This section will present a general distribution table showing the patient acquisition channels for all patients and specific tables indicating the patient acquisition channels for each segment. To avoid a complex presentation, only the highest-valued channels will be included, and channels with less than 1% will not be shown.

Table 6. Patient Acquisition Channels for Cluster 1

Pat. Acq. Channels	Num	%	Avr.M (₺)
Recommendation	1272	48%	46.219
Google Website	683	26%	47.332
Walk-in Patients	469	18%	45.943
WhatsApp	127	5%	47.331
Unknown	32	1%	42.799
Intt - Social Media	25	1%	47.619
Total/Average	**2630**	**100%**	**45.255**

Table 7. Patient Acquisition Channels for Cluster 2

Pat. Acq. Channels	Num.	%	Average M
Walk-in Patients	7	32%	₺ 293.800
Recommendation	7	32%	₺ 332.601
Google Website	5	23%	₺ 280.553
WhatsApp	3	14%	₺ 300.563
Total/Average	22	1	₺ 301.879

Table 8. Patient Acquisition Channels for Cluster 3

Pat. Acq. Channels	Num.	%	Average M
Recommendation	19744	43%	₺ 4.321
Google Website	14884	32%	₺ 3.889
Walk-in Patients	8818	19%	₺ 4.141
WhatsApp	1196	3%	₺ 4.972
Unknown	764	2%	₺ 4.324
Social Media	278	1%	₺ 5.513
Total/Average	**45867**	**100%**	**₺ 5.742**

Table 9. Patient Acquisition Channels for Cluster 4

Pat. Acq. Channels	Num.	%	Average M
Recommendation	175	50%	₺ 138.363
Google Website	82	23%	₺ 143.934
Walk-in Patients	49	14%	₺ 136.537
WhatsApp	33	9%	₺ 143.563
Unknown	3	1%	₺ 162.793
Fair	3	1%	₺ 145.962
Social Media	3	1%	₺ 157.094
Adv.TV/Mag./N.P	2	1%	₺ 183.675
Overseas Branch	1	0%	₺ 120.771
Total/Average	**351**	**100%**	**₺ 148.077**

V. DISCUSSION

The third group, which has the most significant number of members (45,867 people), is the group with the least number of members. In contrast, the second group has the highest average monetary return (304,057 TL). As expected, members in the third group have the lowest average monetary return (4,176 TL). In addition, patients in the second group have the lowest R (recency) and L (loyalty or age) values. This situation can be interpreted as being caused by a subjective factor (high payment amounts). The highest R-value is found in the third group.

Given that the third group represents 94% of the patients, it is normal for this group to have the most up-to-date patients. The highest L value is found in the fourth group, meaning the most loyal patients belong to this group. The second most loyal patient group is the first group. On the other hand, the R values of the first and second groups are the same, placing them in second place regarding recency. In other words, the most loyal and up-to-date patients are found in the fourth and first groups.

When examining which channel patients most commonly came from, it was found that the highest percentage of patients (43%) came via recommendations. This is a critical piece of information. Patients who came through recommendations, either from other patients or from staff, make up the highest proportion in every group. Particularly in the fourth group (most loyal) and the first group (most up-to-date), the percentage of patients coming via recommendation is even higher, at 50% and 48%, respectively. This suggests that the recommendation channel is the most effective regarding patient loyalty and profitability. As expected, the second most common source of patients is the Internet, notably Google and the website. The channel marked as WhatsApp is, in fact, a channel operating on the company's website, so the high effectiveness of Google and the website becomes clearer. The impact of social media appears to be lower. The third group of patients who come through the door (i.e., walk-ins) do not have a channel attribution so that no comments can be made about that. The patients with the highest average return come from fairs (international patients). While the second group consists of patients from overseas branches, the patients from advertisements are also the third group with the highest return. However, the high costs associated with fairs mean whether this channel is the most profitable is a separate issue.

After these evaluations for all patients, conducting separate assessments for each of the four groups would be helpful.

Group 1:

Patients in this group constitute 5% of the total patient population. Their recency and longevity values are higher than average. On the other hand, their financial returns are significantly above average (46,488 TL, compared to the average of 7,568 TL). Although the patients in this group represent 5% of the total patient population, they contribute 33% of the total revenue. The revenue-per-patient ratio (33/5) is 6.6. The distribution of channels through which patients in this group arrived is similar to that of the total patient population, with those arriving via recommendations making up the most significant portion of the group.

Group 2:

This group contains only 22 patients with the highest average revenue (304,057 TL). However, due to the very small number of patients, making meaningful evaluations seems complicated. Indeed, the distribution of arrival channels shows that the percentage of patients coming through the door is equal to that of patients arriving through recommendations in the entire population. This could be considered a random occurrence. Additionally, the patient ratio is extremely small (0.00045), and the revenue-per-patient ratio (0.00045/301,879) is disproportionately high, leading to an unrealistic value (670,842,222). The recency (R) and longevity (L) values for patients in this group are below average. This means these 22 patients are very young and have yet to visit in the most recent period, but their last visit occurred much earlier than average.

Group 3:

This is the largest group, with 45,867 patients (94%). However, the average revenue per patient is only 4,176 TL, roughly one-third of the overall average. The revenue-per-patient ratio is 0.55 (52/94). The distribution of arrival channels for this group is as expected and similar to that of the total patient population. The recency (R) and longevity (L) values are very close to the average, indicating that patients in this group are relatively balanced regarding how recently they have visited and how long they have been patients.

Group 4:

This group contains 351 patients (1% of the total), with an average revenue per patient of 140,540 TL. This is more than ten times the average revenue (13,118 TL). Thus, the revenue-per-patient ratio (13/1) is 13. The channel distribution for patients in this group is similar to the others. These patients represent the most loyal group, with the highest recency values.

Determining the Target Segment (Group)

Based on the above comments and evaluations, the research's purpose will be achieved by identifying the target segment (group).

To identify the hospital's most ideal (profitable and loyal) patient segment, it would be useful to create a separate table that shows the values of L, R, and M, along with the Revenue per Patient/Patient Ratio. This will make it easier to visualize the data. These values are presented in Table 10.

Table 10. Comparison Table for Identifying the Target Segment

Cls.	Number	L	R	M	Rev Pat. Ratio / Patient Ratio.
1	**2630**	**5**	**314**	**46.488**	**6,6**
2	22	0	244	304.057	670.842.222
3	45.867	3	324	4.176	0,55
4	351	**7**	**314**	**140.540**	**13**
Av.		**2,86**	**323**	**7.568**	

The third group, which has the lowest revenue, is also huge, so targeting it would not be appropriate. Similarly, targeting the second group, which is very small, does not make much sense either. After all, the recency and loyalty of the patients in this group are lower than those of the other groups. Therefore, for a more focused approach, targeting patients from Group 4 and Group 1, who have high revenue, are recent, and loyal, would be the most rational approach.

While Group 4 has much higher revenue and loyalty, the small number of patients means that defining their profile would not be statistically significant, and the total revenue potential may be lower. For this reason, Group 1 should also be considered the target segment. Furthermore, the Revenue per Patient/Patient ratio for both groups is significantly higher than that of the largest group.

VI. CONCLUSION

This study serves as an example of the segmentation and targeting stages—a critical phase in strategic marketing—and focuses on segmenting the patients of a healthcare institution and targeting specific segments.

Following the research, as discussed in the literature section, several studies [29], [19] have demonstrated that the K-Means method, which is considered the most suitable method for the healthcare sector, was used for segmentation according to LRFM criteria. The model was revised according to the hospital's dataset and sector-specific dynamics. The model excluded the F-value, and the most suitable segment cluster was identified. Some of the aspects of this study align with and differentiate from other research encountered in the literature search.

Firstly, as many studies have expressed, it is most appropriate for segmentation studies to be done subjectively due to the unique data each brand possesses, which was also confirmed in this study. In fact, the F-value in the LRFM analysis was excluded, making the model more tailored to the firm's specific characteristics.

One of the significant differences between this research and similar studies is the dataset size. For example, studies like Wei et al. [33] and Wuu et al. [36] in the same sector had datasets of 2,258 and 1,462, while this research utilized a dataset of 48,870, significantly increasing the accuracy of the segmentation. Additionally, the number of segments identified in the mentioned studies was much larger, but limiting the segmentation to just four groups in this study made it easier to develop a marketing strategy focused on the target audience.

Another distinction of this research is that an analysis based on patient arrival channels was also conducted. While no similar study exists in the sector, Nakano and Kondo [22] have conducted a comparable analysis. Even though patient arrival channels did not create significant differences across the clusters, this analysis has helped identify essential marketing activities. It was found that referral channels and internet searches are crucial factors for patient acquisition.

Another unique feature of this research is that it includes the revenue per patient/patient ratio as a criterion for selecting the target segment. While this ratio is not the most critical factor, it serves as a comprehensive indicator for evaluation. Accordingly, the clusters with the highest revenue and the most loyal and up-to-date patients were selected as the target segments. Marketing activities directed at these segments are predicted to increase the hospital's long-term profits. However, one of the most necessary activities, patient profiling (persona) studies, could not be conducted due to data limitations. Nevertheless, based on the data obtained in this study, several recommendations can be made for hospital managers.

Recommendations Regarding Patients:

Targeting all patients in Group 1 and Group 4: These patients should be thanked and offered small gifts or discounts.

To improve the referral channel, "Bring a Friend" campaigns should be organized for staff and customers, with promotions offered to those who refer others.

Collect as much customer data as possible (demographic, psychological, sociological) to gain deeper insights.

Detailed analyses of patients' actions should be conducted, and records should be kept up-to-date to improve segmentation.

Patient characteristics in Groups 1 and 4 should be extracted to carry out future "look-alike sales" activities, and marketing budgets should be allocated to target potential customers with similar profiles.

Recommendations Regarding Patient Arrival Channels:

From a digital marketing perspective, the company's resources should be spent more on search engine (Google) searches than social media. The website should continually be updated, and the WhatsApp line should be prioritized.

When the proportion of walk-in patients is high, it is essential to ask them how they heard about the hospital and how they arrived, then define these channels in the system.

Profitability analysis for fairs and international patients should be conducted. While revenue per patient looks pretty high for these channels, especially considering the small number of international patients, ROI (Return on Investment) analysis should be included to evaluate their profitability.

Limitations of the Study and Future Research:

The data collected on patients could be more extensive, meaning the analyses conducted in this study are also constrained. In similar studies, sufficient customer data is assumed to be available, and the task of extracting patient profiles for target segments must be added to such research. Future research could also investigate the costs associated with patient arrival channels. Moreover, analyses of the procedures conducted could be broken down by segment, allowing identification of the most profitable and high-potential procedures.

Authors' Contributions

The authors' contributions to the paper are equal.

Statement of Conflicts of Interest

There is no conflict of interest between the authors.

Statement of Research and Publication Ethics

The authors declare that this study complies with Research and Publication Ethics

References

[1] M. Alves Gomes, & T. Meisen, (2023). "A review of customer segmentation methods for personalized customer targeting in e-commerce use cases", Information Systems and e-Business Management, 21(3), 527-570.M.

[2] H.H., Channg., S.F., Tsay., '' Integrating SOM and K-man in Data Mining Clustering: An Empirical Study of CRM and Profitability Evaluation''. *Journal of Information Management*, 11(4), 161–203 (2004).

[3] Chen X, Fang Y, Yang M, Nie F, Zhao Z, Huang J.Z. Purtreeclust: a clustering algorithm for customer segmentation from massive customer transaction data. IEEE Trans Knowl Data Eng 30(3):559–572. https://doi.org/10.1109/TKDE.2017.2763620 (2018)

[4] Y.S., Chen, C.H., Cheng, C.J., Lai, , C.Y., Hsu, & Syu, H.J. Identifying patients in target customer segments using a two-stage clustering-classification approach: A hospital-based assessment. Computers in biology and medicine, 42(2), 213-221 (2012).

[5] A.J., Christy, A., Umamakeswari, L., Priyatharsini, & A., Neyaa. ''RFM ranking–An effective approach to customer segmentation.'' *Journal of King Saud University-Computer and Information Sciences*, 33(10), 1251–1257, (2021).

[6] S.,Dolnicar, B. Grün, & F., Leisch. Market segmentation analysis: Understanding, doing, and making it useful. Springer Nature.(2018).

[7] S.,Dolnicar, & K., Lazarevski,. Methodological reasons for the theory/practice divide in market segmentation. *Journal of Marketing Management*, 25(3-4), 357–373. https://doi.org/10.1362/026725709X429791, (2009).

[8] D., Dryglas, & M., Salamaga,. Segmentation by push motives in health tourism destinations: A case study of Polish spa resorts. *Journal of Destination Marketing & Management*, 9, 234–246. https://doi.org/10.1016/j.jdmm.2018.01.008, (2018).

[9] O., Ergun, ''Makine Öğrenmesi Algoritmaları İle Müşteri Segmentasyonu Ve Hepsiburada E-Ticaret Platformu Üzerine Bir Uygulama.'' Uludağ Üniversitesi, Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi, (2023).

[10] Z.A., Hallab, Y., Yoon, & M., Uysal. Segmentation based on the healthy-living attitude: A market's travel behaviour. *Journal of Hospitality and Leisure Marketing*, 10(3/4), 185–198. https://doi.org/10.1300/J150v10n03_12 , (2003).

[11] S., Han, Y., Ye, X., Fu, & Z., Chen. Category role aided market segmentation approach to convenience store chain category management. Decision Support Systems, 57, 296–308. https://doi.org/10.1016/j.dss.2013.09.017 , (2014).

[12] A.M., Hughes, A*Strategic database marketing*. Chicago: Probus Publishing Company, (1994).

[13] S., Kanca, T., Özcan, & Y., Çelikbilek. Bir Tekstil Perakendecisinin Müşterileri İçin RFM Modeli ile Müşteri Segmentasyonu. *The Journal of International Scientific Researches*, *8*(3), 393-409, (2023).

[14] U., Kaymak. ''Fuzzy target selection using RFM variables, in Proceedings of the IFSA World Congress and 20th NAFIPS International Conference,'' vol. 2, 2001, pp. 1038–1043, (2001).

[15] U., Konuş, P.C, Verhoef, & S.A., Neslin. Multichannel shopper segments and their covariates. *Journal of Retailing*, 84(4), 398-413, (2008).

[16] P., Kotler & G., Armstrong. Principles of Marketing, 7Th ed., Englewood Cliffs, NJ: Prentice-Hall, (1996).

[17] J.B., MacQueen, Some methods for classification and analysis of multivariate observations, in: Proceedings of Fifth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1967, pp. 281–297, (1967).

[18] F., Marisa, S.S.S., Ahmad, Z.I.M., Yusof, F., Hunaini and T.M.A., Aziz. Segmentation model of customer lifetime value in small and medium enterprise (SMEs) using K-means clustering and LRFM model. *International Journal of Integrated Engineering*, *11*(3), (2019).

[19] G., McKernan. Customer Segmentatıon Approaches:'' A Comparıson Of Methods Wıth Data From The Medıcare Health Outcomes Survey '' (Doctoral Dissertation, University Of Pittsburgh), (2018).

[20] J.A., McCarty, & M., Hastak. Segmentation approaches in data-mining: A comparison of RFM, CHAID, and logistic regression. *Journal of Business Research*, 60(6), 656–662, (2007).

[21] A., Moghaddam, & Harandi. ''An RFMV Model and Customer Segmentation Based on Variety of Products,'' 155-161, (2017).

[22] S., Nakano, & F.N., Kondo. Customer segmentation with purchase channels and media touchpoints using single source panel data. *Journal of Retailing and Consumer Services*, 41, 142-152, (2018).

[23] U.O., Nnaji, L.B., Benjamin, N.L., Eyo-Udo, & E.A., Etukudoh. A review of strategic decision-making in marketing through big data and analytics. Magna Scientia Advanced Research and Reviews, 11(1), 084-091, (2024).

[24] N.T., Nwosu, S.O., Babatunde, & T., Ijomah. Enhancing customer experience and market penetration through advanced data analytics in the health industry. *World Journal of Advanced Research and Reviews*, *22*(3), 1157-1170, (2024).

[25] P., Özkan, & İ.D., Kocakoç. Sağlık sektöründe LRFM analizi ile pazar bölümlendirme. *Kuşadası: Türkiye*, (2019).

[26] E., Sarıoğlu, & M., İnel. Müşteri Segmentasyon Modellerinin Karşılaştırılması Üzerine Ampirik Bir Araştırma. *Öneri Dergisi*, *19*(62), 130-145, (2024).

[27] J.M.C., Schijns, G.J., Schroder. Segment selection by relationship strength, *J. Direct Mark*. 10 (3) (1996) 69–79.), (1996).

[28] T.T., Shi, X.R., Liu, & J.J., Li. Market segmentation by travel motivations under a transforming economy: Evidence from the Monte Carlo of the Orient. Sustainability, 10(10), 3395. https://doi.org/10.3390/su10103395 , (2018).

[29] L., Torkzadeh, H., Jalilian, M., Zolfagharian, H., Torkzadeh, M., Bakhshi, & R., Khodayari-Zarnaq. Market segmentation in the health tourism industry: ''A systematic review of approach and criteria.'' *Journal of Policy Research in Tourism, Leisure and Events*, 16(1), 69-88, (2024).

[30] Y.C., Tsao, P.V.R.P., Raj, & V., Yu. Product substitution in different weights and brands considering customer segmentation and panic buying behaviour. *Industrial Marketing Management*, *77*, 209–220, (2019).

[31] M., Wedel, W., Kamakura.: Market Segmentation: Conceptual and Methodological Foundations. Kluwer Academic Publishers, Boston, 2nd edn. (2000)

[32] J.T., Wei, S.Y., Lin, & H.H., Wu. A review of the application of RFM model. *African journal of business management*, *4*(19), 4199, (2010).

[33] J.T., Wei, S.Y., Lin, C.C., Weng ve H.H., Wu, : "A Case Study of Applying LRFM Model in Market Segmentation of A Children's Dental Clinic,". Expert Systems With Applications, 39(5), 5529-5533, (2012).

[34] J.-T.,Wei, M.-C., Lee, H.-K., Chen, ve H.-H., Wu. ''Customer Relationship Management in the Hairdressing Industry: An Application of Data Mining Techniques. Expert Systems with Applications,'' 2013, 40(7), 7513- 7518, (2013).

[35] H.F., Witschel, S., Loo, & K., Riesen. How to support customer segmentation with useful cluster descriptions. In Advances in Data

Mining: Applications and Theoretical Aspects: 15th Industrial Conference, ICDM 2015, Hamburg, Germany, July 11-24, 2015, Proceedings 15 (pp. 17-31). Springer International Publishing, (2015).

[36] H.-H., Wu, S.-Y., Lin, & Liu, C.-W., Liu. Analyzing patients' values by applying cluster analysis and LRFM model in a pediatric dental clinic in Taiwan. *The Scientific World Journal*, 2014. https://doi.org/10.1155/2014/685495 , (2014).

[37] A.X., Yang. How to develop new approaches to RFM segmentation. *Journal of Targeting, Measurement and Analysis for Marketing*, *13*, 50–60, (2004).

[38] I.C., Yeh, K.J., Yang ve T.M., Ting: "Knowledge Discovery on RFM Model Using Bernoulli Sequence". Expert Systems with Applications, 36:5866–5871, (2008).